Adaptation of the GreenLab growth model to plants with complex architectures and multi-scale analysis of source-sink relationships for parametric identification.

Veronique Letort

▶ **To cite this version:**

Veronique Letort. Adaptation of the GreenLab growth model to plants with complex architectures and multi-scale analysis of source-sink relationships for parametric identification.. Vegetal Biology. Ecole Centrale Paris, 2008. English. NNT : . tel-00365172

HAL Id: tel-00365172
https://theses.hal.science/tel-00365172

Submitted on 2 Mar 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.

ÉCOLE CENTRALE DES ARTS ET MANUFACTURES «ÉCOLE CENTRALE PARIS »

THÈSE

présentée par: Véronique LETORT
pour l'obtention du GRADE DE DOCTEUR

Spécialité: Mathématiques Appliquées

Laboratoire d'accueil: Mathématiques Appliquées aux Systèmes (MAS)

Adaptation du modèle de croissance GreenLab aux plantes à architecture complexe et analyse multi-échelle des relations source-puits pour l'identification paramétrique.

Adaptation of the GreenLab growth model to plants with complex architectures and multi-scale analysis of source-sink relationships for parametric identification.

Date de soutenance : 28/05/2008

Devant un jury composé de :

M. Winfried KURTH	University of Technology, Cottbus	(Rapporteur)
M. Thierry FOURCAUD	CIRAD Montpellier	(Rapporteur)
M. Thiéry CONSTANT	INRA Nancy	
M. Paul-Henry COURNÈDE	Ecole Centrale Paris	
M. Yan GUO	China Agricultural University	
M. Ep HEUVELINK	Wageningen University	
M. Jérémie LECOEUR	SupAgro Montpellier	(Président du jury)
M. Philippe de REFFYE	CIRAD-INRIA	(Directeur de thèse)

N° ordre : 2008 - 24

Remerciements

Grands mercis à :

Thanks very much to:

谢谢大家!

- Philippe de Reffye 老师, pour les liens scientifiques et amicaux mis en place durant ces années, pour m'avoir fait partager sa passion des plantes et de leur modélisation selon St GreenLab, pour toutes les connaissances transmises, incluant les proverbes africains et chinois, et pour les bons moments passés au cours de cette thèse: à travailler face à l'écran d'ordinateur ou bien à la machine à café, à se promener dans le vent froid du parc du Palais d'été (idéal pour neutraliser le décalage horaire), dans l'avion (surtout surclassés en classe affaires!), pour l'initiation à la musique classique, point sur lequel je ne me suis pas montrée très bonne élève...

- Paul-Henry Cournède pour son enthousiasme et son implication dans le travail présenté ici, pour ses conseils scientifiques et stratégiques, pour m'avoir initiée et associée au développement de notre cher DigiPlante, pour m'avoir guidée sur les parties maths applis de la thèse, pour les nombreuses relectures d'articles et de chapitres de thèse, même lorsque je les lui soumettais très en retard, pour l'aide dans les moments difficiles (habituellement les veilles de dead-lines !) et le bon repas préparé Océane !

- Au directeur du laboratoire MAS, Christian Saguez et au directeur de la recherche de Centrale, Jean-Hubert Schmidt pour m'avoir fourni les moyens financiers et matériels de poursuivre cette thèse

- Amélie Mathieu pour les nombreux conseils donnés en tant que "soeur aînée" (au sens chinois du terme !), les échanges sur les (très rares !) bugs de Digiplante, les relectures de chapitres , les discussions comparatives des fittings arabido-colza et autres, et les pique-niques au parc de Sceaux

- Sylvie Dervin pour son aide précieuse dans les méandres administratifs, ses dessins et ses messages sympathiques au tableau, ses entrées-surprises dans notre bureau et aussi ses éternuements divertissants dès qu'un chat montre le bout d'un poil

- Angélique Christophe pour le travail accompli sur notre chère Arabido, pour son accueil et les discussions pendant la (longue!) rédaction de l'article

- Jérémie Lecoeur pour toutes les idées et les discussions intéressantes, ainsi que pour avoir accepté de présider ma soutenance de thèse

- Winfried Kurth and Thierry Fourcaud, for their careful reading of this thesis manuscript as reviewers and for their valuable comments

- Ep Heuvelink, for his welcoming in Wageningen, for friendly and scientific discussions, and for taking part of the PhD defense as member of the committee

- Thiéry Constant pour sa participation à mon jury de thèse, ainsi que Gérard Nepveu et toute l'équipe du Lerfob (notamment Manu, Claude et Alain) pour leur accueil au labo et leur aide pour l'étude du hêtre

- Tous les amis d'AMAP, notamment Daniel Barthélémy, Sylvie Sabatier, Marc Jeager, Jef Barczi, Hervé Rey, Thierry Fourcaud, Alexia Stokes, Yves Caraglio, François de Coligny et Daniel Auclair

- ainsi que Camilo Zalamea et Patrick Heuret pour m'avoir permis de découvrir la forêt guyanaise et la plus merveilleuse de toutes les plantes, Cecropia

- Valentin, Susi, Toundé et Valentino, la valeureuse équipe Cecropia du module FTH2007, qui a bravé la chaleur, les serpents, les moustiques, les pous d'agouti et autres bestioles toutes plus affectueuses les unes que les autres pour collecter de superbes jeux de données sur tous les arbres choisis, même ceux qui n'avaient pas forcément "souper bien poussé"

- Kang MengZhen 师姐 and Ma YunTao 师姐 for their kind welcoming, for friendship and scientific joint work

- Guo Hong for friendship and for the intensive work on our young pines and on the huge set of data on adult pine, Hong LingXia for her kindness and for taking care of me each time I went to CAF, and Pr Lei

- All people from LIAMA, and particularly Pr. Hu BaoGang and Hong BiZhen for welcoming me in LIAMA and for their help in many occasions

- Chinese friends and colleagues from CAU: Pr. Guo Yan, Zheng BangYou, Ma Yuntao, Zhan ZhiGang, Li Dong, Wang Feng

- and Pr Zhang BaoGui, pour son acceuil à tous mes séjours à Pékin et pour nos nombreuses discussions amicales et scientifiques (les mystérieux zigzags de la tomate...!)

- Tous les membres du labo MAS et en particulier Qi Rui, Marlène, Sylvain, Vincent (les meilleurs co-bureaux "comiques" du monde), Marc, Aurélie, Cédric,

Takuya (l'ordre d'énumération n'a aucune vocation à représenter une hiérarchie quelconque dans le bureau des personnes concernées), Xing Le, Julien et XiuJuan pour les règlements mis en place pour veiller au bon déroulement des pauses (horaires, durée, chemin à emprunter selon la météo pour se rendre au bat'ens...) et les conviviales soirées pizza-jungle-speed-loup-garou-elixir.

- Anne-Laure pour ses relectures, ses conseils avisés sur la rédaction de la thèse (cf l'annexe C!) et ses bons gâteaux
- aux amis de longue ou très longue date, en particulier Gwénola, Cédric, Jeanne, Michaël, Laetitia, Jaya, Slaweck, Lucie
- toute ma famille ; en particulier mes parents, Steph, Jean, Gaëlle et Damien pour leur soutien matériel et moral pendant les trois ans, les références biblio tapées et toutes les ondes positives envoyées pendant la soutenance (sur des distances allant de deux mètres à 15 000 km...)

«Les plantes invitent l'homme à des recherches mathématiques. »

[Jean, R. V., Phytomathématique, Montréal, 1978]

«La vérité ce n'est pas ce qui se démontre, c'est ce qui simplifie. »

[Antoine de Saint-Exupéry, Terre des Hommes, 1939]

«Sans théorie il n'y a pas de mesures valables.»
«L'essentiel de la plante est invisible pour les yeux. On ne voit bien qu'à travers les modèles. »

[Phil LaoShi, extraits choisis, 2008]

Contents

Introduction

0.1 Introduction (version française)

This introduction is written in French as required by ECP graduate school administration. A translation follows.

0.1.1 La plante, un système complexe ?

De manière générale, une plante ou un peuplement végétal peuvent être vus comme un système complexe, au sens mathématique du terme, c'est-à-dire comme un système constitué d'un grand nombre d'entités hétérogènes dans lequel des interactions fortes créent différents niveaux d'organisation et induisent un comportement holistique (*i.e.* dont certaines propriétés émergentes ne peuvent être déduites de l'étude indépendante des sous-sytèmes). Dès lors, l'approche réductioniste prônée par exemple par Descartes dans son "*Discours de la Méthode*", c'est-dire la subdivision d'un problème en sous-problèmes indépendants, ne peut pas être appliquée : il est nécessaire d'analyser le système dans son ensemble [Ricard, 2003].
Une caractéristique des systèmes complexes est que leur étude nécessite de manière intrinsèque une approche interdisciplinaire et multi-échelle. L'interdisciplinarité est liée au fait que les questions posées et les méthodes développées pour y répondre sont issues de différents domaines scientifiques. Elle permet ainsi de confronter les développements théoriques aux problèmes spécifiques posés par les applications aux plantes réelles. D'autre part, la complexité inhérente aux systèmes étudiés requiert une analyse à plusieurs échelles d'espace et de temps. La théorie des hiérarchies décrit en particulier des systèmes pouvant être représentés comme des empilements de sous-systèmes emboités ("nested hierarchy"). Cette définition peut s'appliquer aux plantes, pour lesquelles les différents niveaux peuvent être par exemple les organes, les phytomères, les unités de croissance, les axes, les complexes branchés, la plante toute entière [Godin and Caraglio, 1998]. Le comportement des structures de chaque niveau conduit à l'émergence de comportements collectifs et d'organisations à des niveaux supérieurs. Les interactions entre les différents niveaux sont donc vues comme des contraintes ou des propriétés émergentes selon le point de vue choisi.
Dans de nombreux cas, les systèmes complexes font preuve d'une grande robustesse à

des perturbations même à large échelle ou multi-dimensionnelles et montrent une capacité inhérente à s'adapter ou bien à persister dans une situation stable. Perry [1995] caractérise ces systèmes de "critiques" ou "métastables" : robustes pour des perturbations auxquelles les composants du système sont adaptés mais sujets à de brusques changements si certains seuils sont dépassés. Cela se vérifie pour les plantes qui exhibent des propriétés macroscopiques une échelle n relativement robustes (qui sont la base des modélisations s'appuyant sur les principes "d'équilibre fonctionnel"[Mäkelä, 2003]) même si à l'échelle $n-1$ les processus physiologiques mis en jeu sont complexes et parfois très différents selon les conditions de croissance. Si les niveaux sont suffisamment isolés, le développement d'un modèle "mécaniste" peut reposer sur l'étude du comportement des constituants du système à l'échelle inférieure uniquement : inclure un plus grand nombre de niveaux différents diminue la robustesse et la stabilité du modèle [Mäkelä, 2003].

Ces théories se sont en grande partie développées pour des applications en écologie ou à l'échelle du paysage [Li, 2000]. Cependant Mäkelä [2003] souligne que les pratiques sylviculturales tendent à rapprocher la plantation exploitée d'un peuplement naturel, c'est-à-dire un peuplement plus hétérogène en termes d'espèces présentes, de classes d'âges et de disposition spatiale. L'auteur en tire la conclusion que l'on doit considérer «l'écosystème forestier comme une entité hiérarchisée consituée de niveaux d'organisation à différentes échelles spatiales et temporelles ». De manière plus modeste, nous soulevons ici la question des changements d'échelle spatiale au niveau de la plante seule. En ce qui concerne leur architecture, la description des plantes comme des systèmes modulaires mettant en jeu des phénomènes multi-échelles est couramment utilisée [Barthélémy and Caraglio, 2007]. Cela a naturellement guidé les premières représentations de la forme des plantes vers des méthodes basées sur la notion d'autosimilarité comme la théorie des fractales. Même si ce type de représentation a vite révélé ses limites, notamment du point de vue de la souplesse de représentation, c'est selon les mêmes principes qu'ont été développés les concepts de modèle architectural et d'unité architecturale. D'un point de vue architectural, la notion d'analyse multi-échelle est donc une notion bien étudiée et pour laquelle un formalisme et des outils adaptés ont été développés [Godin and Caraglio, 1998]. Reste à intégrer le fonctionnement dans ce cadre, c'est-à-dire effectuer une analyse multi-échelle d'un modèle structure-fonction.

0.1.2 Complexité du modèle et identification paramétrique : un équilibre à trouver.

Comme pour d'autres systèmes complexes, la confrontation aux observations réelles pose des questions liées à la reconstruction de la dynamique du système à partir des données, souvent sources de problèmes inverses difficiles. Cela soulève également le problème de l'acquisition de données et de la définition d'un niveau pertinent de mesures

et des variables associées. Nous allons donc chercher les compromis possibles entre la complexité des processus à décrire et les simplifications apportées par la modélisation.

Nous présentons dans cette thèse des travaux appliqués au modèle de croissance des plantes GreenLab, initié par Philippe de Reffye. Le premier chapitre de la thèse en synthétise les principales caractéristiques en les situant par rapport aux méthodes communément employées dans d'autres modèles de la bibliographie. Dans ce chapitre, nous nous sommes principalement concentrés sur les modèles de croissance pour les arbres, qui sont l'objetif premier de notre étude, même si certaines problématiques spécifiques aux herbacées sont aussi abordées. Rappelons ici de manière brève les principales composantes de GreenLab.

GreenLab est un modèle structure-fonction qui simule les processus de production de biomasse et sa répartition dans la plante entière, à l'échelle de l'organe. L'organogénèse est simulée par un automate double-échelle qui détermine les règles de branchaison entre les axes, qui sont classés en différentes catégories selon leur stade de différentiation (âge physiologique). Cet automate se décline en 3 versions selon le type de plante à décrire et l'objectif visé : déterministe ($GL1$), stochastique ($GL2$) ou bien avec rétroaction de la photosynthèse sur l'organogénèse ($GL3$). La production de biomasse est calculée à chaque pas de temps selon la surface foliaire active par une fonction qui intègre la compétition lumineuse à l'échelle de la plante (selon un formalisme de type loi de Beer-Lambert). Cette biomasse est répartie entre les organes en expansion selon leur demande, indépendamment de leur position ("pool" commun de biomasse), suivant un modèle de compétition de puits.

Les travaux d'Amélie Mathieu [Mathieu, 2006] ont permis des progrès dans le domaine de la simulation de plantes ramifiés à l'aide du modèle GreenLab. L'arbre est simulé comme un système capable de s'auto-réguler, selon les conditions dans lesquelles se déroule sa croissance. On dispose ainsi d'un outil efficace pour représenter la plasticité du développement topologique des arbres selon les conditions environnementales et selon la compétition qu'ils subissent de la part de leurs voisins. On peut également simuler de manière plus réaliste les différents stades de croissance dans la vie d'un arbre, avec notamment l'apparition progressive des ordres de branchements ("effet de base"). Le modèle permet également de simuler des phénomènes cycliques observés chez certains arbres ou herbacées comme l'alternance de périodes de fructification et de branchaison [Mathieu et al., 2008].

L'un des objectifs majeurs de GreenLab est la prédiction et l'optimisation de la croissance des plantes. Une étape importante est donc l'ajustement du modèle sur des plantes réelles. Cela repose sur des ajustements en parallèle (*multi-fitting*) des masses de chacun des organes et permet de retracer la dynamique de la répartition de biomasse au cours de la vie de la plante et les seuils de branchaison et/ou fructification en fonction du rapport de l'offre de biomasse sur la demande. Les précédents travaux réalisés autour de GreenLab ont permis son identification paramétrique pour plusieurs plantes

agronomiques ou herbacées, dont par exemple :
– le blé [Zhan et al., 2000], [Kang et al., 2007b]
– le tournesol [Guo et al., 2003]
– le maïs [Guo et al., 2006] [Ma et al., 2007]
– la tomate [Dong, 2006] [Dong et al., 2008]
– le chrysanthème [Kang et al., 2006b]
– le riz [Mathieu, 2006]
– le caféier [Mathieu, 2006]
– le jeune hêtre [Mathieu, 2006]

Un des enjeux cruciaux de ces études, en plus de la calibration du modèle en tant que telle, est l'étude de la stabilité des paramètres : il est important de pouvoir déterminer quel est le domaine de validité du modèle, sa robustesse et son pouvoir prédictif. Ma et al. [2007] ont analysé la variabilité des paramètres de GreenLab pour le maïs en réponse à trois types de variabilité phénotypique : au sein d'une même population, au cours de ses stades de développement et sous différentes conditions climatiques. Le modèle n'explique pas la variabilité des plantes au sein d'une même population, puisque la version utilisée était déterministe, avec un jeu de paramètres supposé spécifique à l'espèce. En revanche une bonne stabilité des paramètres entre les stades de croissance a été constatée. Concernant les conditions de croissance, le coefficient de variation des mesures de biomasse atteint 18% alors que la variation correspondante des paramètres est de 10%. Cela signifie que les deux facteurs environnementaux choisis (ETP pour la production de biomasse et température pour le développment) contrôlent la majeure partie des variations phénotypiques dues aux variations environnementales. En revanche, des changements significatifs des valeurs de certains paramètres ont été obtenus par Ma et al. [2008] (sur le maïs) et Louarn et al. [2007] (sur la tomate) pour des plantes croissant au sein de plantations de différentes densités. Cela montre que, si les résultats restent fiables dans un intervalle limité de conditions environnementales, des améliorations restent à apporter pour relier les variations des paramètres aux variations environnementales influentes.

Les applications citées ci-dessus concernent des plantes dont les ordres de ramifications restent faibles (au maximum 3 pour les talles du riz). Pour ce type de plantes, une description détaillée et une prise de mesures organe par organe sont possibles. En revanche, cela devient irréalisable lorsque les ordres de ramifications sont plus élevés. La construction d'un fichier-cible [1] complet imposerait de peser et mesurer séparément toutes les feuilles, tous les entre-noeuds et tous les fruits de la plante. S'il s'agit d'un arbre âgé et avec une architecture complexe, ce travail devient rapidement beaucoup trop fastidieux.
Ainsi, même si la simulation est performante, nous sommes limités dans la pratique par

[1] On appelle fichier-cible le fichier regroupant l'ensemble des données récoltées sur la plante et utilisées comme cible pour l'identification paramétrique

le travail de récolte de données. Les informations réelles auxquelles nous avons accès sont partielles ou plus grossières que celles que nous pouvons simuler. Cela implique une analyse de l'effet de l'agrégation de données ainsi qu'un travail de simplification du modèle, de manière à adapter le niveau de détail de la simulation à celui que l'on peut espérer obtenir en pratique à partir des mesures. C'est l'objectif du travail présenté ici.
La question à laquelle nous cherchons à répondre dans cette thèse est la suivante :
Comment appréhender la prise de mesure et l'identification paramétrique pour des plantes branchées ayant un modèle architectural complexe ?
On peut pour traiter cette question distinguer les sous-problèmes suivants :

- Une description détaillée du système architectural complet n'étant pas accessible, la plante simulée aura une architecture différente de celle de la plante réelle : quel est l'effet d'un changement de structure topologique sur la production et l'allocation de biomasse dans le modèle ?
- Pour un ajustement du modèle sur une plante réelle, quelles données concernant sa topologie peuvent être incluses dans la cible ?
- Si aucune donnée topologique n'est disponible, est-il utile de simuler le développement topologique et dans ce cas, comment fixer les règles d'une architecture simplifiée par défaut ?
- Quelles sont les équations et les paramètres régissant le comportement d'un modèle simplifié qui conserverait des équivalences avec le modèle complet ? Quel choix peut-on envisager pour les niveaux d'agrégation et pour les variables à conserver lors de ces changements d'échelle ?

0.1.3 Présentation générale de la démarche

L'objectif principal visé est donc l'ajustement du modèle GreenLab aux arbres. On verra néanmoins au cours du manuscrit que les méthodes développées dans ce but s'appliquent en fait dans un cadre plus large, pour des plantes ramifiées de types variés.
Pour répondre à cette problématique, nous avons tout d'abord étudié le comportement du modèle dans sa formulation actuelle. Nous rappelons en premier lieu les grandes caractéristiques et les principaux choix de modélisation de GreenLab, que nous analysons à la lumière d'une revue bibliographique succinte (chapitre 1). Nous détaillons ensuite les équations du modèle et son formalisme pour les parties concernant la mise en place de la structure (chapitre 2) et le fonctionnement physiologique (chapitre 3), en nous attardant sur les nouveautés introduites depuis la thèse d'Amélie Mathieu [Mathieu, 2006]. L'introduction d'un nouveau type d'équation pour le calcul de la production de biomasse (basée sur la loi de Beer) modifie les propriétés du modèle. Nous montrons en particulier que, sous certaines conditions, la croissance de deux plantes peut être identique, du point de vue de la biomasse affectée à leurs compartiments, alors que leurs paramètres de développement et d'architecture sont différents (chapitre 4). Ces caractéristiques sont importantes dans une optique d'analyse multi-échelle de la plante.

Enfin, une étude particulière a été menée sur les conséquences du choix de modélisation de l'allocation de biomasse à la croissance secondaire sur le comportement du modèle. En effet, trois modes de calcul de cette allocation ont été introduits et l'étude de leurs implications respectives sur la croissance peut aider à faire un choix. Les autres critères qui guident ce choix sont bien sûr l'adéquation du modèle aux données mesurées et la possible interprétation physiologique des variables de contrôle choisies (Nombres de feuilles, rapport offre-demande Q/D, production de biomasse Q).

Pour compléter la présentation du modèle, nous résumons brièvement en chapitre 5 un travail finalisé durant la thèse qui concerne l'étude des apports possibles de la modélisation de croissance des plantes à la génétique quantitative. Nous avons mis en place le couplage de GreenLab avec un modèle génétique très simple, reliant par des multiplications matricielles les valeurs des paramètres du modèle à des variables représentant le génome de la plante. La définition de ce génotype virtuel a permis de simuler des croisements entre plantes et donc de reproduire la chaîne complète des processus à mettre en oeuvre pour extraire l'information génétique associée aux paramètres du modèle à partir de mesures sur les plantes. Cela passe par l'inversion successive des deux modèles : le modèle de croissance GreenLab et le modèle génétique. L'inversion du modèle génétique est résolue en utilisant les techniques développées par les généticiens pour la détection de QTL (Quantitative Trait Loci), notamment le logiciel QTL Cartographer. Bien que ce travail soit resté entièrement théorique, il a permis de souligner l'intérêt de s'appuyer sur la modélisation pour fournir de nouveaux critères pour la sélection génétique.

Dans une seconde partie, nous nous intéressons à la description et à l'ajustement du modèle selon différents niveaux de détails et d'agrégation des mesures. Pour répondre au problème soulevé, nous avons confronté le modèle complet à des plantes réelles sur des cibles simplifiées. Cela nous a amenés à aborder le développement de modèles simplifiés, avec différents niveaux d'agrégation des variables.

Plus précisément, nous rappelons tout d'abord la procédure et l'ensemble des mesures utilisées dans les applications précédentes pour l'ajustement du modèle en l'illustrant par quelques applications réalisées au cours de la thèse (jeunes pins, *Cecropia sciadophylla*, chapitre 7.1). Puis nous présentons les différents niveaux de simplification des mesures que nous avons considérés. Nous avons envisagé principalement trois niveaux : le niveau 0 ou cible "cumulée" pour lequel seules des données par compartiment sont accessibles, le niveau 1 ou cible "Lollipop" pour lequel on se focalise sur le tronc, le reste de l'arbre étant décrit par compartiments et enfin le niveau 2 ou cible "Rattle" pour lequel l'agrégation des données se fait seulement pour les branches d'ordre 2, le tronc étant décrit à l'échelle de l'organe. Ces choix se justifient par les objectifs envisagés pour chaque type de cible (par exemple, le niveau 2 permet d'envisager des applications dans le domaine de l'étude de contraintes biomécaniques dans le tronc) et par la faisabilité pratique du protocole expérimental associé. Les niveaux choisis ont

aussi l'avantage d'être plus cohérents avec le type de données mesurées lors des inventaires forestiers ou générées par les modèles de croissance de plantations classiquement employés (en particulier pour le niveau 1). Nous avons distingué deux classes de cibles selon leur nature : des cibles décrivant les poids et les dimensions d'organes ou de compartiments de la plante (associées donc à son fonctionnement) et des cibles décrivant des nombres d'organes et/ou leurs connexions topologiques (associées au développement). Nous nous servons de ces différents types de cibles pour analyser la croissance de plantes selon les données accessibles (chapitre 8). Cette analyse est aussi très dépendante de la version du modèle que l'on choisit (déterministe $GL1$, stochastique $GL2$, déterministe avec rétroaction de l'état trophique sur le développment $GL3$). Nous présentons les problématiques spécifiques aux types de situations rencontrées et les solutions retenues en nous appuyant sur les applications à des plantes réelles réalisées dans le cadre de la thèse. Une plante emblématique a été choisie pour chaque version du modèle : le pin (*tabulaeformis*) pour la version $GL1$, le blé tallé pour la version $GL2$ et le hêtre pour la version $GL3$. Cela nous amène finalement à présenter une ébauche de typologie pour l'analyse de plantes à l'aide du modèle GreenLab.

Ces considérations nous ont amenés à envisager une étude plus théorique de ces niveaux de simplifications (chapitre 9). On cherche à vérifier des équivalences entre le modèle complet et différents modèles simplifiés. L'intérêt des modèles simplifiés est d'augmenter la significativité des paramètres et d'en réduire le nombre. Les équivalences sont imposées en cohérence avec le niveau d'agrégation des données correspondant aux simplifications choisies. Il s'agit par exemple d'imposer la conservation de la production totale de biomasse, de la demande totale de la plante (ce qui assure la conservation d'une variable-clé de GreenLab, le rapport offre sur demande) et de l'allocation par compartiments. On cherche à écrire un modèle agrégé à partir des équations du modèle complet : on considère que l'on connaît les paramètres du modèle complet et l'on cherche à en extraire ceux du modèle simplifié. Les équations issues de cette procédure permettent de conserver les équivalences recherchées mais ne sont pas forcément les plus pertinentes lorsqu'on les considère indépendamment du modèle complet. En particulier, la répartition de la biomasse entre les cernes d'une structure est fortement dépendante de sa topologie ce qui rend difficile le maintien d'équivalences strictes lors des changements d'échelle. Il en va de même pour les variables dont les variations sont fonction de l'état trophique de la plante (variables "$GL3$"). Pour l'application future de ces équations à la croissance de plantes réelles, nous proposons donc des modifications du modèle simplifié qui entrainent la non-conservation des équivalences avec le modèle complet mais dont les principes sont en adéquation avec les choix de modélisation de GreenLab.

0.1.4 Remarques préliminaires à la lecture du document

Nous donnons ici quelques précisions sur le vocabulaire employé. Ces précisions seront rappelées au cours de la lecture.

- On nommera *modèle complet* le modèle GreenLab présenté dans la première partie, qui peut être décliné selon les versions GL1, GL2 ou bien GL3 et dans lequel on décrit toute la structure branchée, jusqu'aux ordres les plus élevés. Par extension, on désigne par *arbre complet* un arbre simulé selon ce modèle.
- Par opposition, l'arbre simplifié est un arbre que l'on décrit de manière simplifiée (ou agrégée ou cumulée) et que l'on simule à l'aide d'un modèle simplifié. Ce modèle simplifié peut être de différents types, selon la plante considérée et le niveau de détails que l'on peut espérer atteindre en pratique lors des mesures.
- Compartiments : certaines mesures pour une description simplifiée de la plante sont décrites par compartiments. Un compartiment est constitué de l'ensemble des organes de même type contenus dans une structure donnée. Par exemple, une structure mesurée de façon simplifiée compte un compartiment bois, un compartiment feuilles et un compartiment fleurs ou fruits.
- Remarque : Nous employons de manière abusive le terme "biomasse" et le terme "poids" pour désigner la quantité de matière (fraiche ou sèche) contenue dans un organe.
- Nous employons ici indifféremment les dénominations "métamère" et "phytomère".

Pour faciliter la lecture du manuscrit, nous avons cherché à employer systématiquement les mêmes notations pour les différents paramètres des équations et notamment pour les nombreux indices nécessaires. En particulier, nous essayons dans la mesure du possible de respecter la convention suivante : les entiers notés en indice représentent les âges physiologiques, ceux notés entre parenthèses représentent des âges chronologiques ou des cycles de croissance et les lettres en exposant représentent les types d'organes. Le lecteur trouvera en annexe A un tableau récapitulatif de ces notations.

Les logiciels : Digiplante, GreenScilab

Les résultats de simulation et d'ajustement présentés dans cette these ont été réalisés à l'aide des logiciels Digiplante et GreeenScilab. La participation au développement de ces deux logiciels a fait partie du travail de thèse.

Digiplante est développé dans le laboratoire de Mathématiques Appliquées aux Systèmes (MAS) de l'Ecole Centrale Paris. Il est basé sur le langage objet C++, qui est un langage particulièrement bien adapté à la modélisation de la croissance des plantes du fait de la structure modulaire d'une plante. En effet, une plante peut être décomposée en unités élémentaires qui ont des propriétés similaires et qui interagissent en parallèle entre elles en permanence. Digiplante comprend la version déterministe du modèle GreenLab (GL1) ainsi que l'introduction de la rétroaction de l'état trophique de la plante sur son développment (GL3).

GreenScilab est une boite à outils du logiciel libre SciLab. Elle est développée principalement au LIAMA. La simulation et les ajustements sont légèrement moins performants mais le libre accès à son code en fait un vecteur privilégié pour la diffusion du modèle et pour l'enseignement (utilisation pour des démonstrations et des TPs de modélisation). GreenScilab inclut la version stochastique du modèle (GL2).

Les collaborations

Une partie importante de mon travail de thèse a été consacrée aux applications sur des plantes réelles. Celles-ci se sont déroulées dans le cadre de collaborations avec des instituts de recherche agronomique ou en foresterie :

- L'Université d'Agriculture de Chine (China Agricultural University, CAU), située à Pékin : coton (Li Dong), riz (Zheng BangYou), pin (Wang Feng) sous la direction du Pr Guo Yan, tomate (Pr Zhang BaoGui).
- l'Institut Forestier de Chine (Chinese Academy of Forestry, CAF), situé à Pékin : pin *tabulaeformis* (Guo Hong et Hong LingXia)
- Le Laboratoire d'Etude des Ressources Forêt-Bois (LERFoB, Champenoux) : hêtre (Thiéry Constant, Gérard Nepveu)
- Le Laboratoire d'Ecophysiologie des Plantes sous Stress Environnementaux (LEPSE, Montpellier SupAgro - INRA) : *Arabidopsis thaliana* (Angélique Christophe, Jérémie Lecoeur)
- Le département Horticultural Supply Chain Group de l'université de Wageningen aux Pays-Bas : blé, chrysanthème (Kang MengZhen)
- L'UMR Botanique et Bioinformatique de l'Architecture des Plantes (AMAP) du CIRAD (Montpellier) : *Cecropia sciadophylla* (Patrick Heuret, Camilo Zalamea, Daniel Barthélémy)

0.2 Introduction (English version)

0.2.1 Plants, complex systems?

Plants and plant populations can be considered as complex systems, in a mathematical sense. Complex systems consist of high numbers of heterogeneous entities between which multi-scale interactions generate holistic behaviours (*i.e.* some emergent properties of the system cannot be deduced from the independent studies of its components). Therefore, a reductionist approach (that is to say the decomposition of problems into independent sub-problems) is not adapted: it is necessary to analyze the system as a whole [Ricard, 2003].
Dealing with complex systems intrinsically requires multi-disciplinary and multi-scale approaches. Indeed, the problems raised and the methods developed to address them

come from different scientific disciplines. It is also important to confront theoretical developments to real systems. The inherent complexity of these systems implies that their analysis must be performed at different spatial and temporal scales. The theory of hierarchies describes systems that can be represented as nested sub-systems. This definition can be applied to plants, the different levels being for instance organs, phytomers, growth units, axes, structures, plants [Godin and Caraglio, 1998]. The properties of structures at each level generate the emergence of collective behaviours and high-level organization. Inter-level interactions are considered as constraints or as emergent properties.

In many cases, complex systems reveal to be very robust and insensitive to small perturbations. Perry [1995] characterizes these systems as "critical" or "metastable" : robust for some perturbations but affected by rapid changes if some thresholds are exceeded. This phenomenon can be observed for plants: they exhibit relatively robust properties at a level n (that are the bases of some modelling approaches that rely on "functional balance" principles [Mäkelä, 2003]) although at level $n-1$ the physiological processes involved may be of very different natures depending on the growth conditions. According to Mäkelä [2003], if the levels are independent enough, the development of a mechanistic model can be based on studies of the interactions between the components of the adjacent underlying level only: including processes at smaller scales would reduce the robustness and stability of the system.

These theories have been mainly developed for applications at landscape scale in ecology [Li, 2000]. Mäkelä [2003] underlines that forest management practices tend to manage exploited stands like natural forests, *i.e.* introducing more and more heterogeneity regarding species, age classes and spatial distributions. The author concludes that forest ecosystems must be considered as hierarchical entities consisting of organization levels at different spatial and temporal scales. More modestly, we raise here the question of changes of scales for single plants only. Regarding plant architecture, it is widely admitted that plants are modular systems that involve multi-scale phenomena [Barthélémy and Caraglio, 2007]. In particular, this idea has led the first simulations of virtual plants to methods based on self-similarity such as the fractal theory. Similar principles have been applied to define the concepts of architectural model and architectural unit. So the concepts, the adequate formalism and the tools for the multi-scale analysis of plant architecture have been already developed [Godin and Caraglio, 1998]. But a further step remains to be achieved: the integration of functional processes into this framework, *i.e.* the multi-scale analysis of a functional-structural model.

0.2.2 Trade-off between model complexity and parameter identification

As for most complex systems, confrontation with real data raises several problems related to the reconstruction of the system dynamics from data, which is often source of difficult inverse problems. It also raises the problem of data acquisition and definition of a relevant description scale for measurement data and variables. It implies finding a trade-off between the complexity of real phenomena and the simplifications induced by the modeling work.
The work presented in this thesis is applied to the GreenLab model of plant growth. The first chapter synthesizes the main characteristics of this model with a short bibliography review of other models of the same kind. We focus on tree growth models although the case of herbaceous is also addressed. Let us remind here the main principles of GreenLab:
GreenLab is a structural-functional model simulating the processes of biomass production and allocation into organs at whole-plant scale. Organogenesis is driven by a dual-scale automaton determining the topological rules of organ emission and organization. The state variable of this automaton is the differentiation state of apical meristems, called their physiological age. Three versions of the model can be considered: deterministic ($GL1$), stochastic ($GL2$) or mecanistic, *i.e.* deterministic with feedback of photosynthesis on organogenesis ($GL3$). Biomass production is computed at each time step (growth cycle) depending on the plant total foliar surface and taking into account the effects of self-shadowing between leaves. Biomass is allocated to expanding organs regardless of their position (common pool of biomass) according to a source-sink model.
The work of Amélie Mathieu [Mathieu, 2006] has brought significant advances allowing realistic simulations of branched plants with GreenLab. Trees are considered as self-regulating systems with several physiological and developmental processes being influenced by their internal trophic state. It allows reproducing the tree architectural plasticity in response to environmental or ontogenic changes (e.g. progressive appearance of higher ramification orders in branches at different growth stages). The model also generates cyclic patterns as emergent property, similarly to what can be observed on real plants (e.g. rhythmic appearance of fruits) [Mathieu et al., 2008].

A major objective of GreenLab is the prediction and optimization of plant growth. Fitting the model on real data is therefore a crucial step. It relies on parallel fitting (*multi-fitting*) of individual organ mass. It allows tracking back the dynamics of biomass allocation through the plant life. The previous works with GreenLab have dealt with its parameter identification for several plants such as:

- wheat [Zhan et al., 2000], [Kang et al., 2007b]

- sunflower [Guo et al., 2003]
- maize [Guo et al., 2006] [Ma et al., 2007]
- tomato [Dong, 2006] [Dong et al., 2008]
- chrysanthemum [Kang et al., 2006b]
- rice [Mathieu, 2006]
- coffee tree [Mathieu, 2006]
- young beech trees [Mathieu, 2006]

These applications concern plants with a relatively low branching order (at most three for wheat tillers). Detailed description and measurements are thus possible. But it is not feasible for plants with higher branching orders. Gathering the set of data needed for the identification of the model would require weighting and measuring independently all leaves, internodes and fruits of any target plant. It represents a very tedious work for an old plant with a complex architecture. Thus, although simulation is efficient, limitations arise from the procedure of data collection. The level of details available for the collection of data is coarser than that of simulations. It implies that some work is needed to analyze the effects of data aggregation and to study the possibilities of model simplifications so that the scale of simulation is consistent with that of observations. This is our objective.

The main question raised in this work is the following:
How to deal with parameter identification and the associated experimental protocol for branched plants having complex architectures ?
More precisely, several questions arise:

- In the process of fitting the model on plants, since detailed descriptions of complete real systems are not available, simulated plants will necessary have a topological structure different from those of real plants. It raises the question: what are the effects of changes in topological structure on biomass production and allocation in the model ?
- For model parameterization, what kind of topological data can be included in the target file ?
- If no topological data are available, how to set some default rules for the plant topological development ?

- What equations and which parameters can be defined to drive the growth of a plant simulated with a simplified model where some aggregated key-variables would be conserved ? What are the possible levels of simplification and which key-variables are the most relevant ones ?

0.2.3 Presentation of the approach

Our main objective is the identification of the model for trees. However we will see that the methods developed have a larger field of applications and can be applied to different kinds of branched plants.
To address the questions presented above, we first analyzed the model behaviour in its current formulation. We remind the main characteristics and modelling choices of GreenLab at the light of a brief bibliography review (chapter 1). Then we detail the model equations and the dedicated formalism that was developed for the description of the plant topological development (chapter 2) and its physiological functioning (chapter 3). In these chapters, we focus on some new features of the model that were introduced since the PhD of Amélie Mathieu [Mathieu, 2006]. In particular, the introduction of a new kind of equation for biomass production (based on the Beer-Lambert law) has deep implications on the model behaviour. We show that, under specific assumptions, the growth of two plants can be identical for their compartment biomass although their topological parameters are different (chapter 4). These properties are important in the context of a multi-scale analysis of plant growth. A particular attention was also given to the consequences of different modelling choices for secondary growth. Indeed, three approaches were proposed, so analyzing the properties of the associated model for each case provides interesting information to help choosing one of them. The other criteria are of course the accordance of the model to real data and the relevance of the main variables chosen, in terms of their physiological interpretation (Number of leaves, ratio of biomass supply to demand Q/D, biomass production Q).
In addition to the presentation of the growth model itself, we present in chapter 5 a work about the potential benefits of plant growth modelling for quantitative genetics. We introduced the coupling of GreenLab with a simple genetic model: it uses matrix multiplications to link the model parameters to virtual variables representing the plant genome. Defining a virtual genome allowed simulating plant reproduction and thus simulating the complete chain to extract genetic information related to the model parameters from phenotipic data. It requires the successive inversion of two models: the growth model GreenLab and the genetic model. It is based on techniques developed by geneticists for QTL detection (Quantitative Trait Loci). Although that work remains entirely theoretical, it underlines that growth modelling can bring new criteria for genetic selection.

In a second part, we focus on the model description and its parameterization with

different levels of simplification and data aggregation. To address the problem, we confronted the complete model to real plants with simplified targets. It has led to the definition of simplified models with different levels of variable aggregation.

More precisely, we first remind the procedure and the set of data used for parameter identification of the complete model. We illustrate it with some of the applications realized during the PhD (young pines, *Cecropia sciadophylla*, chapter 7.1). Then we present the different levels of measurement simplifications that were considered. Three levels were defined: level 0 or "cumulated target" for which only compartment data are available, level 1 or "lollipop" target for which the trunk is detailed and the crown is described at compartment level and level 2 or "rattle" target for which data aggregation is done only at branch level. These choices can be justified by the objectives defined for each level (for instance, level 2 allows considering applications in predicting biomechanical stresses in the trunk) and by the practical feasibility of the associated experimental protocol. The aggregation levels are also chosen in adequacy with the kind of data classically collected for forest inventories and generated by forestry models. We distinguished two kinds of target data: those containing mass and dimensions of organs and compartments (related to plant functioning) and those containing the topology of the plant (related to plant development). This analysis is also strongly dependent on the version of the model ($GL1$, $GL2$, $GL3$). We present some specific problems and solutions developed for some real plants that were studied during the PhD. One representative plant was chosen for each of the model version: pine (*tabulaeformis*) for $GL1$ version, tillering wheat for $GL2$ and beech tree for $GL3$. It eventually lead to defining an attempt of typology for the analysis of plants with GreeenLab.

This work has also led to a more theoretical study of the different aggregation levels (chapter 9). The objective was to keep some equivalences between simplified and complete models. The interest of simplified models is to reduce the number of parameters and to improve their significance. The choice of the variables to keep is consistent with the level of data aggregation. It can be for instance the total biomass production and the whole plant demand (which insures that the key-variable of the biomass supply to demand ratio is conserved) or the biomass allocated to compartments. To define the equations of the simplified models from those of the complete model, we first use the method of variable aggregation. It allows keeping strict equivalences but the equations that arise from this procedure are not necessary the most relevant ones when considered independently from the complete model. For instance, biomass partitioning for ring growth is strongly dependent on the plant topology and therefore it is difficult to keep these equivalences with changes of scales. The same problem is raised for some variables that depend on the trophic state of the plant ("$GL3$" variables). Therefore, for future applications of these equations to the simulation of real plant growth, we propose modifications of the simplified model. These modifications imply non-conservations of equivalences with the complete model but their principles are consistent with the GreenLab philosophy.

0.2.4 Preliminary remarks

We give here some vocabulary remarks.

- We call *complete model* the current version of the GreenLab model as presented in the first part (with its $GL1$ to $GL3$ versions). The whole branching structure is described, including the highest branching orders. By extension, *complete trees* refer to trees simulated according to this model.

- By opposition, *simplified tree* refers to a tree described in a simplified manner and simulated with a *simplified model*. Simplified models can be of different kinds depending on the plant and on the level of details that can be reached in practice.

- Some measurement data can be gathered at *compartment* scale. Compartments consist of the set of organs of the same kind borne by any structure. For instance, simplified measurements of structures can include wood, leaf and flower or fruit compartments.

- Remark: we abusively use the terms of "biomass" and "weight" to refer to the quantity of fresh or dry matter of organs.

- We equally use in the words "metamer" and "phytomer".

To facilitate the reader's task, we tried systematically to use the same notations for the model parameters and their index. In particular, the following conventions were adopted and applied as widely as possible: subscript integers represent physiological ages, those between brackets represent chronological ages or growth cycles and letters in exponents represent organ types. A list of some parameters and notations can be found in Appendix A.

Softwares: Digiplante, GreenScilab

The results of simulation and fitting presented in this thesis were performed using the Digiplante and GreeenScilab softwares. It was part of the PhD work to contribute to the development of these softwares.
Digiplante is developed at the laboratory of applied mathematics (MAS) in Ecole Centrale of Paris. It is based on the object-oriented language C++ which is adapted to plant growth modelling due to plant modular structure. Indeed, plants can be decomposed into elementary units that have similar properties and that interact in parallel. Digiplante includes the deterministic version of GreenLab ($GL1$) and the version with feedback of the trophic state of the plant on its development ($GL3$).

GreenScilab is a toolbox of the open source software SciLab. It is mainly developed at LIAMA (Sino-french Laboratory of automatic, computer sciences and applied mathematics). It includes the deterministic ($GL1$) and the stochastic ($GL2$) versions of GreenLab. As it is open source, it is a privileged support to teaching sessions on the model.

Collaborations

An important part of the PhD work was dedicated to applications on real data. This work was realized in collaboration with agronomy or forestry research institutes:

- China Agricultural University (CAU), in Beijing: cotton (Li Dong), rice (Zheng BangYou), pine (Wang Feng) under the direction of Pr Guo Yan, tomato (Pr Zhang BaoGui).
- The institute of Forest resource information techniques, (Chinese Academy of Forestry, CAF), in Beijing: Pinus *tabulaeformis* (Guo Hong and Hong LingXia)
- The laboratory of wood and forest resources (LERFoB, Champenoux): beech tree (Thiéry Constant, Gérard Nepveu)
- The laboratory of plant ecophysiology under environmental stress (LEPSE JRU, Montpellier SupAgro - INRA): *Arabidopsis thaliana* (Angélique Christophe, Jérémie Lecoeur)
- The Horticultural Supply Chains group of Wageningen university, the Netherlands: wheat, chrysanthemum (Kang MengZhen, Ma Yuntao, Ep Heuvelink)
- The joint research unit Botany and computational plant architecture (AMAP), CIRAD (Montpellier): *Cecropia sciadophylla* (Patrick Heuret, Camilo Zalamea, Daniel Barthélémy)

Part I

The GreenLab model: description and analysis

Chapter 1

Principles and hypotheses in the bibliographical context

In this chapter, we present synthetically some main features of the GreenLab model and we compare it with other existing models of plant growth.

The bibliography references address mainly tree growth models but other plants are considered too, especially for descriptions of the physiological bases and justifications of modelling choices for allocation processes.

1.1 Objectives of the model

In any modelling approach, it is important to identify clearly the model goals. The characteristics of most models are strongly linked to their application fields. The purpose of the GreenLab model is to be a generic mathematical model of plant growth for agronomy and forestry applications. As such, its modelling concepts should be robust and general enough to be applied to herbaceous as well as trees. It should produce a performing simulation tool that simulates the minimal set of processes involved in the growth of most plants. It deals with components and processes at mesoscale. It aims at answering the following question: from the basic mechanisms that are common to most plants, what observed phenomena can be reproduced ? It is oriented toward applications for decision support systems to improve crop yield and forest management. Therefore, particular focus is given to parameter identification, optimization and optimal control. GreenLab has inherited from the knowledge acquired on one hand from the architectural models based on botanical analysis and on the other hand from process-based models describing ecophysiological processes. It is the result of a joint work between researchers from several scientific disciplines: botany, ecophysiology, agronomy, applied mathematics and computer sciences.

1.2 A discrete-time and -space model

Spatial and temporal scales are of crucial importance as they are major determinants in the model development. They must be carefully chosen according to the goals and the potential applications. GreenLab is a dynamic discrete model, from both a temporal and a spatial point of views. Concerning the temporal aspect, the growth dynamics is sampled at different dates and the time units are called *growth cycles* (GC). Consequently, equations of the model are recurrent equations (and not differential equations) that appear under the form: $X_{n+1} = f(X_n, U_n)$, where X_n is the vector of state variables of the system at cycle n and U_n the vector of control variables that can influence the plant growth. This formalism is detailed in the next chapter. An advantage of using a discrete time scale is that it avoids facing the time-consuming problem of solving partial derivative equations. The determination of the cycle duration depends on the plant species studied and is detailed in chapters 2.2 and 2.3. Regarding spatial description, the plant is decomposed into sets of organs. Six kinds of organs are mainly considered in the model:

- internode (portion of stem situated between two nodes. It includes the pith and several rings)
- blade (leaf lamina)
- petiole (leaf stem)
- flower (reproductive structure)
- fruit (ripened ovary)
- bud (embryonic shoot)

It is thus assumed that all vascular plants can be described from organized combinations of these kinds of generic organs.

1.2.1 Choice of temporal scale

In GreenLab, the time step is based on the plant life-time and on its specific rhythm of growth. For trees in temperate zones, it is generally one year. This time step seems adequate to simulate the growth over more than 100 years, a typical duration for simulating the long-term effects of silvicultural practices. It allows neglecting some short-term phenomena such as the dynamics of organ expansion, considered as immediate (two or three weeks) compared to the growth cycle duration (one year) for trees like Mapple trees or Horse chestnut trees. However this assumption does not hold for some trees, for example in the case of polycyclism (*i.e.* several growth flushes per year) or for some tropical trees (e.g. Eucalyptus) that do not exhibit visible growth units and have a fast continuous growth.

1.2.2 Choice of spatial scale

Regarding spatial decomposition, organ scale appears as a natural scale for considering the growth of individual plants. The choice of this level of decomposition unit allows dealing with plants of diverse sizes from herbaceous such as Arabidopsis (20 cm high) to trees such as Beech tree or Pine tree (more than 10 m). Indeed, plants can be seen as modular systems where modules can be considered at different levels, the smallest units being phytomers in our study. A phytomer is defined as the entity formed by an internode, its leaves (and roots if any), and axillary buds (see section 2.1).

As stated in [Minchin and Lacointe, 2005] who recommend to switch from a reductionist approach to a holistic approach, there is currently a general trend to integrate processes at molecular scales, in the hope of creating more mechanistic models with growth processes appearing as emergent properties. On the contrary, GreenLab is voluntarily restricted to an analysis of the plant at the mesoscale. We consider that this scale is appropriate (and sufficient) for most applications in crop yield prediction or forest management. On the other hand, compared to simulations at compartment levels (e.g. root/shoot, crown/stem/roots), organ scale has several assets for tree growth models, as well for their potential applications as for the constraints on building the model. It allows assessing the importance of structural development as an adaptation mechanism in trees, for example to predict the effects of disturbances like pruning or thinning or of an abrupt change in environmental conditions (e.g. light regime after gap formation). Models at organ scale can also take into account biomechanical properties that determine wood quality and that are highly related to branch positions. They provide more accurate variables to include competition between trees and to simulate the individual-based growth of heterogeneous stands. The choice of organ scale also suppresses some problems met at compartment scale when developing models. Perhaps the most important advantage is that no a priori crown shape has to be chosen. For instance, in [Mäkelä, 1986], the lower limit of the live crown depends linearly on tree height. The compartment demands have to be made artificially varying with a competition index ([Mäkelä, 1986]) or with soil water potential (for allocation to roots) (e.g. [Zhang et al., 1994]). On the contrary for organ-based models, compartment demands are simply the sum of individual organ demands and their dynamics are driven by the tree ontogenetic development through variations of its number of organs. No senescence rate has to be artificially incorporated in compartments: it is easily assessed from field observations at organ scale. Table 1.1 gathers the spatial and temporal scales chosen in some other tree growth models.

Such characteristics can be found for other models in [Le Roux et al., 2001]. Note that when a daily computation step is considered, the model scope is generally limited to young individuals with a growth extension of no more than a few years. Among the models referenced here, only the LIGNUM model has similar characteristics to GreenLab for both spatial and temporal scales.

Model	Target trees	Spatial scale	Temporal scale
SIMWALL	young walnut tree	organ	hourly
ECOPHYS	poplar	organ	hourly
EMILION	maritime pine	organ	hourly
Zhang 1994	red pine	compartment	daily
L-PEACH	peach tree	organ	daily
Sterck 2005	generic	organ	10 days
LIGNUM	generic	organ	1 year
GreenLab	generic	organ	1 year

Table 1.1: Spatial and temporal scales of some tree growth models. The models are referenced from: [Balandier et al., 2000], [Rauscher et al., 1990], [Bosc, 2000], [Zhang et al., 1994], [Allen et al., 2005], [Sterck et al., 2005], [Perttunen et al., 1996]

1.3 A model relying on a sound mathematical formalism

A dedicated mathematical formalism was developed to describe the plant topology through a recurrent formula ([de Reffye et al., 2003]; [Cournède et al., 2006]). Given this formula, the number of organs and their spatial organization are completely determined at each time step, at least for the deterministic version of GreenLab. In the same way, the computation of biomass production at a given cycle can be explicitly written according to the state variables of the model at the previous cycles. This mathematical formalism has several advantages:

- the simulation software can be easily checked to detect possible errors and bugs: it prevents from building a complicated black box whose outputs cannot be calculated without running simulations.
- the model behaviour can be analytically studied and limit conditions can be defined
- the simulation speed is enhanced since growth simulations rely on one key equation and does not need to consider organ-by-organ calculations
- it allows using mathematical algorithms for model parameterization and optimal control problems.

More generally, Kurth [1996] presents some arguments in favour of the use of an explicit formalism to describe plant architectures. Equations are more elegant and condensed than a complicated algorithm or flowchart that must be analyzed to understand the model. The model gets a higher level of transparency and of universality when an exact

specification language, designed for that purpose, is used. It allows applying general theories that have been previously developed in other disciplines.

1.3.1 Model identification

A main characteristic of GreenLab, compared to other models, is the way chosen for its identification (sensu Walter and Pronzato [1994]). In most models, parameters are taken from bibliography or they are calibrated independently for each of the processes involved in the model. On the contrary, GreenLab is associated to a dedicated fitting procedure to calibrate the model to different species. All different outputs (weights of organs and of compartments) are fitted together in parallel, including data from several growth stages of the plant [Guo et al., 2006]. As such, it can be considered as a solver to determine the source-sink dynamics in a plant from discrete measurement data. Contrary to other fields of application of discrete dynamic models, the measurement process is destructive. It implies that the growth of a particular single plant cannot be followed through its whole life span. Rather, several individuals are measured and considered as several realizations of the same plant.

1.4 A structural-functional model

1.4.1 Different classes of models

Kurth [1994] classifies tree growth models into three categories: aggregated models, process models and morphological models. Aggregated models deals with production of forest stands and often rely on statistical relationships between global key variables such as crown volume, wood volume, basal area, height, diameter at breast height.... Process-based models are more explanatory in the sense that variations of the state variables are the emergent result of interactions between underlying processes: photosynthesis, assimilate allocation between plant compartments. Morphological models, also called geometrical models or architectural models, focus on the 3D architecture of plants. Organs are only decorative and do not play any functional role. They have led to the emergence of several kinds of formalisms and simulation platforms allowing the integration of plant structures. Recently, a new class has emerged that mixes the specificities of both process-based and morphological models: functional-structural plant models (FSPM), as reviewed in [Sievänen et al., 2000] and [Prusinkiewicz, 2004], incorporate functional processes into the framework of tree architecture. Coupling process-based models and architectural models has become a key issue in the fields of tree growth prediction as well as crop yield prediction.
Although functional-structural models could appear as the most achieved modelling approach, they are not yet used for industrial applications. Each of these four categories of models has specificities that make it suitable to a particular user community. Table

1.2 gives the major trends (non exhaustive) for the application field of each kind of models.

Model	Application field
Aggregated models	forest management
Process-based models	crop yield prediction
Morphological models	video games, landscape/urbanism, design
Functional-structural models	research and teaching

Table 1.2: Different types of models and their main application fields

GreenLab has some specificities that position it at an intermediate level between process-based models and functional-structural models. It has inherited of the notions of self-similarity and multi-level organization of plant topology from the morphological models. It also includes the main laws of process-based models, although the environmental control is not yet as detailed. The 3D architecture is simulated but only partially taken into account in the functional part. For several processes, only the numbers of organs appeared at each cycle are important and not their spatial organization. Those choices avoid some drawbacks inherent to functional-structural models, as presented below.
In the following paragraphs, we review the main characteristics of those four classes of models. We discuss the reasons that have led to the emergence of functional-structural models and we explain the positioning of GreenLab.

1.4.2 Aggregated models

Decision support systems for forest management generally rely on aggregated models based on statistical analysis of data from the past. They are calibrated for specific geographic areas and only for the major tree species of the region. They have good predictive abilities because they are not limited by the lack of biological knowledge. However, adding new input factors implies building a new model. And extrapolation to other species or new locations is often impossible: when they are used outside their ranges, their predictive values are limited [Marcelis et al., 1998].

1.4.3 Process-based models

These restrictions are partially solved by process-based models. They rely on biological knowledge and thus the experimental data needed for their development is more focused. They are more flexible because they integrate mechanistic interactions instead of being mere data tables that need to be adapted to each new situation. They provide more reliable predictions in changing environmental conditions. They have been mainly applied to crop yield production and, in forestry field, to studies of the effects of global

warming and other long-term climatic changes on large scales (e.g. [McMurtrie et al., 2001]).

They have benefited from the advances of research in plant physiology. Most of them consider plants or stands as sets of compartments that can produce or exchange matter and information with each others and with their environment.

Process-based models have provided useful tools to predict crop yield under changing conditions (temperature, light, CO2, water, nutrients) or to optimize water and fertilizer supplies. However, they have revealed several limitations that have restricted their applications. They only integrate an empirical control of environmental stresses at compartment level ([Jeuffroy et al., 2002]) and do not account for the architectural response to environmental factors, such as tillering or organ abortion ([Dingkuhn, 1996]; [Luquet et al., 2007]). Yet yield determination is influenced by architectural plasticity, especially under stress conditions, and this influence has to be considered at the phytomer level instead of the square meter scale. Moreover, modellers often encounter difficulties to get reliable computation of leaf area index (LAI) which is mostly the main component of biomass production modules ([Marcelis et al., 1998]; [Heuvelink, 1999]). They also have difficulties to deal with the inter-plant variability and to handle the often complex interactions between all the different physiological modules: process-based models are often presented as a complicated flowchart showing all the interactions between sub-models that have been calibrated and tested independently but that are difficult to integrate into a common system ([Heuvelink, 1999]). These drawbacks result from the fact that process-based models do not take into account plant morphogenesis: at compartment level, since all organs are mixed together, the memory of the growth process is lost and so is the architectural plasticity that reflects feedbacks between growth and development processes. The endogenous parameters that control both plant development and plant growth are useful key-components for yield prediction.

Moreover, for trees, architecture is required in many applications. For instance, the 3D structure influences assimilate partitioning, light interception capacity and gas exchanges in the foliage [Sievänen et al., 2000]. Modelling the effects of pruning and thinning strategies for fruit trees or for forest management imposes to have access to plant structure. Wood quality is greatly influenced by biomechanical constraints in stem which depend on tree shape and structure ([Fourcaud et al., 2003]; [Jirasek et al., 2000]).

3D plant structures have not been considered in process-based models. There was a general feeling that their implementation would be difficult, that it would require high-performance computers and time-consuming experimental data collection and that the resulting model would be too complex and difficult to manage. However the experience of morphological models proved that those difficulties could be overcome. Thanks to the rapid development of computer graphics and of dedicated formalisms, it is now possible to take into account morphogenesis.

1.4.4 3D models and methods for plant structure analysis

The first representations of tree structure were based on the notion of self-similarity, based on the existence of repetitive patterns at different scales in plants. This observation has led naturally to the use of fractal rules (repetition of similar shapes at different scales in an object) or iterated function systems (IFS) where several iteration of affine transformations are applied to an initial figure. Different approaches to this concept are listed in [Kurth, 1994]. It can generate beautiful tree-like shapes and landscape pictures that look realistic. Those techniques are suitable for visual purposes but not for integration into process-based models. Indeed, their major drawback is that it is generally not possible to consider independently the components of the structure (e.g. removing one branch), as they are all linked together by multi-scale iterative processes. And the simulated trees are not faithful to botany as they are generated by automatic application of mathematical rules with no botanical knowledge included. Furthermore, while it is quite easy to generate the IFS resulting image (its attractor), the inverse problem has not necessary a unique solution [Shonkwiler et al., 1991] and its resolution generally requires computationally heavy heuristic algorithms such as genetic algorithms [Cofino et al., 2000].
Prusinkiewicz and Lindenmayer [1990] introduced the formalism of L-Systems for the simulation of plant structure. L-systems are defined as parallel rewriting systems operating on strings. They are particularly suitable to describe the dynamics of structure development with discrete time step (inherent to the concept of «rewriting»), as the meristems of the plant also have a parallel functioning. They consist of:

- an alphabet, i.e. a set of symbols representing the constitutive elements of the plant structure (e.g. internode, leaf, bud)
- a seed or initial string
- a set of replacement rules

L-Systems form a language dedicated to represent the growth of a branching structure incorporating its spatial organization. A dedicated platform (L-Studio) is developed in the University of Calgary to integrate L-Systems and to provide frameworks for the integration of growth models and for the 3D visualizations of their outputs (e.g. in [Prusinkiewicz, 1998]). To take into consideration some attributes of the elements (e.g. internode length, branching angle), parametric L-systems allow associating a finite sequence of real numbers to each symbol of the alphabet. The application of the replacement rules can be conditioned by external constraints and further extensions have been developed such as stochastic L-Systems (with given probabilities for application of the rules), context-sensitive L-Systems (the rewriting rules for a module depend on the following and/or preceding modules), relational growth grammars (graph rewriting rules that allows representing network structures) [Kurth, 1996]; [Kurth, 2007]. Dedicated programming languages and modelling platforms have been developed within this

framework: the «L+C »language and L-studio platform [Karwowski and Prusinkiewicz, 2003], the XL language and GroIMP platform [Kniemeyer et al., 2007].
A complementary approach has been developed since the 1970s by the AMAP research group (CIRAD) after the pioneering work of Philippe de Reffye. It is based on an analysis of tree architecture using the theory of stochastic processes ([de Reffye, 1979]; [de Reffye et al., 1988]). Tree structures result from the activity and trajectory of meristems that can be characterized by their physiological age, as defined in the next chapter. Meristems can take a finite number of states and thus tree structures can be seen as the outputs of stochastic processes that can be analyzed from the renewal theory, reliability theory, Markov chains .... It provides compact ways to represent some architectural patterns of trees and thus provides tools to analyze their ontogenetic trends and to disentangle genetic and environmental influences on architectural plasticity. It allows producing simulations of realistic stands with stochastic trees having correct branching distributions (consistent with sample measurements).

Description of tree architecture: MTG formalism

The AMAP team has also developed tools to represent plants as rooted multi-scale tree graphs (MTG) [Godin and Caraglio, 1998]. Indeed, the choice of the description scale is a crucial problem for trees. Topological models are inherently subjective [Godin and Caraglio, 1998]: it depends on the botanical knowledge of the observer, on its perception (e.g. according to the distance tree-observer, the observer may not be able to distinguish the smallest modules or on the contrary, may only be able to detail a part of the structure and to get a rough description of the rest of the tree), on the presence of morphological markers (e.g. old scars tend to deform and to disappear due to secondary growth), on the work-force needed for the measurement processes, that make impossible an exhaustive description of the whole tree. Therefore it is important to have an analysis method flexible enough to integrate various levels of description in the same plant and to integrate the correspondences between those different levels. This has been achieved through the MTG formalism. The plant is represented as a set of finite quotiented directed graphs, a graph being defined as a set of vertices (V), edges (E) and a mapping function from E into VxV. Each edge can be associated to either a branch insertion or to a succession link. A change of scale is represented by a projection of a quotiented graph (partition of the image set of vertices).
That formalism provides an essential framework to record data with a multi-level organization consistent with the tree structure and with different values of attributes that can be input for each vertex. Thus it allows processing statistical analysis and building models. Guédon et al. [2001] studied the branching and axillary flowering sequences for Apple tree, Tassili cypress and Vanilla using Markovian models. Then, Durand et al. [2005] introduced the formalism of hidden Markov tree to account for the branching structure of plants (and not only sequences on a linear pathway). They calibrated the

transition probability matrix on Apple tree and bush willow. That kind of analysis can reveal patterns that are not directly apparent in the data but those patterns are not modelled as a result of functional processes with environmental control. This is where structural-functional models can play a role.

1.4.5 Functional-structural models

While process-based models were mainly developed by the communities of plant physiologist and agronomy, geometric models have emerged in parallel from the communities of computer sciences and applied mathematics. Thus integrating those two modelling approaches requires strong interdisciplinary links. Two categories of FSPMs can be distinguished: some of them have been created by integrating physiological knowledge into a geometric model and others started from a process-based model into which data describing the 3D structure of the plant were incorporated [Sievänen et al., 2000]. Most FSPMs include the same basic processes, simulated at each time step:

- set up of new architectural units
- biomass production
- biomass partitioning

A definition of FSPM can be faound in Vos et al. [2007], in introduction to their book on the use of FSPM for predicting crop production:

> «Functional-structural plant models, FSPMs or virtual plant models, are the terms used to refer to models explicitly describing the development over time of the three-dimensional (3D) architecture or structure of plants as governed by physiological processes, the timing of which is determined by environmental factors. FSPMs are particularly suited to analyse problems in which the spatial structure of the plant or plant canopy is an essential factor contributing to the explanation of the behaviour of the system of study. Applications of FSPMs therefore include the study of plant competition (intra-plant, inter-plant, inter-species), and analysis of the effects of plant configuration and plant manipulation (e.g. pruning and harvesting) on the quantity and quality of the produce. »

For every FSPM, crucial questions to address concern the rules driving the set up of the new architectural units, biomass acquisition and partitioning. Their prior characteristic is to integrate effects of interactions between architectural and plysiological processes. For example, plant architecture can influence the functional processes through different ways:

- assimilate production depends in particular on light interception which is determined according to crown structure (for trees) or to leaf organization and orientation (for trees and herbaceous)
- water transport from roots to leaves depends on the hydraulic resistance of the branched network
- assimilate transport and partitioning depend on the relative positions of sources and sinks.

Reciprocally, effects of plant functional variables on architecture can be for instance:

- organ (e.g. fruit) abortion
- tillers or branches appearance
- organ senescence or branch death

In that case, the changes in the plant architecture are generally modelled according to thresholds of a functional variable (e.g. ratio of biomass supply over demand in GreenLab, Index of internal competition in Ecomeristem [Luquet et al., 2007]). For instance, if the biomass supply is less than the needs for assimilates for the maintenance of an organ, then it will not grow and possibly begin to be senescent.

Among the existing FSMs, we distinguish two approaches, each having a distinct applicative scope: (i) detailed models simulating the growth of young trees and of fruit trees, with simulated growth periods of no more than a few years; (ii) simplified models with efficient algorithms allowing longer simulated periods.
Models of the first type provide an accurate description of source-sink interactions within fruit trees or young trees at organ level. The biomass production of each leaf is usually computed on a hourly or daily basis according to the amount of intercepted radiations. Dynamics of allocation patterns are defined with various degrees of details; for instance using a transport-resistance model in L-PEACH ([Grossman and DeJong, 1994]; [Allen et al., 2005]), with a matrix of measured allocation coefficients for each source-sink pair in ECOPHYS [Rauscher et al., 1990], using fitted regression equations for allocation fractions in [Zhang et al., 1994] or including a function representing the source-sink distance in SIMWAL [Balandier et al., 2000]. Those models can address problems such as the effects of local environmental conditions on the growth of each organ individually or the analysis of branch carbon autonomy within a growing season (EMILION, [Bosc, 2000]). They incorporate numerous environmental and internal variables to represent the climatic control on physiological processes. Simulations are generally run for growth periods ranging from several months to a few years.
L-PEACH takes into account an index of water stress calculated according to root water uptake capacity in the soil in comparison with tree water demand (determined

according to the light exposure of each leaf). Bud fates in each zone of a shoot are regulated by the quantity of carbon available, using hidden semi-Markov chains to reproduce the observed stochastic patterns of Peach tree branching systems [Costes et al., 2007]. It accounts for local environmental and endogenous effects on the growth of each fruit, according to its location in tree architecture. Qualitative user-defined shapes of curves are introduced to relate the different physiological variables. The results are qualitatively satisfying but the authors underline that more experimental work is needed for calibration of the model, especially to investigate the values of carbon availability threshold for bud breakouts in each zone. After being calibrated, it aims at being used for decision support systems in horticulture.

SIMWAL simulates structural growth and carbon partitioning among organs of young walnut trees (no more than five years). Photosynthesis is computed at hourly time step and at leaf level considering direct and diffuse light from each azimuthal direction, crowns being simplified to ellipsoids when trees grow larger. A detailed algorithm of carbon partitioning includes processes such as respiration, reserves and adaptation of the demand to the current growth conditions. It account for a large set of environmental conditions (temperature, radiation and air CO2 concentration) and pruning interventions. Recently, the simulation platform PIAF-1 incorporated SIMWAL and offers more flexibility: its modular structure allows selecting the processes to include and different scales for spatial description [Lacointe and Donès, 2007].

The second type of models are more oriented towards forestry applications where the growth period of interest can largely exceed one hundred years. It raises specific problems that are different from those encountered when studying young trees or for horticulture. Only a few functional-structural models cope with the problem of tree simulation at that time scale. Rather than a detailed reproduction of tree structure, their models allow exploring the global effects of model assumptions and parameter variations [Sterck and Schieving, 2007]. They generally account for a more limited set of variables (e.g. C balance) and processes with a yearly time step.

In LIGNUM, the representation of tree architecture is based on elementary units (tree segments, branching points and buds) that were recently translated into L language [Perttunen et al., 2005]. Assimilate production is uppermost used for respiration and then distributed to different growing parts: new shoot segments, new root segments and thickening of existing shoot segments. Allocation is based on empirically-derived rules that aim at mimic functional balances (e.g. pipe model principles [Shinozaki et al., 1964]). The production of each segment is proportional to the amount of intercepted radiation coming from each sky sector. It is calculated according to the amount of transmitted radiations through the foliage of shading segments (if a beam hits a woody part, its propagation is interrupted). The length of new segments depends on the availability of photosynthetates, on local light conditions of mother segments and on its location in the crown, characterized by its branching order or by a vigour index calculated from the

ratio between cross-sectional areas at the initial branching whorl. LIGNUM was validated on coniferous (e.g. Scots pine [Perttunen et al., 1996] and broadleaved trees (e.g. Sugar maple [Perttunen et al., 2001]). Light interception through the crown with mutual shading of leaves and local light effects on architectural and morphological growth are the key points of LIGNUM.

Light influence is also the main environmental constraint introduced in the model developed by Sterck et al. [2005]. The probability of new shoot emergence is related to its branching order and to the level of daily light intensity; the number of new metamers is limited by the available carbon pool filled by photosynthetic production at the previous time step. The construction cost of a new internode comprises losses due to growth respiration and to the parallel construction of a woody pipe running from the leaf basis to the roots. The plasticity of tree shape is mainly due to adaptations of numbers of new metamers and not of their morphology.

Limitations of FSPMs

A drawback of FSPMs is that they become computationally heavy when a detailed architecture is considered. Even if several strategies (object-oriented programming, parallel processors...) can be used to reduce simulation times, it is still a limitation in many models. For instance, Isebrands et al. [2000] used a parallel computing strategy, called "component object model", to extend the scope of their model. They can simulate the first three years of a juvenile poplar growing under multiple stress (weather: CO2, O3, light, humidity, temperature; soil: water, nitrogen) with hourly computations of light interception and shading by individual leaves (up to 7000 leaves, the third year). But even when calculations are distributed on five parallel processors, the computational time is still about one hour. Sterck et al. [2005] limit artificially their simulations: they constraint the total number of leaves to be below 5000 to keep a reasonable simulation time (without specifying the value). To do so, they choose a very high value for leaf area to pipe cross-sectional area ratio: thus the high cost in assimilates for building the pipes reduces the growth of the tree (at most 5m tall).

Moreover, FSPMs still generally lack the same level of predictive ability under various environmental conditions as PBMs. The main efforts aimed at integrating functioning aspects into a complete description of plant structure and dynamics, but effects of environmental control variables were often neglected. So their application fields are currently restricted to research and teaching [Le Roux et al., 2001]. However, they have a lot of applications providing other disciplines with a framework: studies of biomechanical stresses, of light interception, of root growth, of dynamics of plant-environment interactions in heterogeneous environments [Sievänen et al., 2000], of irregular planting with different species and uneven-aged stands, of competition for resources, of silvicultural practices (thining, pruning)....

1.4.6 Positioning of GreenLab

In GreenLab, the plant 3D structure is computed using a dual-scale automaton, based on the concept of physiological age (see Chapter 2). It provides an efficient way of parallelizing the computation by factorization of similar parts of the plant: simulation of a 20 year-old tree with thousands of phytomers only requires a few seconds. It also provides a natural formalism to describe plant structure at each growth cycle in a very compact form. The botanical knowledge on ontogenetic trajectories acquired from the morphological analysis of the AMAP team can be integrated in a simplified way to define the potential architecture of a given tree species (see part II). GreenLab can be considered situated at an intermediate level between process-based models and functional-structural models. Plant 3D architecture is defined and represented but only partially used in the functioning part of the model. In GreenLab, architecture has an influence mainly for simulation of trees, for which hydraulic resistances of structure networks can be considered and partitioning of ring biomass for secondary growth is strongly dependent on tree topology. But for other functional processes modelled in GreenLab, only the numbers of organs are important and not their precise location. Indeed, the numbers of organs per category of physiological and chronological ages have effect on demand calculation for biomass partitioning. There is no effect of architecture in the new formulation of biomass production (see Chapter 3). GreenLab takes into account neither the detailed influence of architecture on the light interception and radiative balance nor the effects of source-sink distances for biomass partitioning. Dong [2006] compared two methods for computation of light interception in GreenLab for a tomato canopy: the equation of the Beer-Lambert law and an algorithm of detailed light interception using radiosity on a hourly basis. They concluded that both methods gave similar quality of results.
To conclude, GreenLab has inherited from the major concepts of morphological and process-based models and has incorporated them with strong simplifications to avoid the limitations cited above. Only a simplified representation of plant structure is considered and topology has influence on a limited number of functional processes (e.g. the biomass production of a plant is a simple function of its total leaf area; allocation to new organs only depends on their number per category).

1.5 A source-sink model

The core of the functional part of GreenLab relies on the principle of source-sink equilibrium, also called model of supply-demand in reference to economics models. In this paragraph, we focus on the modelling choices for the sink balance. The source function will be detailed in chapter 3. As the main interest of foresters is to predict the stand response to different forest management strategies, in particular to thinning, the allocation part of models is of burning interest. Moreover allocation is the most re-

active process allowing the plant to adapt to changes in environment [Deleuze, 1996]. Regarding crop growth, carbon allocation to seed or fruit compartment is also a major component of the yield (e.g. in [Yin et al., 2002]). However, this is often the least mechanistic part of models and there is no general consensus on how to model it [Le Roux et al., 2001]. Indeed, the biological processes involved have complex interactions at different scales and are difficult to measure independently from each other. Canell and Dewar [1994] state that «allocation is the outcome of many processes rather than a process in its own right ». Consequently, when browsing the different models under development, it is sometimes possible to find an assumption and the opposite one to drive the simulation of biomass allocation.
In GreenLab, the exchanged matter is the amount of carbon, which is the main component of photosynthetates and thus the main constituent of plant structure (80 to 90% of dry matter in a growing organ in standard conditions [Brouwer, 1962]). In most models, other elements are often considered as control variables, that is to say their presence can enhance or inhibit processes linked to carbon assimilation and distribution. Some organs, defined as sources, contain autotrophic cells (usually green due to the presence of chlorophyll), *i.e.* they have the ability to synthesize the essential molecules needed for growth and maintenance of the plant. Assimilates are partitioned between demanding organs, called *sinks*, that are heterotrophic organs, relying on supplies from the sources. They represent expanding organs of the plant. The quantity of assimilates that they can get is determined according to a coefficient called sink strength. But apart from sink-based models, other methods have been developed to represent the allocation processes in plants.

1.5.1 Different approaches to model allocation

In several reviews on carbon allocation in functional-structural tree models ([Deleuze, 1996], [Lacointe, 2000], [Le Roux et al., 2001]), authors have classified the different methods in four categories:

- empirical allocation coefficients
- functional balance or growth-rule based models (teleonomic models)
- transport-resistance models
- interactions among sinks: hierarchical or relative sinks.

Among those four methods, two very different approaches, in their philosophy, can be distinguished: one focuses on the mechanisms of allocation and the other one on its results. The first class consists of transport-resistance models, including diffusion-reaction models. They are derived from the Münch theory that describes the physiological basis of flow transfers in phloem and in xylem. They aim at simulating carbon allocation

as an emergent property of transport mechanisms. As underlined by DeJong (1999), «dry matter partitioning does not direct the growth of the tree but is the result of the growth and development of the organs that make up the tree ». On the other hand, the three other categories of methods (empirical allocation coefficients, teleonomic models, sink competition models) focus on the result of allocation and do not account for its physiological causes. They deal with allocation by imposing growth rules, allometries or sink coefficients to drive the partitioning. They are more empirical than models of the first class, as they rely on observations of the growth results under given environmental conditions.

Most models need to be positioned for those three key questions:

1. origin of assimilates: (a) individual sources (taking into account source-sink distances) vs (b) common pool
2. assumptions for assimilate allocation: (a) priority or hierarchical vs (b) proportional vs (c) transport-resistance (mechanistic)
3. assumptions for definition of sink dynamics : (a) saturable vs (b) unlimited.

Only teleonomic models escape those questions. This approach is also defined as evolutionary as it considers that plants optimize different growth compartment according to environmental constraints. Mäkelä [1990] presents the assets of the functional carbon balance, the pipe model theory and the height growth game. The author privileges this approach to that of transport-resistance and sink-source models, stating that «removing the problem of allocation, however, this method appears to have created one of demand». We do not detail the teleonomic approach here but instead, we prefer to discuss the consequences of the different choices in the related assumptions presented above. In GreenLab, the choices are: (1b) common pool, (2b) proportional allocation and (3b) unlimited sinks.

Note that Heuvelink and Marcelis [2007] add a fifth category that relies on canonical modelling concepts. Canonical modelling is described in [Renton et al., 2005] as an appropriate framework for model development, especially in the case of functional-structural models. It consists of representing the interactions (fluxes or controls) between the main compartments. The construction of this flowchart mainly relies on qualitative knowledge and thus appears as a modelling approach at an intermediate level between empirical models and mechanistic models. Variations of their associated state variables can be driven for example by the generalized mass action system [Renton et al., 2007] with different possible shapes for the mathematical expression of fluxes (e.g. power functions). Only global outputs are fitted against measurements. Thus it does not require a detailed knowledge about underlying processes that would need to be calibrated, as it is classically done. In that sense, this approach is similar to that

of GreenLab. However, they need to be linked with an external representation of the plant architecture, using for example the L-System formalism [Renton et al., 2005].

1.5.2 Physiological bases of biomass allocation

Transport of carbohydrates from their sites of synthesis (mainly in mature leaves) to their sites of utilization and storage takes place in the sieve tubes of the phloem (external side of the cambium). The set of processes of loading and unloading of the phloem as well as carbon transfers is called *translocation*. We give here a simple description of translocation physiology, oriented to a modelling point of view (in particular, we do not evoke the influences of hormone signals and regulations).

Phloem loading

Photosynthesis produces glucides (triose phosphates) that are either transformed into starch and temporarily stored in the leaves or transported under soluble form (sucrose) to the phloem. The regulation between starch and sucrose is what physiologists usually define as "allocation" ([Deleuze, 1996]). Starch storage allows the plant to maintain a relatively constant loading of the phloem despite the day/night alternation and short-term environmental fluctuations. Studies of allocation between starch and sucrose under different conditions suggest that a fairly steady rate of translocation throughout the 24-hour period is a priority for most plants [Taiz and Zeiger, 2006]. Consequently, in most models, time steps are long enough (more than several days) to neglect this temporary storage.

Phloem unloading

Phloem unloading is related to carbohydrate consumption at sink sites. Those carbohydrates are partitioned within cells between maintenance and growth respiration. Carbohydrate requirements of a sink greatly influence phloem flows into it. The notion of sink emerged in the 1970s to define the capacity of a growing zone to get carbohydrates when competing with other growing zones of a plant [Farrar, 1993]. Although this notion has been commonly used for decades, its expressions in models vary depending on authors. Warren-Wilson [1967] defines it as the mass of sink tissues times sink activity (its rate of uptake of photosynthates per unit mass of sink tissue):

$$Sink\ strength(g.d^{-1}) = sink\ size(g) \cdot sink\ activity(g.g^{-1}.d^{-1})$$

Sink size can be considered as the number of active cells. For example, fruit development proceeds in three phases: cell division, expansion and ripening. From their experiments under limiting or non-limiting conditions and comparing the results at different stages in the plant development, Bertin et al. [2001] report that cell division is affected by

the level of assimilate competition among sinks. The other term of sink strength, sink activity, is more difficult to define and may include various processes: unloading from the sieve elements, metabolism in the cell wall, uptake from the apoplast and metabolic processes that use photosynthates in either growth or storage. Sink activity is also related to the presence of sucrose-splitting enzymes that catalyze the first step in sucrose utilization. The abundance of this enzyme is driven by a gene whose expression is regulated by carbohydrate supply (*i.e.* the more carbohydrate supply, the more enzymes enhancing sink activities) [Taiz and Zeiger, 2006].

Transfer in the phloem

Translocation is driven by the mass-flow principle that was described by Münch in 1930 after his famous experiment (Figure 1.1). It states that flow speed in the phloem is driven by osmotically-derived hydrostatic gradient pressure.

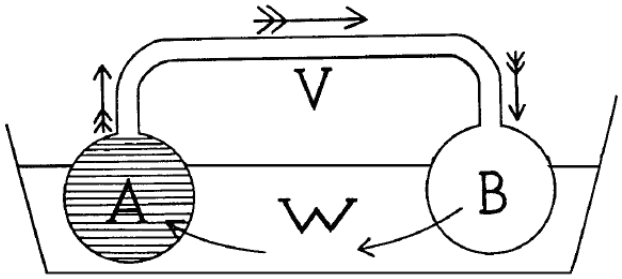


Figure 1.1: Münch experiment

Two cells with semi-permeable membranes (permeable to water, impermeable to sugars), denoted A and B, are placed in a recipient full of water (W) and linked by a tube (V). At the initial state, cell A is full of sugars and cell B is full of water. This difference in sugar concentrations creates a gradient of osmotic pressure that induces a mass flow from A towards B in the tube and in parallel, a hydrostatic pressure gradient that induces a water flow from B to A in the recipient. The same mechanisms occur in the sieve elements of the phloem: the concentration gradient is maintained by assimilates production by the sources at one end and unloading by sinks at the other end. Thus the phenomenon is automatically regulated by the sink and source activity that generate the pressure gradients [Taiz and Zeiger, 2006]. However, an exact and complete formulation of the Münch theory implies complex equations with often instable resolution schemes (see [Deleuze, 1996] for a commented review of models based on the Münch theory). Nevertheless, they have confirmed that a good compromise was to describe translocation using a diffusion model, where the flow rate J is yielded by the Poiseuille-Hagen law:

$$J = k \cdot P \tag{1.1}$$

where k depends on the pipe radius a, on the fluid viscosity η and on the tube length L ([Lacointe, 2000]). P is the osmotically generated hydrostatic pressure that is related to solute concentration C using Van't Hoff's equation (related to the ideal gas law):

$$J = C/R \tag{1.2}$$

with

$$R = 8 \cdot L \cdot \eta/(\Re \cdot T \cdot \pi \cdot a^4) \tag{1.3}$$

where $\Re$ is the gas constant and T the temperature. The diffusion model provides the same long-term behaviour as the Münch theory (in steady state), but short-term reactions are not considered and the resistivity parameter has to be empirically estimated: it has no direct physiological reality since diffusion is not the real phenomenon. Moreover, measurements of phloem flows are difficult to realize experimentally. They require in vivo and non-invasive methods, as sieve tubes are very sensitive to mechanical probing ([Deleuze, 1996], [Minchin, 2004]).
Note that in fact, differences in solute concentrations between sources and sinks are not the only driving force for phloem transport. Numerical simulations of Münch's basic hypothesis have shown that it cannot account for the observed flow speed in files of sieve tubes longer than several meters. There is continuously and simultaneously leakage and reloading all along the pathway ([Minchin, 2004]). This is associated with secondary growth that acts as a boost motor for translocation along the phloem. And of course, not only physical signals (turgor pressure) but also hormonal signals affect source-sink relationships.

1.5.3 Discussion on the transport term

Although the transport-resistance methods are conceptually justified and in adequacy with the real phenomena, they are generally too complex to be fully integrated in structural-functional models. They have been applied in models at compartment levels for instance in [Berninger et al., 2000] or [Deleuze and Houllier, 1997]. Regarding plants with complex architecture, Nikinmaa et al. [2004] applies a 1D reaction-diffusion scheme to a young branching tree and uses finite-element methods to solve it. But he recognizes that it is computationally heavy and does not detail how to parameterize it. One of the most achieved approach is implemented in the L-PEACH model [Allen et al., 2005]. Plants are represented as networks of sources and/or sinks connected by conductive elements, the driving force being the osmotically generated concentration gradient, as described by the Münch theory. Using an analogy with an electric network to build an equivalent circuit (folding phase), the carbon flows at each time step are calculated (unfolding phase). The behavior of the resulting system is time-dependent and highly non-linear. The equations of carbon flows in each component of the plant network are solved using the L-System formalism.

In all cases, the Achilles' heel of those methods remains the determination of sink activity. As, finally, phloem unloading plays a great part in the regulation of sucrose transfer, sink activity is the major actor. But it is still difficult to determine empirically and it is often done under strong physiological hypotheses (see for instance the material and methods part in [Bancal and Soltani, 2002]). Most importantly, the necessity of taking into account resistance terms in allocation modelling is still under question, especially for crops. Thornley (1998) states that transport-resistance principles should be included in every process-based model as it takes into account the «only two significant processes: transport and biochemical conversion ». However, the relative influence of each process needs to be investigated to assess the significance of each of them. There are evidences, both experimental and theoretical, that it may not be necessary to include transport terms in allocation models. From their experimental results on tomatoes, Slack and Calvert (1977) conclude that dry matter partitioning is influenced by pathway resistances. But Heuvelink [1996] re-examined their interpretation at the light of his model (TOMSIM) simulations and suggested that the observed data could be explained by the temporal shift of truss growth, without necessary invoking the effects of transport resistance. Indeed, the importance of phenology is underlined in [Wardlaw, 1990]. The appearance order of new sinks and the timing of their development are key factors when calculating the demand of the plant in an organ-based model [Letort et al., 2007a].

In [Heuvelink, 1995], the author reports the results of experiments with different pruning strategies in tomato plants to investigate the influence of source-sink transport. He compares double-shoot plants for which the fruit proportions are 100-0 (no fruits on one shoot, all fruits on the other shoot) and 50-50 (half of the fruits was left on each shoot). He concluded that, at least concerning assimilate allocation to fruits, the experimental results support the assumption of common pool and that of a negligible effect of transport resistance. In [De Reffye et al., 2008], the authors show the example of the growth of a non photosynthetic shoot of *Hedera helix*. The size of the non photosynthetic leaves is similar to that of photosynthetic ones. It means that the source leaves supply in the same proportions their own shoot and the 'sink' shoot. It reinforces the theory that distance has a negligible effect. Bancal and Soltani [2002] pointed out that, in every transport-resistance model, resistance should be dependent on the fluid viscosity and hence on its concentration. They corrected the model of [Minchin et al., 1993] in that way and studied it both numerically and with parameters from measurements on wheat grain filling. They compared the partitioning coefficients output in their corrected model to the former version of the model and to the classical sink-based partitioning model. They showed that the resistance to flux propagation has an influence on the partitioning coefficients only in pathologic cases of very low source activity, that is to say nearly dying plants. This means that in most cases, flows are mainly driven by sinks. They concluded that resistance terms in model could be abandoned in most cases as they are only a mathematical burden whose parameter values are very difficult to measure experimentally. But is must be underlined that their study

concerns only a very simple sink-source network and would be worth being extended to more complex cases.

1.5.4 The common pool concept

All these evidences lead to the conclusion that it is possible to build a model where carbon partitioning is mainly driven by sinks and where the effects of source-sink distance are considered negligible, at least for crops. From a pure mathematical point of view, assuming that the transfer resistance is equal to zero implies that all sinks have a direct access within the step of growth to each source. This has for immediate consequence the assumption of a centralistic carbon pool. But in some models (e.g. SIMWAL [Balandier et al., 2000], ECOPHYS [Lacointe et al., 2002]) sink strength is defined as a property of the couple (source i, sink j) as a decreasing function of the distance between i and j. This definition is biologically relevant but we can wonder whether the main limiting factor is geometrical distance or the topological organization of source and sinks (*i.e.* the number of other sinks in a source-sink pathway).
Wardlaw [1990] states that distance between sources and sinks is not a major factor in limiting the growth. Experiments on apple and soybean have shown that increasing the distance from source leaves to the fruits induced no significant reduction of the fruit growth rate. Wardlaw [1990] also reports from the experiments of Palit (1985) with ^{14}C that a sink is generally supplied with photosynthetates from the nearest sources: the top leaves supply carbon to the apical bud and young growing leaves, the central leaves supply carbon to the stem and the basal leaves supply carbon to the lower stem and the roots. This is confirmed by the coefficients of the allocation matrix of the ECOPHYS model that controls the assimilate partitioning in a young poplar [Rauscher et al., 1990]. Following displacements of assimilates issued from each source leaf, they reveal a preferential local allocation. The experimental results of Pallas et al. [2008] also suggest that the assumption of a common assimilate pool does not hold for grapevine. Their pruning experiments revealed that cluster weight is more affected by its distance to the main assimilate sources than by the global level of trophic competition in the plant. Wardlaw [1990] reports that large sink dominate the overall supply and that smaller sinks need to rely on local supplies. This could mean that what is important is the topology of sink organization (organ relative positions) more than the geometry (distances between organs).
As assimilate flows in the plant network are driven by osmotic gradient, it is reasonable to assume that, at each intersection, the assimilate proportion that flows into each branch is proportional to the demand of each terminal substructure. The assimilate proportion kept by the leaf depends on its attractive ability but also on the demand load of the rest of the plant (generating a potentially high osmotic gradient that can reduce the assimilate proportion stored in the leaf). The same principle is applied to assimilate allocation to other organs. The plant can be seen as a network where

organs are represented by taps, as in [Wardlaw, 1990] (see his figure 5 page 359). The assimilate flow generated by a source leaf is thus partitioned into each structure proportionally to its cumulative sink. Thus after successive simplifications of the ratios at each intersection, the quantity of biomass attributed to each organ is equal to its sink multiplied by the ratio of biomass production over the total plant demand. When several sources are considered, each organ receives the same proportion from each source leaf:

$$q_i = \sum_{sources\ k} c_i \cdot \frac{Q_k}{D} = c_i \cdot \frac{Q}{D} \tag{1.4}$$

where Q is the plant production, defined as the sum of the contributions of each leaf, D the plant demand, defined as the sum of the contributions of each organ demand (c_i). So this way of considering biomass allocation is equivalent to the equations derived from the common pool assumption concept (or centralistic concept, as denominated in [Kurth, 1994]). Under those assumptions, it is not necessary to consider independently each source (hypothesis (1a)) as the result can be simplified to the common pool hypothesis (hypothesis (1b)).

From the same initial assumptions, another choice for sink definition would be to consider c_i as the proportion of biomass that can be taken by an organ from the flow at its borders, regardless of the demand of the rest of the plant. It means that sinks would be considered as absolute instead of relative values. In that case, the equations cannot be reduced to the same kind of simplified form. The result would be a hierarchical model (hypothesis 2a), with sink priorities depending on their positions relatively to the source. It allows introducing reserves to store assimilates when the demand does not match the supply. But it implies that leaves would have the highest priority as they would be necessarily the first ones to receive a given proportion of the biomass they produce. This proportion being independent of the changes of the rest of the structure, the resulting model would be not as flexible to model the plasticity of the plant response to environmental or internal changes. However this priority concept has been used for models of assimilate allocation, but mainly at compartment scale, to give priority to some vital processes (e.g. respiration as in most models including this process [Le Roux et al., 2001]). At compartment levels, the priority orders are generally accepted as:

$$Seeds > fleshy\ fruit\ parts = shoot\ apices\ and\ leaves > cambium > roots > storage$$

[Minchin and Lacointe, 2005]. Several models mix hierarchical models (in a first step) with proportional sub-models (for intra-compartment partitioning). According to the common pool assumption, source organs contribute (but not necessarily in the same proportions) to providing assimilates to all organs of the plant, including themselves. As each leaf has both source and sink roles, it implies that, paradoxically, assimilates required for its growth do not come mainly from the leaf itself but from all the active leaves of the plant. Although this is not biologically relevant at this scale, it is justified at a higher scale, as shown above from mathematical and experimental evidences. So

here again we see that the choice of modelling concepts is strongly dependent on the investigation scale: the common pool assumption is wrong at microscopic scale but justified when the integrated result at mesoscale is considered.

Regarding trees, the same conclusions cannot be directly drawn: influence of source-sink distance needs to be taken into account. «Decrease in C flux between source and sink with increasing pathway length, [is] a major characteristic within trees.»([Lacointe, 2000]) Kurth [1994] opposes the concept of common pool to the concept of *branch-autonomy*. It assumes that newly produced assimilates are locally consumed in their source shoot or in the organs located on the path downwards to the roots (Figure 1.2). In particular, that concept was adopted in [de Reffye et al., 1993]. The authors showed that this mechanism was similar to the pipe model and in accordance with the Pressler law (relating the sapwood sectional area to the above foliage quantity).

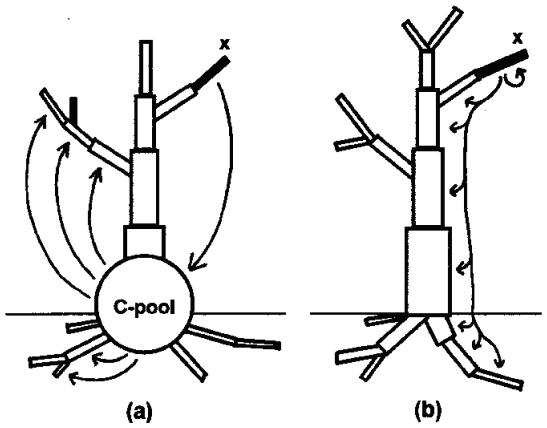


Figure 1.2: Common pool vs Branch autonomy concepts, from [Kurth, 1996].

However, the assumption of common pool can still hold as a modelling simplification. It only requires a definition of sink strength depending on the organ position in the tree topology. In GreenLab, this is done for the cambial growth demand of each internode (see Section 3.2.3) and the same method could be easily applied to other kinds of demand.

1.5.5 Definition of sink strength

Besides the modelling choices for the mechanisms of assimilate allocation, another key point is to determine the shape of sink strength variation. So, finally, we keep raising the problem of sink definition and determination as the major factor in allocation models. Even for transport-resistance models, the sink shape evolution is determinant (e.g. chosen as a Michaelis-Menten function in [Minchin et al., 1993]). Here again

we meet the problem of choosing between a definition that would be more biologically based and the practical applicability for modelling. Authors generally refer to the macroscopic results of biomass allocation: they define the sink of an organ according to the quantity of biomass that it can get for its incremental growth, without considering the underlying biochemical processes such as cell division and metabolism.

An important choice to make is to decide whether sinks should be unlimited or saturable. Indeed, it is observed in many cases that growth of an organ is limited by the saturation of its storage capacity. For instance, Heuvelink [1996] noticed that the fruit growth rate for tomato were the same when all fruits were pruned except one per truss and when two fruits were left. They interpret the result by assuming that the growth of the fruits is sink-limited and that they have reached their maximal potential growth rate. However, it is difficult to distinguish between sink saturation and organ size saturation. Defining saturable sinks allows a natural introduction of a reserve pool to store the quantity of assimilate supply that exceeds the demand. Those assimilates are available for the next growth step. The reserve pool plays a role of passive buffer to regulate the variations of the ratio of supply over demand. But the definition of maximal growth rate is likely to be dependent on the source supplies that can vary in time during organ growth. Organ maximal size has the advantage of being easily measured: it is the maximal size reached by the organ in the most favourable conditions. This maximal organ size would result from physical and mechanical limitations. For instance, experiments on sunflower (see [Rey, 2003, chap. 2]) have shown that when the cap is suppressed, some organ dimensions (leaf area, stem diameter) do not differ from the measurements on an isolated plant. It can even induce explosions of some superficial tissues.

In the definition of [Warren-Wilson, 1967], sink strength is defined as the mass flow per time unit into a sink. This definition is the first one attempting to give some biological foundations to the notion of sink strength, although the way to measure sink activity remains unclear. But mass flow is not an intrinsic property of the sink but a function of carbohydrate supply, transfer pathway and the competition of other sinks. So this definition is not adequate. Patrick [1993] defines it as the potential import into a sink under non limiting conditions. This definition has been since adopted in several models. For instance, Heuvelink [1996] defines the sink shape as the first derivative of a Richards function. It is measured in non-limiting supply conditions, assuming that the growth rate is the maximal possible one. Then biomass is distributed to each organ in proportion to its potential growth rate. If the supply is higher than the sum of the potential growth rates, biomass is stored and available for the next growth cycle. But that maximal import may be difficult to determine experimentally and remains dependent on the environmental conditions. Moreover, the potential growth rate is difficult to disentangle from the variations of the assimilate supply, related for example to variations of the leaf area index.

To overcome these difficulties raised by experimental measurements of sink activity, GreenLab can provide an alternative approach. By fitting the model on morphological

data, it extracts the sink dynamics of individual organs. In this context, Greenlab can be seen a source-sink solver (see the discussion part in [Dong et al., 2008]).
To summarize, GreenLab sinks are relative and get biomass from a common pool of assimilates and no possible saturation is considered. No potential or maximal growth rates are defined but a boundary to the growth of an organ can be set by defining a maximal organ size if needed.

Chapter 2

Organogenesis and topological rules

In the following chapters, we present the current developments of GreenLab. A very complete description and analysis can be found in [Mathieu, 2006]. We focus here on the new features that have been introduced during the phd.
We call *organogenesis* the series of processes resulting in the creation of new organs from the functioning of meristems. The term *apical meristem*, or apex, refers to meristematic cells in a bud located at the tip of a stem. The spatial organization of organs forms the tree structure or its *topological structure*. More precisely, describing the topological structure of trees consist in their decomposition into elementary constituents and in the identification of their connections.

2.1 Modular description of plants

A plant can be considered as a population of basic structural units, called metamers or phytomers [White, 1979]. It is the entity formed by a node, its associated leaf (or leaves), its axillary buds, the subtending internode and potentially roots (Figure 2.1).
Modelling of tree architecture was initiated in 1978 by the botanists Francis Hallé and Rudolphe Oldemann [Hallé et al., 1978]. They introduced the principles of architectural organization in trees, as a result of the growth processes: rhythmic or continuous growth, apical or lateral flowering, axis differentiation. They identified 23 architectural models that represent the different classes of tree development. Their work is based on the description of plants as modular organisms that develop by the repetition of elementary botanical units [Barthélémy and Caraglio, 2007]. Plants can be decomposed into a set of constituents that have common nature and properties. As most plants share the same basic growth processes, their structures reveal two main kinds of modularity: nodal modularity, resulting from the apical process and axial modularity, resulting from the branching process [Godin and Caraglio, 1998]. Describing the topological structure of a plant means identifying its elementary constituents and the connections between them. Natural ways of multi-scale decomposition have been identified by botanists.

They are based on observations of repetitive structures that have similar patterns at different scales in tree architecture. They also rely on morphological markers (cataphylls, nodes, scars) to separate elementary units. Morphological markers can give some clues about the past development of a plant. It allows extrapolating the dynamics of development from static measurements at a given date. Barthélémy and Caraglio [2007] have reviewed four main features used as criteria to describe architecture. Some of these notions will be referred to in the following parts of the manuscript; they are briefly presented in Table 2.1.
Each species has a genetic predisposition to some of these features, leading to particular privileged architectural strategies, but it is also affected by the environmental conditions. Architectural models provide criteria to define a typology of plants. The architectural model is the result, visible at the whole-plant level, of a particular hierarchical organization and repetitive sequence of the same structures, named *architectural units* [Barthélémy et al., 1997]. Indeed, axes in a plant can be sorted according to their morphological or functional characteristics. Each category is produced during a particular phase in the meristematic development, called its differentiation state. This phase is characterized by the notion of *physiological age* (PA) of the apical meristem. It is determined a posteriori according to a morphological analysis of the botanical entity produced by the meristem. For most plant species, five or six physiological ages are sufficient to cover all the categories of axes observed in a tree. For instance, architecture of *Cedrus atlantica* (Pinaceae) is composed of trunk, branches, branchlets, twigs and brachyblasts ([Sabatier and Barthélémy, 1999]).

2.2 Dynamics of these modular structures: growth unit and growth cycle

In GreenLab, the time step for simulation is not determined *a priori* but is based on the specific development of the plant considered. As plant topology is modular, plant growth can be seen as a cyclic phenomenon whose period corresponds to the duration of the set up of a new module. That period is called *growth cycle*. As stated above, plant growth can be either rhythmic or continuous. Rhythmic growth is defined by the alternation of periods of activity of the apical meristems and of resting periods. The active period corresponds to a growth flush and creates a *growth unit* (Figure 2.1). The set of growth units produced during one year forms the annual shoot. For trees, the case of polycyclism (definition in Table 2.1) is not considered in the current version of GreenLab. Hence the growth cycle lasts generally one year for temperate trees. The choice of growth cycle as time step simplifies the model as it allows neglecting some short-time variations such as high-frequency variations of environmental variables.
For plants with continuous growth, the growth cycle can be defined as the duration (called plastochron) between the initiations of two successive phytomers. It has been

Growth processes	
Determinate	The apex aborts or transforms after some period of functioning
Undeterminate	The apex indefinitely maintains its growth potential
Rhythmic	Shoot growth has marked endogenous periodicity
Continuous	Shoots show no marked cessation of extension
Preformation	All organs of the future elongated shoot are present at an embryonic stage in a bud before the elongation of the shoot
Neoformation	More organs than those included in the bud are elongated
Monocyclism	The annual shoot consists of one growth unit
Polycyclism	The annual shoot is made of a succession of several growth units
Branching processes	
Terminal	The apical meristem directly splits into to two or more new axes
Lateral	The embryonic cells are located just aside the initiated leaf (axillary meristems)
Immediate	The lateral axis elongate immediately after meristem initiation (syllleptic)
Delayed	The lateral axis elongate after a phase during which the lateral meristem remains inactive (proleptic)
Monopodial	Branching pattern associated to undeterminate growth
Sympodial	Branching pattern associated to determinate growth: several branches may develop after the death or transformation of the apical meristem. [1]
Branching order	Ordinal numbers used to define a branched system: the main stem is of order 1; lateral axes it gives rise to are of order 2 and so on.
Acrotonic	Prevalent development of lateral axes in the distal part of a parent shoot
Basitonic	Prevalent development of lateral axes in the median part
Morphological differentiation of axes	
Orthotropic	General orientation of axes is vertical with radial symmetry
Plagiotropic	Axes have a general horizontal to slanted orientation with bilateral symmetry (leaves and branches generally arranged in one plane)
Position of reproductive structure	
Lateral	The reproductive structure forms from a lateral meristem
Terminal	The reproductive structure forms from a terminal meristem

Table 2.1: Table of botanical architectural concepts, summarized from [Barthélémy and Caraglio, 2007]

observed that the number of initiated phytomers depends linearly on cumulated thermal time. Cumulated thermal time (CTT) is defined as the sum of the average daily

[1] Even if not edified by the same meristem, a sequence of such lateral branches may be considered as an axis.

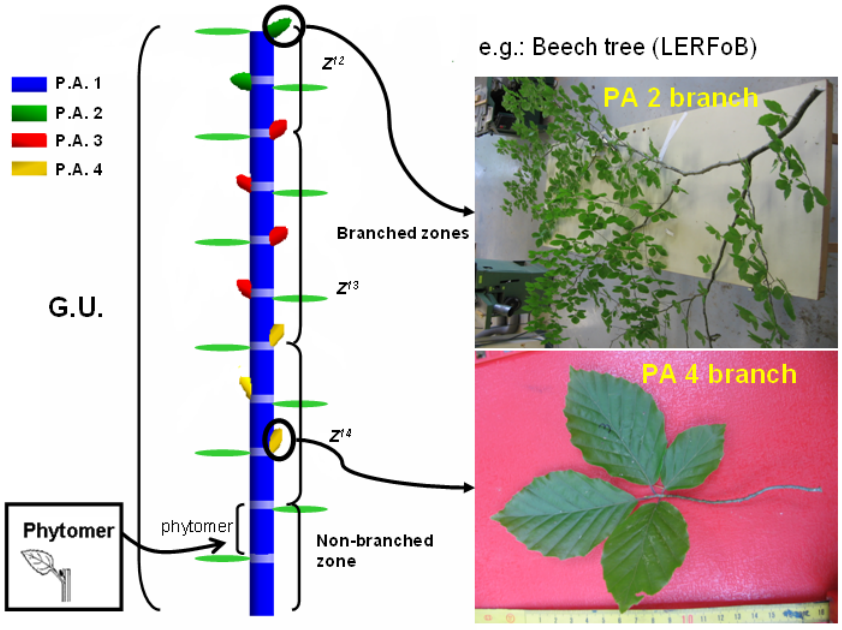


Figure 2.1: Schematic representation of a growth unit (G.U.) of physiological age 1 (PA) with axillary buds of PAs 2, 3 and 4. A zone Z_{pq} is characterized by its PA p and the PA q of its axillary bud (see paragraph 2.4.2).

temperatures between two dates. A reference value, called base temperature (T_b), is subtracted to the average daily temperature $T(d)$. More precisely, cumulated thermal time from day 1 to day n can be calculated as:

$$CTT(n) = \sum_{d=1}^{n} (T(d) - T_b)^+ \tag{2.1}$$

with the usual definition of function $^+$ (takes the positive values of the argument and zero everywhere else). The base temperature depends on the species and on the experimental conditions. For instance Granier et al. [2002] set 3°C as the base temperature for *Arabidopsis thaliana*. However, the determination of plastochron is often experimentally tedious because dissections under binocular microscope are required to count the number of primordia present in apical meristems. Often this number cannot be determined with accuracy due to the small size of leaf embryos and to the necessity of choosing an arbitrary limit size below which a leaf is not considered. Furthermore, it is difficult to determine the number of primordia initially contained in the seed. To overcome these difficulties, growth cycle can be defined according to the phyllochron, that is to say the duration, in cumulated thermal time, between the expansions of two successive phytomers. It is also ambiguous because the number of leaves that have appeared depends on the limit size chosen. However, as the model considers biomass

partitioning, a leaf really needs to be taken into account in the carbon balance only when it is a non-negligible sink or source. So using a phyllochron basis to define the growth cycle can also be a relevant choice. In GreenLab, the time step for simulation is based on organogenesis rate and is common for all the processes. It is kept constant even when organogenesis has ended. However, it could be easily dissociated for the functional part if for instance daily computations of photosynthesis and organ expansion are needed. The number of growth cycles spent since plant emergence is called its *chronological age*. More generally, the age, expressed in growth cycles, of an organ or an axis since its initiation is defined as its chronological age.

2.3 Rhythm ratio for development rate

2.3.1 Development rate at different scales: axis, plant, stand

For some plants, the emission rate of phytomers is not constant. Firstly, it can change temporally, during the growth, after particular events of the plant phenology. For instance, in *Arabidopsis thaliana*, the speed of phytomer emission increases drastically after inflorescence induction (more than 2 times higher than during the first growth phase, for plants grown in standard non-limiting conditions in growth chamber [Christophe et al., 2008]). In rice, phyllochron is shorter during the reproductive phase than during the vegetative phase. Secondly, phytomer emission rate can also change spatially in cases when some axes have a different development rate than the rest of the plant. For instance, cotton tree sympodial branches have a development rate about half of that of the main stem. During the interval needed for initiation of two phytomers of the main stem, only one phytomer is initiated on each branch. Thirdly, phytomer emission rate can be different from one plant to another of the same stand, even if their sowing and emergence dates are similar (e.g. sugar beet). To simulate the stand growth, it is necessary to synchronize the schedulers for each plant, so that the incident radiations during one growth cycle are the same. Thus, if two plants exhibit the same leaf area, their biomass production per growth cycle is the same. The inverse method allows fitting the growth rhythms for each plant. Thus it is possible to simulate a stand with a distribution of different dates for plant emergence and a distribution of growth rhythms. Those parameters have a strong influence on the further competitive capacity of each plant.

2.3.2 Simulation of rhythm ratio

This variation in development rate is simulated using a parameter called rhythm ratio. The growth cycle duration is defined arbitrarily by choosing the plastochron or phyllochron value of a given growth phase or a given axis (usually the main stem). If the rhythm ratio is higher for branches, it means that several phytomers appear during one

growth cycle. On the contrary case (rhythm ratio less than one), the apical meristem undergoes pauses: there is no phytomer production at some growth cycles. It must be underlined that the duration of the growth cycle, based on thermal time relationship, does not change. The time step of the model is constant throughout the whole simulation. If the growth cycle duration would vary, it would imply that the cycle duration δt should be taken into account in the biomass production equation (multiplied by δt instead of 1) and changes in the sink variation functions.
The number of phytomers produced at each growth cycle is an integer and needs to be computed so that a given rhythm ratio is reproduced. This is done according to the principle of Bresenham's line algorithm [Bresenham, 1965]. This algorithm was designed to determine which points in a n-dimensional raster (e.g. pixels in a screen) should be plotted in order to form a close approximation to a straight line between two given points. We use a similar method to compute the discrete values of numbers of metamer that must be initiated to approximate the real rhythm ratio (see figure 2.2). For example, if the rhythm ratio is less than one, it is possible to represent the axis development as a series of 0 and 1. A simple way is to say that the meristem makes a break when at growth cycles i verifying $\lfloor w \cdot (i-1) \rfloor = \lfloor w \cdot i \rfloor$ where w is the rhythm ratio.

Remark: a more accurate expression would be: $\lfloor w \cdot (i-1) + 0.5 \rfloor = \lfloor w \cdot i + 0.5 \rfloor$ *but as we are only interested in the average slope (the rhythm ratio) and not on the actual values of the series, the former one is as suitable as this one.*

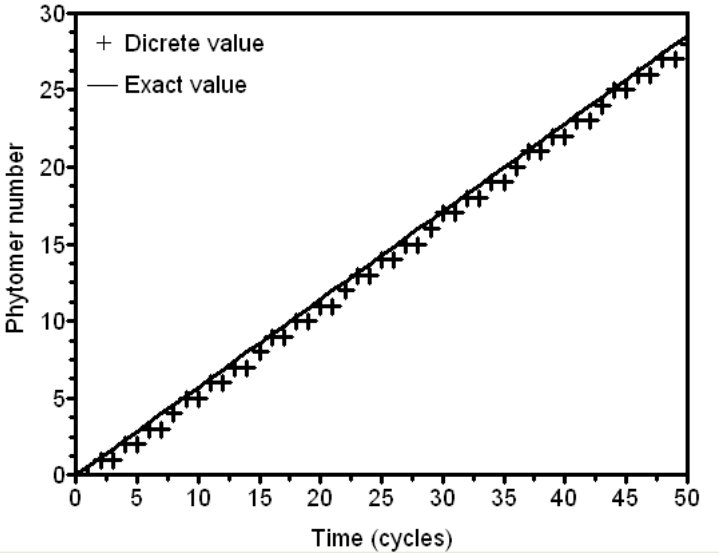


Figure 2.2: Simulation of number of phytomers on an axis with rhythm ratio $w = 0.57$. The corresponding sequence for phytomer emission at each G.C. is: 0.1.0.1.0.1.0.1.1.0.1.0.1.0.1.1.0.1.0.1.0.1.1.0.1.0.1.0.1.1...

2.4 Modelling plant development: formalism

2.4.1 A dual-scale automaton with different kinds of transition rules

The plant topology dynamics is described through the fate of lateral and apical buds on each phytomer. They can give birth to shoots with a set of parameters depending on their location and on the date of their emergence. The set of parameter to use is identified by the physiological age of the bud. GreenLab has inherited the basis concepts developed in AMAPsim software [Barczi et al., 2008]. The sequence of every potential state possibly taken by the meristem forms the reference axis. The topological rules are defined through a finite state automaton implemented using a semi-Markov chain where the occupancy probability distribution follows a binomial law. In GreenLab, the same state variable (physiological age) is used in a dual-scale automaton as introduced in [Zhao et al., 2001]. The basic concepts are the same as in L-systems, where the evolution of the plant topology corresponds to parallel rewriting of a string of symbols [Prusinkiewicz and Lindenmayer, 1990]. The microstate is defined as a phytomer with a given physiological age and bearing potential buds of a given physiological age. A macrostate, representing a growth unit, consists of the repetition of similar microstates and of the succession of different microstates. Inside a given macrostate, all phytomers have the same physiological age and only differ from the physiological ages of their lateral buds. The succession of macrostates forms axes of the plant (see figure 2.3). As the physiological age of an apical or lateral bud is usually higher than that of its bearing phytomer, the automaton can be characterized as a left-right automaton. If the physiological age of a branch is the same than that of the mother axis, the branch is called a *reiteration*. Depending on the version of GreenLab, the transition rules can be deterministic (GL1, [Yan et al., 2004]), stochastic (GL2, [Kang et al., 2003]) or function of the source-sink ratio (GL3, [Mathieu, 2006]). But other kinds of dependencies could be considered: the branching and transition rules could be determined as functions of the time to reproduce for instance the changes due to ontogeny or they could be determined as probabilities with thresholds depending on the source-sink ratio (included in the GL4 version currently under development).

The version of GreenLab to use is determined according to the plant species studied and to the objectives of the modelling. For instance, maize (*Zea Mays*) has a determined growth with a nearly fixed number of organs: the GL1 model is sufficient to deal with this cultivar [Guo et al., 2006]. On the contrary, coffee trees (Coffea robusta) reveal a stochastic development and need to be studied with a probabilistic approach [de Reffye, 1979]. For cucumber plants, a specific objective is to analyze fruit abortion and thus the GL3 version is appropriate [Mathieu et al., 2007].

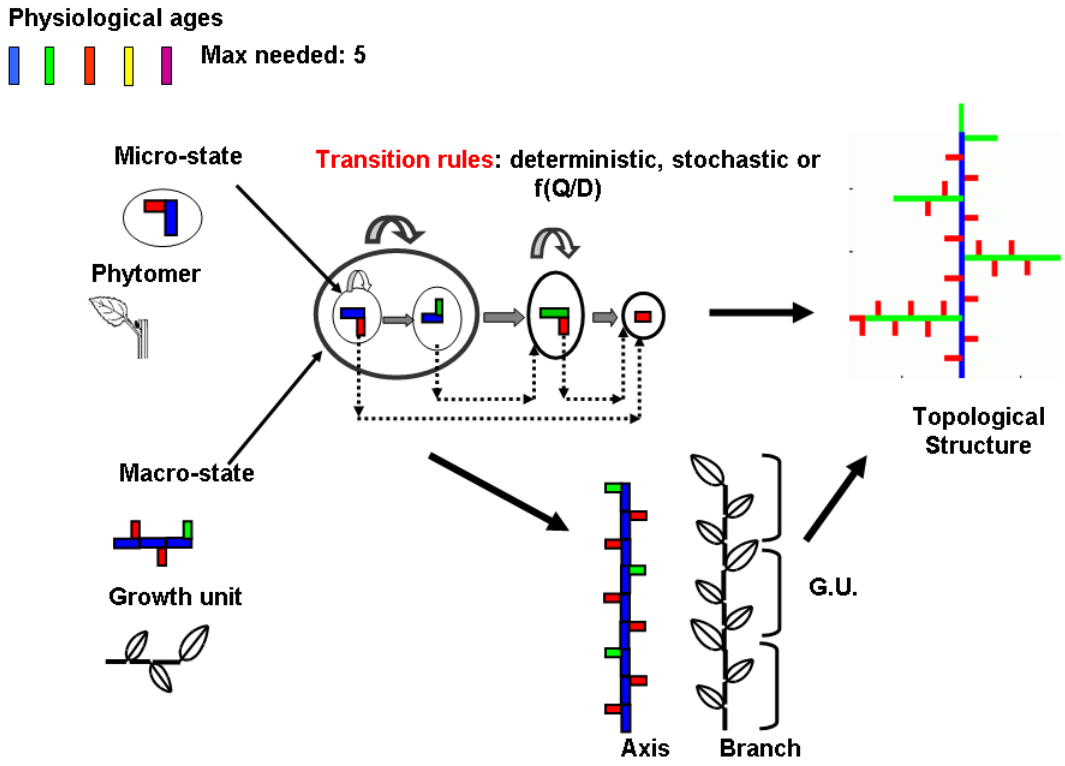


Figure 2.3: Dual-scale automaton for topological rules: deterministic (GL1), stochastic (GL2) or function of the source-sink ratio (GL3)

2.4.2 Growth grammar and structure factorization

Defining the rules of the automaton is analogous to writing a growth grammar ([Kang et al., 2006a]). Each symbol of the alphabet defines a constitutive element of the plant. Following Cournède et al. [2006] and Mathieu [2006], we denote:

- t the chronological age of the plant (the maximal growth cycle)
- $P = 1, 2, ..., P_m + 1$ the physiological ages (PA) where P_m is the maximal PA observed in the target plant
- $m_{pq}(n, t)$ the metamer of chronological age n and of PA p with axillary buds of PA q. By convention, metamers with axillary buds of PA equal to $P_m + 1$ bear no buds.
- $s_p(t)$ the bud of PA p at growth cycle t.

The set of symbols consists of metamers $m_{pq}(n,t)$ and of buds $s_p(t)$. The alphabet needed to represent a plant at age t is:

$$\widetilde{A}_t = \left\{m_{pq}(n,t), (p,q) \in P^2, q \geq p, 1 \leq n \leq t\right\} \bigcup \left\{s_p(t), 1 \leq p \leq P_m\right\} \tag{2.2}$$

These symbols are assembled to form structures using the concatenation operator, represented by the product symbol. A structure is defined as the set of organs generated from an initial bud: it includes the main axis but also all the lateral axes branched on this main axis. The use of structures as unit elements to describe the plant structure differs from the classical method where organs are gathered in axes. It allows a very compact writing of the whole-plant structure as the result of an inductive relationship. A structure at growth cycle t can be unequivocally associated to its basis characteristics, that is to say the physiological age p and the chronological age n of its basal growth unit (except in case of a plant exhibiting reiterated complex). Thus each structure can be described with a decomposition based on its basal growth unit (see figure 2.4). It consists of the metamers forming the basal growth unit, the structure generated by its apical bud and the set of structures generated by its lateral buds (these two structures being of chronological age $n-1$, i.e. they appeared one cycle after their mother growth unit). This decomposition can be transcripted as:

$$S_p(n,t) = \left[\prod_{p \leq q \leq P_m} \left(m_{pq}(n,t)\right)^{u_{pq}(t+1-n)} \left(S_q(n-1,t)\right)^{b_{pq}(t+1-n)}\right] S_p(n-1,t) \tag{2.3}$$

- where $S_p(n,t)$ represents the structure of chronological age n with a basis metamer of PA p, the current growth cycle being t.
- where $u_{pq}(n)$ is the number of repetitions of the microstate (p,q) initiated at cycle n, i.e. the number of phytomers with axillary buds of PA q in a growth unit of PA p.
- where $b_{pq}(n)$ is the number of lateral axes of PA q initiated at cycle n on a growth unit of PA p.

Hereafter we call Z_{pq} the set of metamers belonging to the same growth unit of PA p that have lateral buds of PA q (see Figure 2.1). So $u_{pq}(n)$ is the number of metamers appearing at growth cycle n in a zone Z_{pq}. This equation defines the structure $S_p(n,t)$ according to $S_q(n-1,t)$ and to the metamers of chronological age n. Using the same principle, the structures $S_q(n-1,t)$ can be decomposed and thus recursively until the structures $S_q(0,t)$ representing the buds of physiological age q at growth cycle t. The first growth unit at plant emergence is assumed being of PA 1 and is represented by the structure $S_1(1,1)$. The whole plant at cycle t is represented as the structure with a basis of PA 1 and chronological age t: $S_1(t,t)$. Those equations are valid while the apical bud

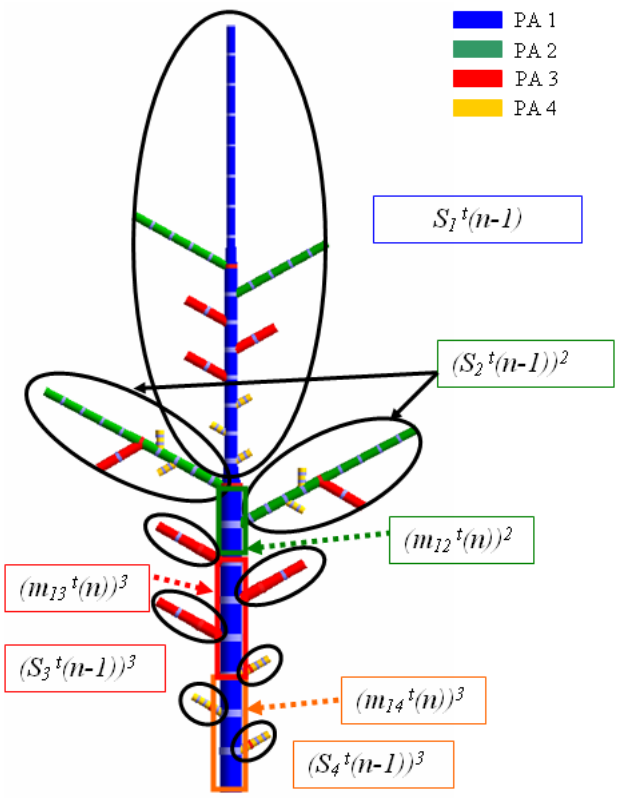


Figure 2.4: Inductive description of plant structure. Decomposition of the plant topology for $t = 3, P_m = 4, u_{14}(1), b_{14}(1) = 3, u_{13}(1) = b_{13}(1) = 3, u_{12}(1) = b_{12}(1) = 2$.

of each axis is of the same physiological age than its bearing phytomer. If a change of state occurs for the apical bud, the axis mutates or dies. The complete equation systems accounting for mutation phenomena can be found in [de Reffye et al., 2003]. In that case, the inductive relation is no longer only based on the basal growth unit but on the set of basal growth units of the main axis before its mutation. Another particular case concerns reiterations with a finite order. It imposes to introduce brackets or another index to distinguish two structures of the same chronological and physiological ages that have different reiteration orders.

2.4.3 Growth morphism

The growth morphism is analogous to defining the rewriting rules in L-Systems. Cournède et al. [2006] define it as:

$$\begin{cases} \Phi\left(m_{pq}(k,n)\right) = m_{pq}(k+1,n+1) \text{with} (p,q) \in P^2, 1 \leq k \leq n \\ \Phi\left(s_p(n)\right) = \left[\prod_{q \geq p} m_{pq}(1,n+1)^{u_{pq}(n+1)} s_q(n)^{b_{pq}(n+1)}\right] s_p(n+1) \end{cases} \tag{2.4}$$

where the notations have been defined in Equation 2.3. The initial condition is given by the seed $s_1(0)$.

The use of a small number of states (usually less than five physiological ages are sufficient to describe a plant) implies that at a given cycle, a lot of identical growth units appear at different locations in the plant, as can be observed in the example of Figure 2.4. More generally, similar parts of the plant are repeated at different positions in its architecture. From Equation 2.3 it can be seen that for example $S_p(n-1)$ is a structure that can be generated from lateral and apical bud of the basal growth unit of $S_p(n)$. It has led to the concept of substructure factorization, as presented in [Yan et al., 2004] and [Cournède et al., 2006]. Each structure type is computed only once and then the structures are stuck together, beginning from the simplest ones (the physiologically oldest), to form the whole-plant structure. A structure being defined by its chronological (n) and physiological (p) ages, only a set of at most $N \times P_m$ structures needs to be computed to simulate a plant at age N. In the stochastic case (GL2), Equation 2.3 is modified as:

$$[S_p(n,t)] = \prod_{p \leq q \leq P_m} (m_{pq}(n,t))^{u_{pq}(t+1-n)} \prod_{1 \leq i \leq b_{pq}(t+1-n)} [S_q(n-1,t)]_i \, [S_p(n-1,t)] \quad (2.5)$$

where $[S_p(n,t)]$ is a realization of the stochastic variable $S_p(n,t)$ [Cournède et al., 2006]. Theoretically, each realization of a structure can be different. It means that the factorization no longer holds in that case. However, it is still possible to keep its benefits by computing only a limited number of each type of structure to create a library of different potential realizations of each type of structure, as represented in Figure 2.5. With a small number of repetitions for each type (e.g. 5), the algorithm is still efficient and it is possible to build realistic realizations of plant structures with little bias from the theoretical stochastic plant [Kang et al., 2003].
Thus the substructure factorization algorithm provides a very efficient way of simulating plants, even with complex architectures. The simulation time is only dependent on the numbers of physiological states present in the plant and its chronological age, not on the total number of phytomers. At the same time, it is a limitation of the model as it restricts its flexibility to reproduce irregular architectures. But if this factorization may not be necessary for herbaceous, it is appropriate for trees. Indeed, when trees present complex branching structures, exact replication of their architecture is unfeasible: firstly because many stochastic events have influenced architecture set up and secondly because a complete description of whole-tree structures would require a huge and tedious experimental work. To simulate stand growth, and particularly radial increment of trees, a simplified description of tree architecture is sufficient. Moreover, the distance from simulated tree to real tree architecture can be quantified. Based on the notion of edit distance between two branched structures, Ferraro et al. [2005] define a way to characterize self-similarity in plants.

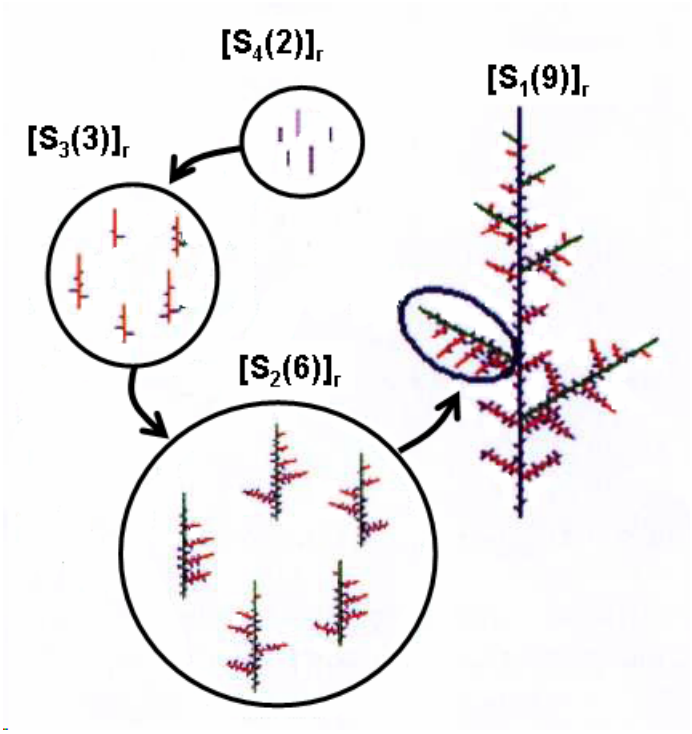


Figure 2.5: Strucutral decomposition of a plant with stochastic structure at growth cycle $t = 9$, from [Kang et al., 2003]. $[S_p(n)]_r$ represents a set of realizations of the stochastic structure $S_p(n)$.

2.5 Numbers of organs and versions of GreenLab

Numbers of organs of given physiological ages in each substructure are the main state variables for computing plant demand. Indeed, except for ring growth (see paragraph 3.2.3), the demand of the plant only depends on the number of expanding organs of each physiological age and their initiation dates. If plant development is known, its morphological growth can be deduced without any reference to its topology. The use of structure formalism provides an analytical method to compute at any time numbers of organs (or moments of numbers of organs for the stochastic version) without requiring to time-consuming methods of Monte-Carlo simulations. The morphism defined in [Cournède et al., 2006] provides the inductive formula to get the vector of size P_m, containing the number of organs of each physiological age in the structure, $\psi(S_p(n,t))$:

$$\psi(S_p(n,t)) = \sum_{p \leq q \leq P_m} u_{pq}(t+1-n) \cdot I_p + b_{pq}(t+1-n) \cdot \psi(S_q(n-1,t)) + \psi(S_p(n-1,t)) \tag{2.6}$$

where the i^{th} component of I_p is $I_p(i) = \delta_{pi}$ (Kronecker's delta) and:

$$\psi(m_{pq}(n,t)) = I_p \tag{2.7}$$

The different versions of GreenLab ($GL1$, $GL2$, $GL3$) are characterized by the functions defining the transition rules for the topological development.

2.5.1 Deterministic version: $GL1$

The transition rules of the automaton are constants. Numbers of metamers (u) and numbers of active buds (b) per growth unit do not depend on growth cycle t:

$$\begin{cases} \forall t, & u_{pq}(t) = u_{pq} \\ & b_{pq}(t) = b_{pq} \end{cases} \tag{2.8}$$

If the growth is undeterminate with no mutations of axes, the number of organs can be written under a simplified matrix form as in [Cournède et al., 2006].

2.5.2 Stochastic version: $GL2$

In the $GL2$ version [Kang et al., 2003], the topological rules to build plant structure are probabilistic: u_{pq} and b_{pq} are stochastic variables. A simulated plant is a particular realization of a plant distribution. In GL2, five processes of development are driven by a random variable for each physiological age p (not only u_{pq} and b_{pq}). The size of probability vectors is equal to the maximal physiological age (P_m):

- survival probability of bud, P_C: the PA p apical bud can die with probability $1 - P_C(p)$ and the growth of the bearing axis is then definitively stopped. The number of growth cycles from the axis initiation to its death follows a truncated geometrical law of parameters (M_A, P_A) where M_A components are the maximal number of growth units in axes of each physiological age.

- starting probability of axillary buds, P_B: the number of apical buds b_{pq} that can give birth to a new branch depends on probability $P_B(p)$. To simplify, it is assumed to be dependent only on the physiological age of the bearing axis and not on the physiological age of the potential lateral axis. $P_B(1)$ is the probability of seed germination (main axis appearance). The number of new branches on a metamer follows a binomial law $B(N_B, P_B)$ where N_B is the vector of bud number on metamers of each physiological age.

- growth probability of apical bud, P_A: at each cycle, the apical bud can produce a new growth unit or pause. If P_B and P_C are equal to 1, the number of growth units into an axis follows a binomial law $B(M_A, P_A)$ where M_A components are the maximal number of growth units in axes of each physiological age (as in first item).

- metamer appearance probability, P_I: the number of phytomers u_{pq} inside each growth unit zone is not a constant. For each phytomer, its appearance is tested against probability $P_i(p)$. Thus the number of metamers per growth unit follows a binomial law $B(\mu, P_I)$ where μ is the maximum total number of metamers per growth unit.

- probability of fruit appearance, P_f: there can be fruit abortion with the probability $1 - P_f(p)$.

Those probabilities can be fixed or variable in time to account for ontogenetic changes through plant life-time. It could also depend on environmental control.
Using these notations, the theoretical mean and variance of numbers of organs in each kind of structure can be computed. Here we only recall the final result (needed for fitting of stochastic development in Part II). We refer to [Kang et al., 2007a] for proofs and details or to [Kang et al., 2006a] for more synthetic formulations based on growth grammars and generating functions. The mean and variance of numbers of organs of physiological age 1 in the plant at growth cycle n are defined in equations 2.9 and 2.10. For the sake of clarity, p_C denotes the first component of the vector P_C (value corresponding to PA 1) and similarly for other probability values.

$$M_{S_n^{1,1}} = \frac{p_C\,(1 - p_C^n)}{1 - P_C} \cdot p_A \cdot \mu \cdot p_I \tag{2.9}$$

$$V_{S_n^{1,1}} = \frac{p_C\,(1 - p_C^n)}{1 - P_C}\left[p_A \mu p_I (1 - p_I) + p_A (1 - p_A)(\mu p_I)^2\right] + \frac{p_C\left[1 - (2n+1)(1 - p_C)(p_C)^n - p_C^{2n+1}\right]}{(1 - p_C)^2}(p_A \mu p_I)^2 \tag{2.10}$$

Similar equations can be written for physiological ages higher than one [Kang et al., 2007a].

2.5.3 Mechanistic version: $GL3$

In the $GL3$ version, the variables driving plant development (in particular u_{pq} and b_{pq}) are influenced by the level of trophic competition inside the plant, as introduced by Mathieu [2006]. This new modelling approach relies on a key state variable of the model: the ratio of biomass supply to plant demand, as detailed in Chapter 3. It requires a new decomposition of growth cycle course: first all active buds are taken into account to calculate the potential demand; it allows setting the number of active buds as a function of the ratio of biomass supply to potential demand and eventually, the number of new organs and the biomass they receive are determined.

Plant are seen as self-regulating systems that can react to exogenous (e.g. environmental stresses that reduce plant production; organ or branch pruning that reduce plant demand) or endogenous (e.g. apparition of new branches or fruits that increase demand) influences. Thus it allows simulating as well the ontogenetic changes in plant topology throughout its growth phases as architectural plasticity in response to environmental changes. In particular, it allows the progressive set up of architectural units in tree architecture, as illustrated in Figure 2.6.

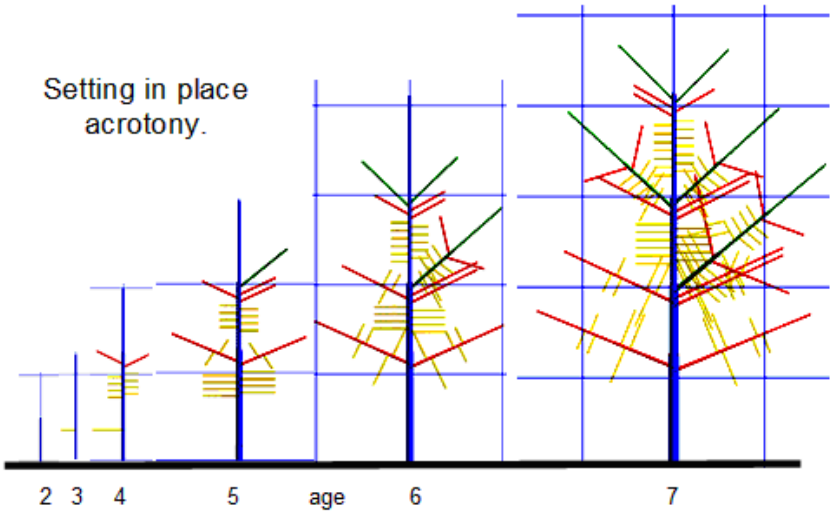


Figure 2.6: GL3: Structural plasticity of the growth units during plant development.

Several topological variables (number of metamers and of lateral axes per growth unit zone, duration of axes growth in G.C.) but also some functional variables such as expansion duration of organs or number of growth cycles of leaf activity. These variables are determined as affine functions of the ratio of biomass supply (Q) to demand (D). For instance, the number of lateral axes per growth unit zone is calculated as:

$$b_{pq}(t) = \left\lfloor B^1_{pq} + B^2_{pq} \cdot \frac{Q(t)}{D(t)} \right\rfloor \tag{2.11}$$

where $\lfloor x \rfloor$ denotes the integer part of x and B^1_{pq}, B^2_{pq} are parameters depending on the physiological ages of the bearing metamers and axillary buds. It implies that the number of active buds that give birth to a new branch is the same for every growth unit of given physiological and chronological ages. So the number of new growth units appearing at each cycle is a piecewise constant function of Q/D.
At a given cycle t, the total number of active buds in zone Z_{pq} for the whole plant is a multiple of the number of newly created growth units of PA p. If there are K new growth units of PA p, the total number of new branches appearing in zones Z_{pq} can take the values $n \cdot K$: (0, K, 2K, 3K,...). For trees, the value of K can be potentially

high (more than 100) so the number of new axes appearing in the tree can only take a limited set of discrete values.

Formula at whole-plant scale. For applications to real trees, it would be appreciable to have more flexibility for the number of branches that can appear at each cycle. This can be achieved with a minimal change of the modelling approach and with an identical number of parameters. Instead of calculating the number of active buds at growth unit scale, the calculation is done at whole-plant scale, with the same formula. The total number of new axes $a_{pq}(t)$ appearing on zones Z_{pq} in the tree depends on the number of existing positions potentially bearing that kind of axes ($N_{pq}(t)$) and on the value of Q/D. It is given by the following equation:

$$a_{pq}(t) = \left[N_{pq}(t) \cdot \left(A^1_{pq} + A^2_{pq} \cdot \frac{Q}{D}(t) \right) \right] \tag{2.12}$$

where $[x]$ represents the rounded value of x.

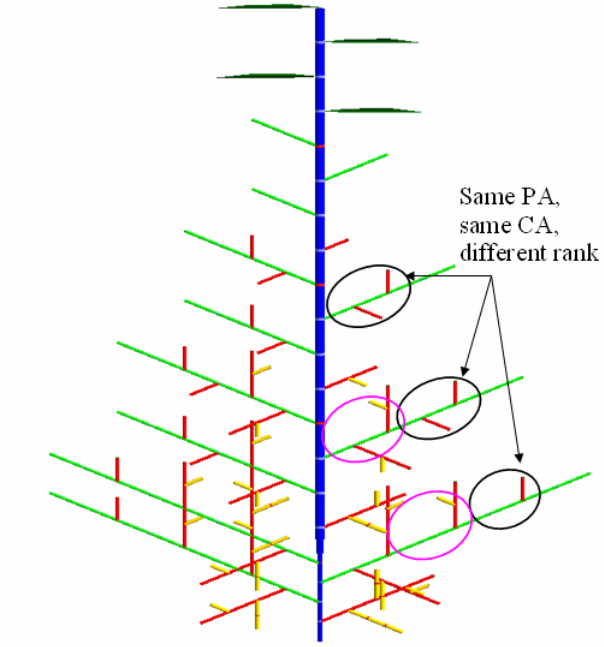


Figure 2.7: Global *GL3*: simulation of plant growth at G.C. t=5, the number of axes being calculated at whole-plant level. Metamers of same physiological age (PA) and chronological age (CA) but different ranks can bear different topological structures.

To compare, the former formulation gives:

$$a_{pq}(t) = N_{pq}(t) \cdot \lfloor B^1_{pq} + B^2_{pq} \cdot \frac{Q}{D}(t) \rfloor \tag{2.13}$$

In the new formulation, the total number of branches that appear can take values in a subdivision of the set of discrete values that are possible in the former formulation. However, even in this new formulation, not all integer values can be reached for the number of new axes. Indeed, these new axes have to be positioned on existing growth units. As done in the case of GL2, we keep the advantages of structure factorization by choosing a number of repetitions for every type of structure. We set that the topology of a structure depends on the rank (ontogenetic age) of its basal growth unit. It means that two metamers with the same physiological and chronological ages but a different rank can have different numbers of active axillary buds, as in Figure 2.7. In that case, we face the problem of choosing how to partition the $a_{pq}(t)$ among the potential positions $N_{pq}(t)$. A possible algorithm is described in section 8.4.1. This new method is illustrated through the example of beech tree analysis presented in that section.

Chapter 3

Physiology

In GreenLab, biomass production and allocation are based on a source-sink model, as presented in the first chapter. We detail here the equations driving morphogenesis, that is to say the organ growth and dimensions at each growth cycle.

3.1 Biomass production : sources

Through the plant life cycle, biomass inputs come from several kinds of sources. The initial input comes from the seed reserves that can be mobilized during several of the first cycles of growth. Then the plant becomes autonomous and produces its biomass by photosynthesis. Other potential sources of biomass, although not primary sources, consist of biomass remobilized from senescent organs or storage compartment into the rest of the plant. We present the GreenLab modelling choices for each of these processes.

3.1.1 Seed

During the first growth phase, the amount of biomass contained in seeds is transferred to seedlings. When fresh biomass is considered, the weight of the seed in GreenLab is equivalent to the total fresh weight of plant before the first leaves can perform significantly photosynthesis. This consists largely of water in the newly produced tissues during the seed germination and very early stage of seedling growth. Thus it surpass greatly the hydrated seed in weight.
Seed emptying can last one or several cycles, depending on species. During this phase, the growth is mainly heterotrophic and therefore the cumulated dry matter content of seed reserves and plant shoot part decreases due to respiration losses. The end of this heterotrophic growth phase can be determined as the end of this decrease [Asch et al., 1999].
In GreenLab, the proportion of seed biomass used for plant growth at each cycle is calculated according to a beta law density function whose parameters can be estimated by fitting. For most plants, it is generally assumed that seed emptying is fast and ends

after one growth cycle only. It means that the biomass included in the seed is equal to the total plant mass (shoot and root) at the first cycle. In fact, this is likely to be true at plant emergence but it may not be a good approximation if the cycle duration is long.

3.1.2 Assimilate production

Source organs

Source organs are mainly leaves but it can also include other kinds of organs potentially containing chlorophyll cells: stems, inflorescence parts, fruits (e.g. rape pods [Allirand et al., 2007]). In trees, stem and branches also have chlorophyll cells to capture CO_2 but their low permeability to gas (e.g. stark permeability is only 10% of needles one) prevent them from synthesizing high proportions of biomass [Deleuze, 1996]. In the model, all organs have the same status, all being potentially sinks and sources. However in current applications, only leaves have been considered as sources.

Hydraulic network of the plant

Dry matter production is related to plant water transpiration [Jones, 1992]. Leaf transpiration generates a pressure gradient that can induce water flows of several hundreds of liters per day in some trees, circulating until heights of more than 100m [Cruiziat, 1991]. That flow is limited by water uptake from roots and by water transport inside the plant architecture. We refer to [Mathieu, 2006] for detailed explanations of the formula presented below. The resistance to water flow circulation is defined for each organ:

- Leaf: $R_{leaf} = \frac{R_1}{S} + R_2$ where S is the blade one-side area and R_1, R_2 are resistance parameters.
- Internode: $R_{in} = \frac{\rho_k \cdot h}{s}$ where ρ_k denotes the layer resistivity of internode of physiological age k, h the internode length and s its transversal section area (Figure 3.1). For herbaceous, internodes are considered as perfect conductive cylinders. For trees, internodes are seen as pipes with concentric layers where only the recently created layers (sapwood) are conductive. The heartwood consists of dead cells (sometimes distinguished by a different colour due to chemical changes) that cannot conduct any flow and whose main role is to contribute to plant mechanical stability.

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,

plant biomass production can be calculated as:

$$Q = \frac{E}{R_{eq}} \tag{3.1}$$

where R_{eq} is the equivalent resistance of the whole network calculated by analogy with the classical laws of electric network (Kirschhoff's laws). We can benefit from the substructure decomposition of the plant presented in section 2.4.2 to get efficient ways of calculating R_{eq}. In the example provided in Figure 3.1, the equivalence resistance is given by the recurrence relation:

$$R_{eq}\left(S_1(n,3)\right) = \frac{\rho_1 \cdot h_1(n)}{s_1(n)} + \frac{1}{\frac{1}{R_{eq}\left(S_2(n-1,3)\right)} + \frac{1}{R_{eq}\left(S_1(n-1,3)\right)}} \tag{3.2}$$

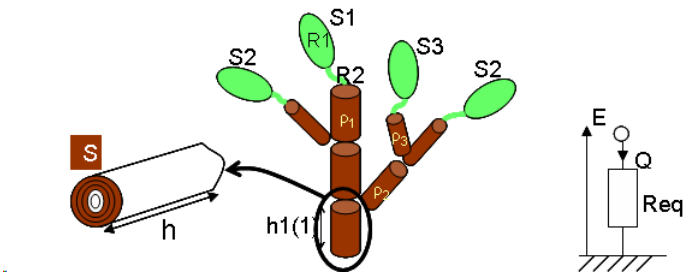


Figure 3.1: Hydraulic resistance of the plant architecture

The general formula can be found in [Mathieu, 2006]. This method could be easily extended to the case of a different potential for each leaf: an equivalent circuit can be computed by the method of successive "foldings" described in [Allen et al., 2005] and [Prusinkiewicz et al., 2007] (based on the Millman theorem, with the evapotranspiration potential at each location represented by a generator with a different electromotive force). However, this representation would not be fully compatible with the structure factorization of the plant. Indeed, structures of same physiological and chronological ages are branched at several locations in the plant and thus do not receive the same amount of radiations, although in our model they would have the same potential.
R_{eq} introduces a limitation of the biomass production to account for high resistivities of axes or for cases where the hydraulic network becomes so long and complex that it influences biomass production. It is negligible in most cases as Cruiziat and Tyree [1990] showed that hydraulic resistance of a young tree (6 m high) consists for 95% in evapotranspiration at blade surface, 3% in the root system and only 2% in the resistance of axes.

Light interception

In the previous versions of GreenLab (from [de Reffye et al., 1997] to [Mathieu, 2006]), only the simulation of individual plants is considered. It is based on the formalism of water use efficiency without self-shadowing of leaves. However leaf transpiration depends on the amount of intercepted radiations. Computing the leaf exposed area requires a detailed description of the plant architecture with the inclination angle and orientation of each leaf. Then the calculation should take into account the directions of incident radiations depending on the moment of the growing period and the phenomena of diffusion, transmission and reflection. An example of such detailed algorithm can be found for the LIGNUM model in [Perttunen et al., 1998].
In GreenLab, a simplified representation of plant architecture is used and there are no accurate measurements of leaf positions and orientations. Consequently, using a detailed algorithm for computing light interception is meaningless. As in many models [Marcelis et al., 1998], a simplified expression of light interception surface (LIS) can be derived from the Beer-Lambert law, assuming that light propagates into a turbid medium:

$$LIS = Sp \cdot \left(1 - e^{-k \cdot \frac{S}{Sp}}\right) \tag{3.3}$$

where k is an extinction coefficient related to leaf inclination and radiation direction. The surface Sp is the potential projection area of the plant silhouette onto a plane orthogonal to the main direction of the incident radiation. The ratio S/Sp represents the ratio of blade surface of the plant to a reference ground surface. It can be compared to the variable "Leaf Area Index" (LAI) that intervenes in the classical formulation of light intercepted by crop stands [Cournède et al., 2008]. However, in GreenLab, the parameters of this equation are considered as empirical, that is to say they are determined by fitting real data of plant production. Indeed, the conditions for Beer law application are not likely to be verified (leaf size infinitely small, leaves randomly located with uniform density) and moreover, the parameters Sp and k should take different values according to the direction of incident light. Thus they can be considered as global parameters that integrate the cumulated effects of light interception in different directions during the whole growth cycle.

Considering the parameters as empirical has also the advantage of allowing more flexibility. In the classical formulations of dry matter production per square meter for crop stands [Howell and Musick, 1984], different viewpoints lead to the same formalism. In the formulations based on water use efficiency (WUE), biomass acquisition is proportional to plant transpiration. In the formulations based on light use efficiency (LUE), it is proportional to the amount of absorbed radiations. In both cases, dry matter production is described as a negative exponential function of LAI. Howell and Musick [1984] compared the two approaches and concluded that computation of dry matter production based on transpiration is the most stable for their set of experiments. They

argue that in most regions of the world, water is usually the main limiting factor of crop production but admit that intercepted light may be easier to measure than crop transpiration, for example using remote sensing methods. In greenhouses, water is usually not a limiting factor. The effect of light interception (more precisely the PAR) is thus preponderant (e.g. TOMSIM [Heuvelink, 1999]).
In GreenLab, the same empirical equation can be used, its parameter meanings being interpreted differently according to the situation (water or light as preponderant factor). The environmental factor E_n can include different kinds of environmental influences (see paragraph 3.1.4).

General formula and parameter analysis

Eventually, we choose the following expression for biomass acquisition at growth cycle t:

$$Q(t) = \frac{E(t)}{R_{eq}} \cdot \frac{Sp}{S(t)} \cdot \left(1 - e^{-k \cdot \frac{S(t)}{Sp}} \right) \tag{3.4}$$

The term $\frac{Sp}{S(t)} \cdot \left(1 - e^{-k \cdot \frac{S(t)}{Sp}}\right)$ is a dimensionless factor representing the limitation of biomass production due to self-shading of leaves (see figure 3.2).

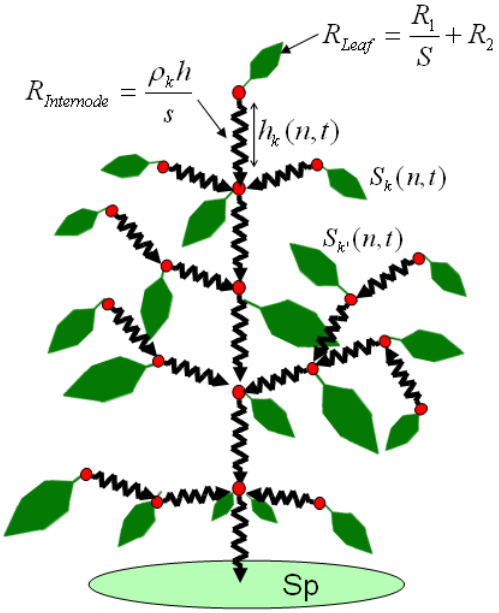


Figure 3.2: Representation of the plant for the general expression of biomass production.

When the hydraulic resistances of petioles and internodes are negligible, the equivalent

resistance writes: $R_{eq} = \dfrac{R_1}{S(t)}$ with S being the sum of all active blade areas. Thus the equation simplifies and can be written as follows:

$$Q(t) = E(t) \cdot \mu \cdot Sp \cdot \left(1 - e^{-k \cdot \frac{S(t)}{Sp}} \right) \tag{3.5}$$

where $\mu = 1/R_1$ is a parameter of biomass conversion efficiency.
Regarding trees, Sp is a parameter representing a reference surface used to account for the effects of self-shading and modulated by the competition of neighbouring plants. The long-term development of tree crown and the step by step metamorphosis of tree architecture make Sp varies with $S(t)$. In GreenLab, as no precise geometrical data are included for crown description, a simple allometric relationship is proposed to compute the variations of the reference surface according to the blade surface:

$$Sp(t) = Sp_0 \cdot \left(\frac{S(t)}{Sp_0} \right)^{\alpha} \tag{3.6}$$

Practically, parameters Sp_0 and α must be identified by inverse method so that the progression of Sp induces the measured value of biomass production.

Effect of parameter Sp

For crops, parameter Sp is considered as a constant through plant life time. It is estimated by fitting as the virtual reference surface that gives the observed biomass production under the measured environment values. It drives the saturation of biomass production: if Sp is small, the maximal potential biomass production is small.
In the simplified case of negligible hydraulic resistances, the biomass production of the plant at young stages does not depend on Sp but it is approximately proportional to the plant leaf area: If $S(t) << Sp$, $Q(t) \approx \mu \cdot k \cdot S(t)$.
Sp influences the growth of the plant at older stages: it controls the limit value of the biomass production (see figure 3.3). It also controls the value of $S(t)$ for which there is "saturation": 99% of the limit value for biomass production is reached when $kS(t)/Sp > 4.6$. It is important to note that above that limit, increasing the plant blade surface does not increase its production.
Lastly, an advantage of this formalism is its easy extension to modelling of stand growth, including effects of competition for light on biomass production of each plant. The parameter Sp can be used to account for density effects on the biomass production (Ma et al. [2007]). Inter-plant competition is quantified by the intersection areas between influence zones of each plant (characterized by the value of Sp for every neighbour plants [Cournède and de Reffye, 2007]).

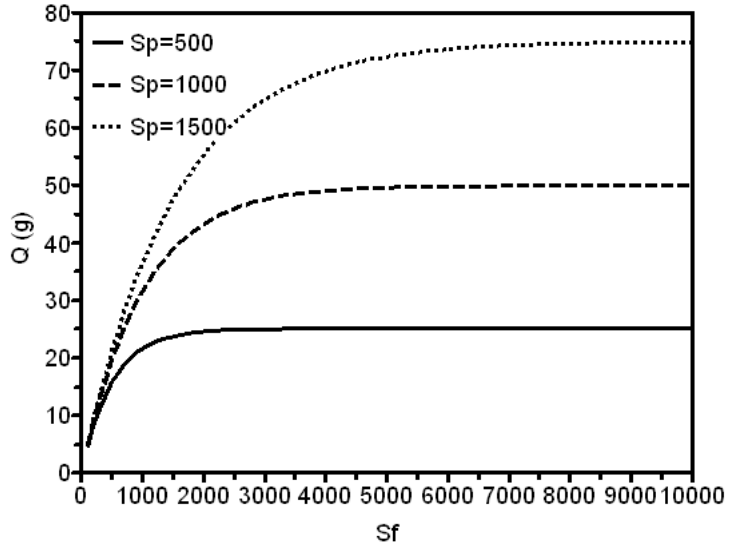


Figure 3.3: Effect of Sp variations (in cm^2) on biomass production (Q) as a function of total leaf area (Sf) with $E = 1$, $k = 1$, $\mu = 0.05$. The limit value is $\mu \cdot k \cdot Sp$.

Effect of parameter α (Variation of Sp)

Parameter α defines the allometry between plant total leaf surface and its reference surface Sp that influences the saturation of biomass production, as presented in the previous paragraph (3.1.2). It accounts for the crown shape. If $\alpha = 0$ it means that the reference surface can be considered as a constant: this can be an effect of high density (tree production is calculated as if its development was constrained inside a cylinder of basal area Sp). If $\alpha = 1$, the reference surface grows proportionally to the total leaf area. It can correspond to grass or to tabular-shaped trees such as *Albizia*. We show in Figure 3.4 the influence of α on the simulated biomass production and plant morphology.
Figure 3.5 shows the influence of α on the architectural development of a tree simulated with the $GL3$ version (see section 2.5.3).

Parameter identification

Introducing a new kind of equation (or *model structure* [Walter and Pronzato, 1994]) raises the problem of parameter identification. Indeed, the comparison between two alternative modelling options with regards to real data provides arguments to privilege one or the other one. The parametric identification of models relies on the comparison between their outputs and observations on real phenomena. The criterion of this comparison is an objective function that can be for example the sum of squared differences

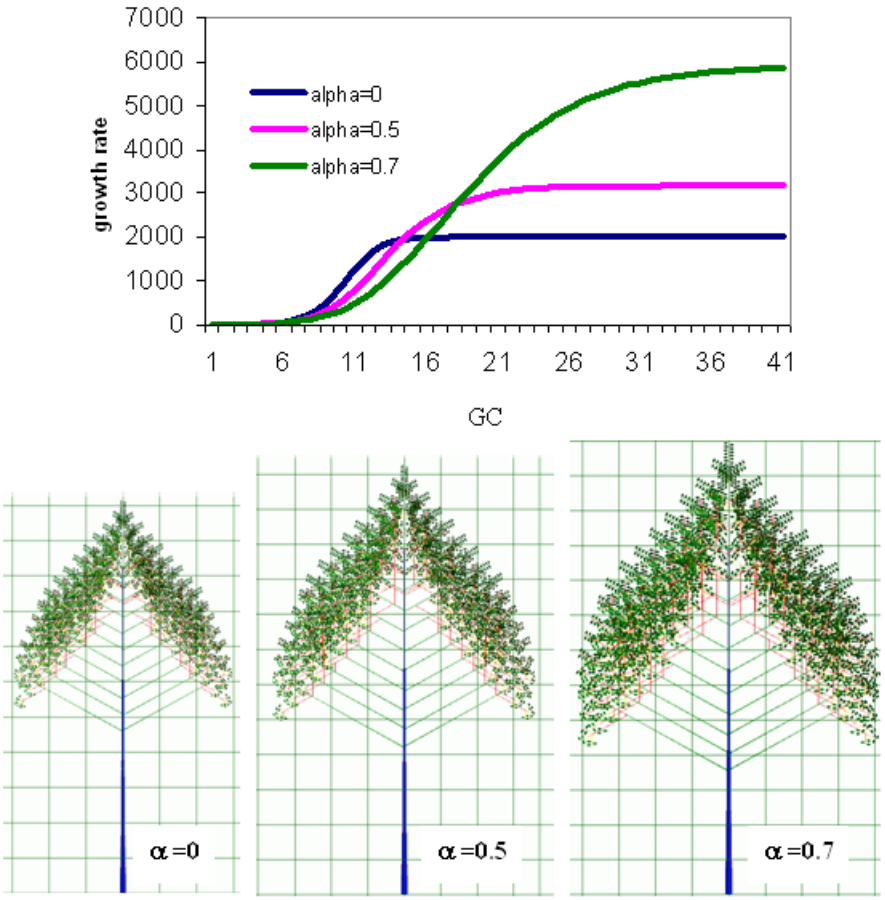


Figure 3.4: Effect of parameter α on biomass production and plant morphology, case $GL1$.

between observed and simulated data (see section 7.2). For a given model structure assumed to be representative of the true process, the procedure consists in finding the most probable parameter values that generate the observed outputs. Different criteria have been developed to discriminate the best model structure. Walter and Pronzato [1994] suggest taking into consideration the accuracy of the estimator and the complexity of the model. The complexity of the model is simply characterized by the dimension of the parameter vector ($dimP$) in the widely used AIC criterion (An Information Criterion). A high dimensional parameter vector penalizes the objective function. Other criteria make different penalization expressions intervene. For instance, in the Final Prediction Error (FPE) criterion, it is:

$$ln\left(\frac{1+dim(P)/n_t}{1-dim(P)/n_t}\right)$$

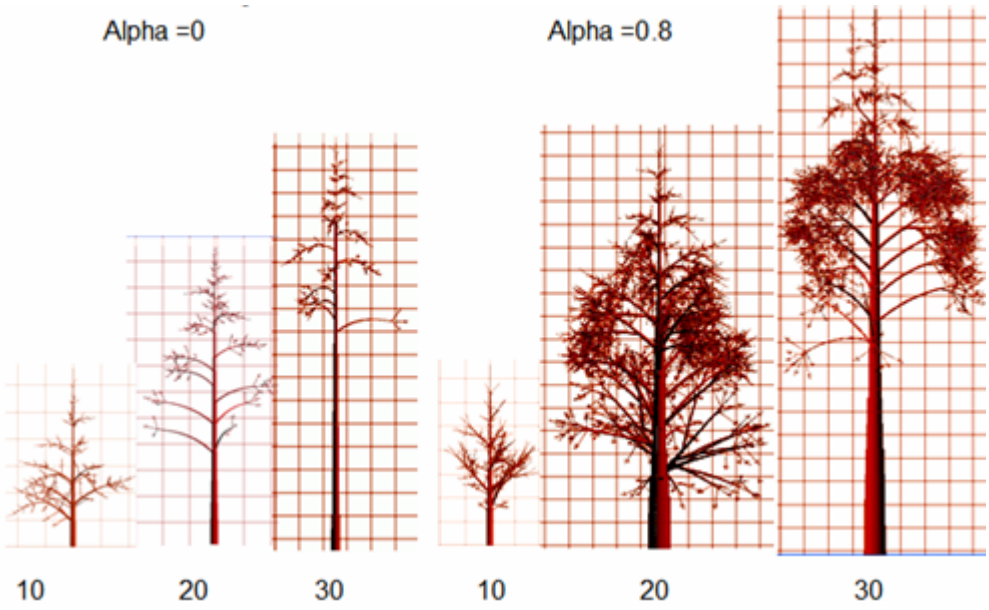


Figure 3.5: Effect of parameter α on plant architectural development in the $GL3$ version.

and in the Bayesian Information Criterion (BIC), it is:

$$dim(P) \cdot \frac{ln\ n_t}{n_t}$$

n_t represents the number of measurement dates: the highest it is, the less important is the parameter vector dimension.
In our case, both equations 3.1 and 3.5 for biomass production have the same number of parameters: R_1 and R_2 for the hydraulic equation (3.1); μSp and k/Sp for the Beer law equation (3.5). Therefore the only criterion to consider is the fitting accuracy. Several tests were performed on isolated plants [de Reffye et al., 2006]. Table 3.1 presents the results for the values of final error (based on the weighted least square criterion as defined in 7.2).

Plant	Error Eq.3.1	Error Eq.3.5
Cotton	6.26	6.06
Wheat	3.27	3.89
Sunflower	46.1	45.9
Maize	47.7	47.8
Tomato	1.76	1.64

Table 3.1: Comparison of fitting accuracy for two model structures for biomass production [de Reffye et al., 2006].

The final error is of the same range for both model structures. Fitting results are of the same quality. Therefore both functions are adequate. The choice of the Beer-Lambert law is reasonable as it is classically used and as it can be interpreted on physical bases. Moreover, introducing the reference surface Sp allows defining the rules of competition for light in a stand, as done in [Cournède et al., 2008].

Variations of RUE

For some plants (e.g. pea [Lecoeur and Ney, 2003]), it has been reported that the radiation use efficiency (RUE) can vary according to the growth stage. By fitting the model on real data, GreenLab can be used as a tool to explore such variations. For example, it was investigated in the case of *Arabidopsis thaliana* (joint work with Angélique Christophe, LEPSE, Montpellier) [Christophe et al., 2008]. A different equation of biomass production was chosen, although the same study could be realized with the equations defined above. It was calculated from Monteith's equation (1977) adapted to the case of an isolated plant and considering only vertical direction of incident radiations. The vertical light interception efficiency was assumed to be equal to the ratio of the plant vertical projected area ($S'(t)$) to its total blade area. It gives:

$$Q(t) = RUE \cdot S'(t) \cdot PAR \tag{3.7}$$

where PAR is the incident photosynthetically active radiation. Dry weights of organs, blade areas and projected area of the basal rosette were measured at five growth stages. Those data were used to build a model of light interception by the rosette leaves and to get the variations of the RUE parameter (g of dry biomass MJ^{-1}) as represented in Figure 3.6.
However, in the results of *Arabidopsis* fitting, it was difficult to disentangle the effects of complex interactions between several phenomena: self-shading (especially for leaves of the basal rosette), leaf senescence, respiration losses (see section 3.1.2), light absorption, reflection and effective variations of radiation use efficiency.

Respiration

Modelling of respiration classically consists of two parts: growth respiration, proportional to biomass production, and maintenance respiration, proportional to biomass of living organs [Le Roux et al., 2001]. In GreenLab, respiration losses are implicitly taken into account at the whole plant scale. The calculation of biomass production (Q) represents the plant *neat production*, that is to say only the biomass amount effectively available for structural growth. It is analogous to considering respiration as a priority sink to which a fixed proportion of the production is systematically allocated [Lacointe, 2000]. However, a more mechanistic approach would consist in considering the individual respiration process of each organ. It would allow a more accurate analysis of

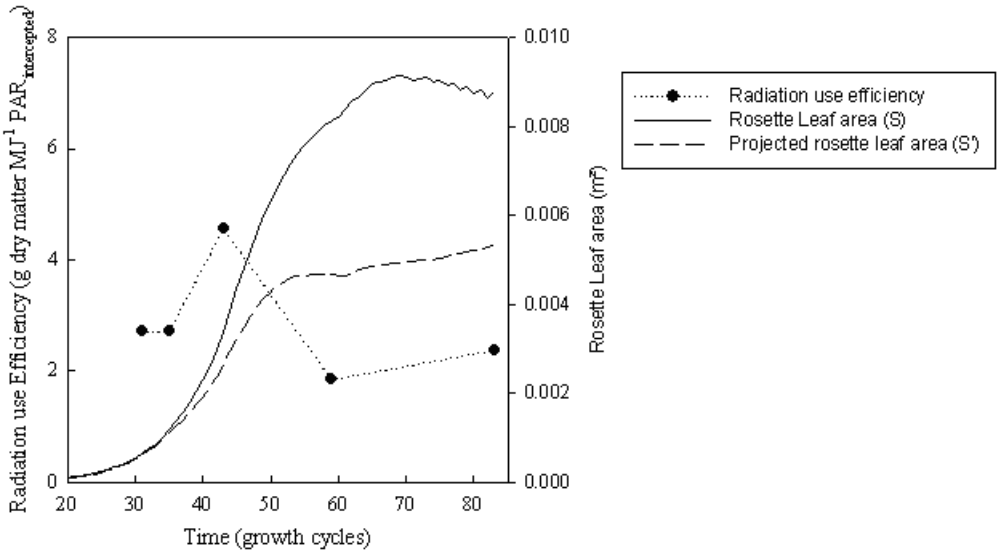


Figure 3.6: Variations of total leaf area, projected leaf area and RUE from fitting of *Arabidopsis thaliana* [Christophe et al., 2008].

biomass production and allocation as not only structural biomass but also construction and maintenance costs could be attributed to every organ.

3.1.3 Remobilization

Biomass can also come from secondary sources, that is to say from re-allocation of biomass already present in the plant organs. In particular, when organs become senescent, they can give back part of their biomass to other growing organs of the plant. That kind of process is modelled for example in VICA by Wernecke et al. [2007] where reversible fluxes can be exchanged between different types of organs.

In GreenLab, it is assumed that after an organ stops growing, it can release part of its biomass into the common pool. It is modelled using three parameters: the time duration between the end of expansion of the organ and the beginning of its emptying (Te), the maximal proportion of biomass that can be given to the common pool (F) and the speed of emptying (k). It is assumed that those parameters only depend on the kind of organ considered and not on its physiological age. Thus the mass of an organ of chronological age n that has finished its expansion ($n \geq T_{exp}$) and that begins emptying, is calculated as follows:

$$\forall n \geq T_{exp} + Te,\ q(n) = q_{max}\left(1 - F \cdot \left(1 - (1-k)^{n-Te-T_{exp}}\right)\right) \tag{3.8}$$

Figure 3.7 presents the shape of the remobilization curve for different values of k.

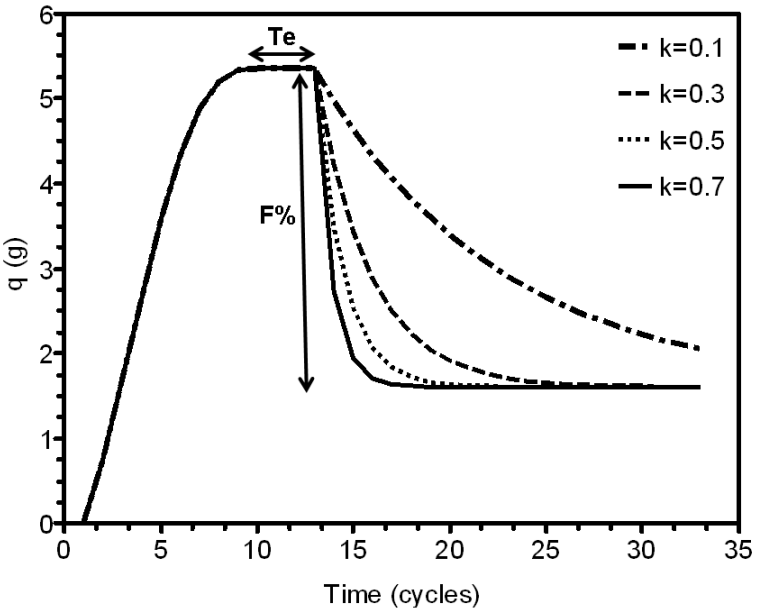


Figure 3.7: Variations of organ biomass with remobilization process, $T_{exp} = 10$, $Te = 3$, $F = 0.7$.

The time duration before the beginning of emptying (Te) must be assessed by the modeller, whereas the proportion (F) and speed (k) of emptying can be fitted from data of organ biomass. An initial value can be input for F from the ratio between the maximal and final weights of an organ. This model for remobilization process was tested for the first time on rice data from the China Agricultural University (Zheng BangYou).

Figure 3.8 presents the fitting results for the compartment of biomass.

All compartments undergo a decrease at the last growth stages, except the fruit compartment (panicles). In the model, the biomass lost by the internode, blade and sheath (denominated as petiole) compartments is given back to the common pool and then reallocated (mainly to the panicle which is the only organ still growing). Table 3.2 presents the parameter values for the remobilization process.

Parameter	Te	k	F
Internodes	4	0.51	0.23
Blades	4	0.046	0.42
Petioles	4	0.037	0.54

Table 3.2: Table of fitted parameters for remobilization in rice (joint work with Zheng BangYou, CAU).

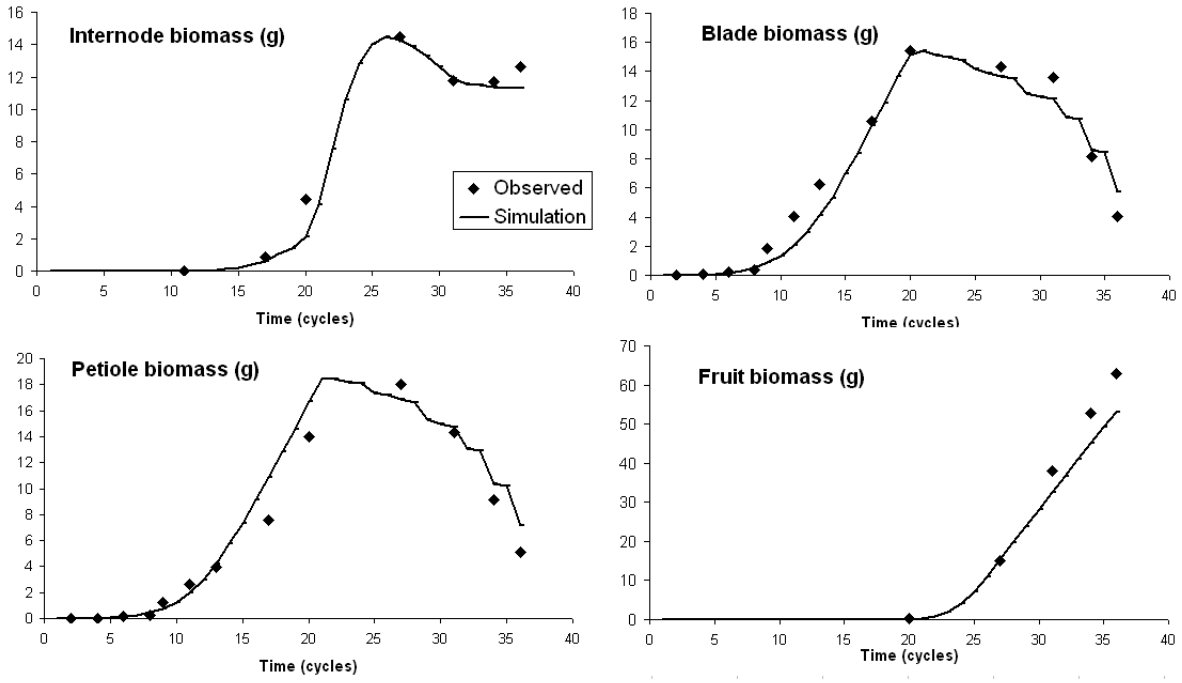


Figure 3.8: Fitting results for compartment biomass of rice cultivar Peiai64S/E32 (joint work with Zheng BangYou, CAU).

It can be observed that the dynamics of the emptying process is different for internodes than for blades and petioles. The speed of emptying is the highest for internodes ($k = 0.51$ compared to less than 0.05) but the final proportion of biomass that they can lose is twice less than for blades and petioles.

3.1.4 Environmental variable

In the current version of GreenLab, the environmental conditions have only effect on biomass production through the global vector variable E. That global variable can incorporate the cumulated effects of temperature, light and water conditions. A proposition for the quantification of these effects is given in [Wu, 2005]: a beta density law is used for temperature, a negative exponential function for light, a rational function for water. It is possible to apply optimal control algorithms to get the best environmental conditions to maximize the growth. An example of optimization of the water supply strategy can be found in [Wu et al., 2005].
The effect of environmental variables is visible in the organ weights at each cycle. An interesting consequence is that it is possible to trace back the environmental variations by fitting the model on measurement data. For example, several plants of *Cecropia sciadophylla* were measured in 2007 in French Guyana (joint work with Patrick Heuret

and Camilo Zalamea (AMAP, CIRAD)). In the case of this tropical tree, the main environmental variable represents rainfalls. The Guyanese climate can be represented by the succession of dry seasons and wet seasons. It can be modulated by a short period around March where the rainfalls rate is reduced ("short summer of March")
For this first fitting, it was chosen to represent the variations of pluviometry using a sinusoid function:

$$E(t) = max(0, E + A \cdot cos\left(\frac{2 \cdot t \cdot \Pi \cdot \omega}{T} + \phi\right) \tag{3.9}$$

where E is a default environmental value characterizing the more or less favourable conditions of growth, T is the period of environmental fluctuations (one year), A is their amplitude and ϕ the phase. The phase can be assessed directly from an estimation of the tree age (the sampling date being known) or fitted. For *Cecropia*, these sinusoidal variations induce a similar periodic variation of internode length, mass (Figure 3.9) and pith diameter.

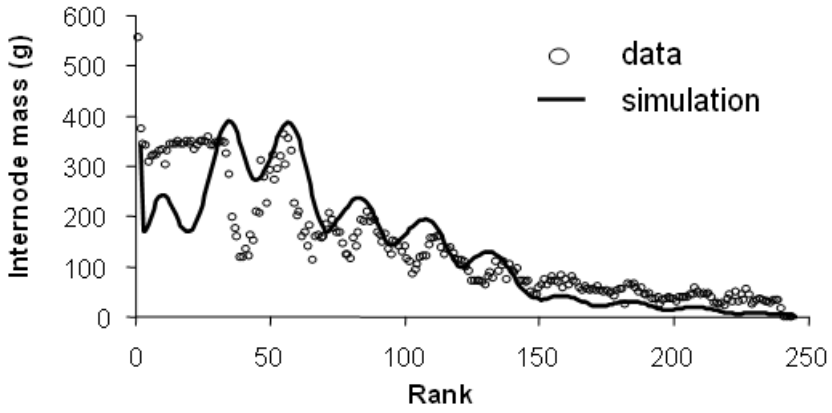


Figure 3.9: *Cecropia*. Internode mass on the main stem: data vs simulation with a sinusoidal environment (joint work with Patrick Heuret and Camilo Zalamea, FTH 2007).

3.2 Biomass allocation : sinks

The pool of biomass production is partitioned between four major compartments: roots, buds, expansion of existing organs, rings. Biomass is first allocated globally to these four compartments then distributed to each organ individually in a second step (Figure 3.10).
As no reserve compartment is considered, the whole biomass pool produced at growth cycle t is distributed at the end of that growth cycle:

$$\forall t \geq 0,\ Q(t) = Q_{bud}(t) + Q_{exp}(t) + Q_{ring}(t) + Q_{root}(t) \tag{3.10}$$

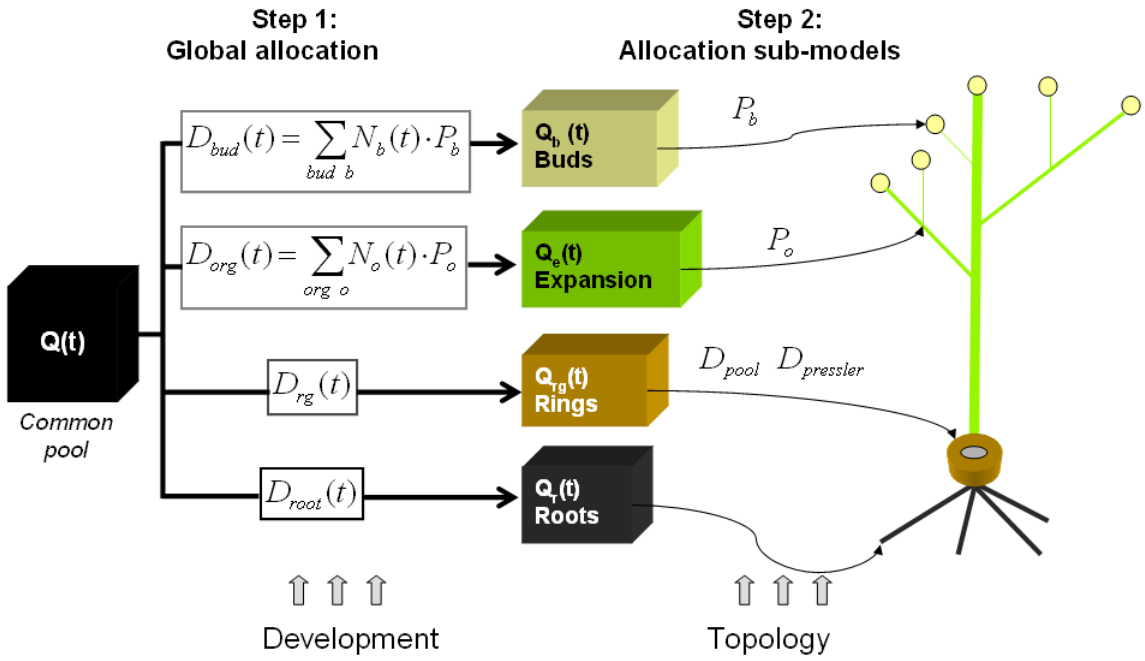


Figure 3.10: Two steps for biomass allocation

where $Q_{bud}(t)$ represents the biomass allocated to new organs (abusively called "buds") at cycle t, $Q_{exp}(t)$ the biomass allocated to the growth of already existing organs (expansion), $Q_{ring}(t)$ the biomass allocated to cambial growth of internodes and $Q_{root}(t)$ the biomass allocated to the growth of the root compartment (Figure 3.10). The biomass attributed to a compartment C is proportional to its demand:

$$\forall C \in \{Bud, Exp, Ring, Root\},\ Q_C(t) = D_C(t) \cdot \frac{Q(t)}{D(t)} \tag{3.11}$$

$D(t)$ is the plant total demand, calculated as the sum of the four compartment demands:

$$D(t) = D_{bud}(t) + D_{exp}(t) + D_{ring}(t) + D_{root}(t) \tag{3.12}$$

Two different modelling choices can be distinguished for the calculation of $D_C(t)$. For buds and expansion, it depends on the underlying number of organs whereas for roots and rings, demand is calculated at compartment level, regardless of the plant topology and development.

In the next paragraphs, we define the demand for each of those four compartments and we describe the intra-compartment partitioning procedure.

3.2.1 Bud and expansion compartments: organ-based demand

There is no formal distinction between the calculation of bud and expansion compartment demands. Both are calculated as the sum of the individual organ demands in the compartment. Thus those compartment demands only depend on the number of organs in each class of physiological and chronological ages. The demand of an organ is expressed as the factor of two components: its sink strength and a function defining its variation throughout its life. As sink strength is defined as a relative value, a reference must be chosen. It is generally assumed that the sink strength of leaves of physiological age 1 is equal to 1. Let $D_o(k,n)$ denote the demand of an organ of physiological age k and chronological age n (with $0 \leq n \leq t$, including $n = 0$ for buds). It is assumed that the variables n and k can be separated: the sink strength of an organ depends only on k and the sink variation function depends only on n:

$$D_k^o(n) = P_k^o \phi^o(n) \text{ with } k \in \{1 \dots P_m\},\ n \in \{0 \dots t\} \tag{3.13}$$

Let $q_k^o(n,t)$ denote the biomass allocated to organ o of physiological age k and chronological age n at growth cycle t:

$$\begin{aligned} q_k^o(n,t) &= D_k^o(n)\frac{Q_{bud}}{D_{bud}}(t) \\ &= P_k^o \phi^o(n)\frac{Q}{D}(t) \end{aligned} \tag{3.14}$$

The sink variation function is defined with a generic shape of beta law density function. It depends on three parameters, a, b and T_{exp}:

$$\phi^o(n) = \frac{1}{N}\left(\frac{n+0.5}{T_{exp}}\right)^{a-1}\left(1-\frac{n+0.5}{T_{exp}}\right)^{b-1} \tag{3.15}$$

T_{exp} represents the duration of organ expansion, expressed in growth cycles. It is estimated from measurements of organ sizes. Parameters a and b drive the curve shape. These two parameters are empirically determined from fitting the model on real data. N is a multiplicative factor introduced to normalize the sink variation function. This factor can be defined following different methods, as detailed in Appendix B. Figures 3.11 represent different shapes of beta law density function, for several values of parameters a and b, and the resulting organ weight in the virtual case of $Q/D = cte = 1$. Note that asymmetrical shapes can be obtained when $a \neq b$.
The main advantages of the beta law were described in Yin et al. [2003] who compared it to six families of functions classically used to represent growth phenomena:

- at initial time t=0, its value is zero (as for the Weibull function)
- the values of the final growth rate is null, which is consistent with the definition of T_{exp} as the expansion duration.

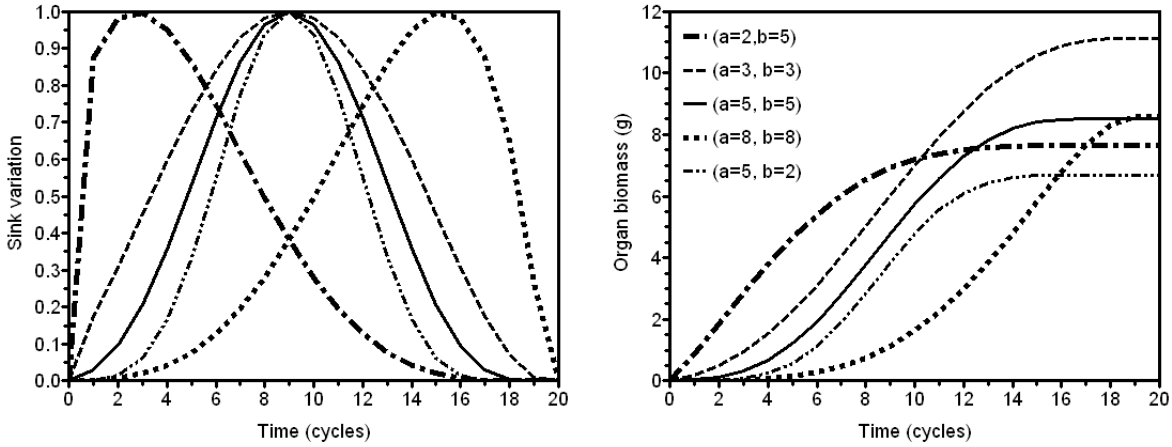


Figure 3.11: Sink variation function and evolution of organ biomass.

- it has a high flexibility and can describe asymmetric growth trajectories (as Richards law)
- it has stable parameters for statistical estimation (similar to Gompertz or logistic functions)

The normalization factor N

The choice of the normalization method is particularly important when the expansion duration of organ can vary for organs of the same kind. We expose four possible ways to choose of N in Appendix B. The choice of the method can influence the interpretation of the notion of "sink strength" (parameter P_k^o). In the three first methods, the sink strength is defined as the cumulated demand coefficients through the organ growth duration. By contrast, in the last method described, the sink strength corresponds to the maximal value taken by organ demand.

Saturation of organ weight

In the current version of GreenLab, no saturation of organ sink or size is set *a priori*. However, Heuvelink [1996] reports that the growth rates of tomato fruits remain unchanged when one tomato or two tomatoes per fruit were left. If the production is stable (case of saturation of the ground surface by blade surface), then biomass allocation to the fruit should be reduced when a second fruit is present as a mechanistic consequence of sink competition. Since it is not the case, a reason can be that there should be a physical or physiological limitation of organ growth, at least for some species. That limitation can be included in the model. For example, let q_{max} denote the maximal weight of an organ. The biomass increment of that organ at chronological age n, the

plant being at cycle t, can be expressed as follows:

$$q_k^o(n,t) = P_k^o \phi^o(n) \frac{Q(t)}{D(t)} \left(1 - \sum_{j=t-n+1}^{t-1} \frac{q(n-(t-j),j)}{q_{max}}\right)^+ \text{ with } 2 \leq n \leq t \qquad (3.16)$$

If the maximal biomass q_{max} is very high, the introduction of the limiting factor has few influence.

Sink variations according to growth conditions

In the model presented above, sinks of organs are independent of the biomass supply and of the changes in environmental conditions. However, it has been often observed that allocation coefficients depend on the level of carbohydrate supply. For instance, if only one leaf is left on a seedling, there is a change in allocation proportion in favour to young leaves [Taiz and Zeiger, 2006]. This effect is the basis of the functional balance principle (presented for instance in [Canell and Dewar, 1994]): the plant invests more in the compartment (shoot or root) that can provide the limiting factor needed for an optimal growth (carbon or nitrogen). Changes in allocation are an emergent property in transport-resistance models as a result of the resistances of the pathways between the different sources and sinks ([Minchin et al., 1993], [Minchin, 2004]).
Ma et al. [2007] analyzed the stability of GreenLab parameters for maize in response to three kinds of phenotypic variability: within populations, among seasons and among growth stages. As the model is deterministic with species-specific parameter values, it does not explain the inter-plant variability. Parameters were found stable for different growth stages and more stable than the phenotypic data (organ weights and dimensions) when the environmental conditions change (the average variation was 18% for phenotypic data whereas it was 10% for parameter values). The authors concluded that the two environmental factors considered (ETP for biomass production and temperture for development are major contributors to phenotypic changes). However, Ma et al. [2008] for maize and Louarn et al. [2007] for tomato found that several parameters varied with plant density: blade resistance, internode sink and fruit sinks. The shape of sink variation functions remained the same but expansion durations also varied.
Therefore, some parameters should be made dependent on the growth conditions. Another possibility is to introduce feedback controls of the state of the plant on sink values. If the sink of an organ depends on the ratio of the biomass supply to the demand (Q/D), it allows simulating processes where the plant adjusts its allocation to different types of organs according to its growing conditions. For instance, the weight of tomato fruits depends on its number of cells, which is determined during a period of a few cycles at its initiation [Bertin et al., 2001]. To model that effect, sinks of internodes of physiological age k and chronological age n can be expressed as the sum of two components: a constant sink Pi_0, representing the demand associated to cell expansion and a variable sink depending on the ratio Q/D at its initiation cycle, representing the

sink strength associated to the number of cells [Mathieu, 2006]. Thus its demand at cycle t is calculated as:

$$D_k^i(n,t) = \left(P_{k,0}^i + P_{k,1}^i \cdot \frac{Q(t-n)}{D(t-n)} \right) \cdot \phi^i(n) \tag{3.17}$$

where $P_{k,0}^i$ and $P_{k,1}^i$ represent constant and mass sinks of internodes of PA k. By fitting on real data, GreenLab can be used as a tool to evaluate whether sinks are constant during the plant growth or if they vary according to Q/D.

Expansion delays

In some herbaceous plants, the final topology becomes visible only at the end of the growth after elongation of lateral axes. In GreenLab, we only consider delays to expansion but not delays to initiation. When there are expansion delays, the whole structure can be already initiated long before its appearance. In that case, phyllochron can be very different from plastochron. For some herbaceous such as *Arabidopsis* or Chrysanthemum, the growth is determinate and it is the mutation of apical bud that triggers the extension of organs and axillaries. This phenomenon has been defined as *apical dominance* in [Cline, 1997]. Development and expansion are completely independent and can even occur in opposite temporal sequences: in *Arabidopsis thaliana*, organ initiation is bottom-up whereas extension of lateral axes is top-down. The acrotonic aspect of its branching structure is a consequence of this top-down extension: the lower lateral axes have a shorter growth duration and they are subject to higher trophic competition due to the simultaneous growth of upper axes. In GreenLab, this is modelled using delays. Organs can be initiated in the meristem but remain undeveloped, in a stage of primordium, for several cycles. They exist inside the meristem but their demand for biomass is zero or negligible. Thus, the complete expression of the demand of an organ of physiological age k ($1 \leq k \leq P_m$), of chronological age n ($1 \leq n \leq t$) and belonging to axis j ($1 \leq j \leq t$) is defined as:

$$D_k^o(n,j) = P_k^o \cdot f^o(n) \cdot d_k^o(n,j) \text{ where } d_k^o(n,j) = \begin{cases} \epsilon \text{ if } n \leq d_k'^o(j) \\ 1 \text{ else} \end{cases} \tag{3.18}$$

These delays can be fixed or dependent on the Q/D ratio. For instance, in growth simulations of *Arabidopsis thaliana*, delays of the lateral axes were set according to the Q/D ratio after the appearance of the main flower, from the top axis to the lower ones:

$$d_k'^o(j) = A^o(k) + B^o(k) \cdot \frac{Q}{D} \left(2(T_F + 4 - j) + j - 1\right) \tag{3.19}$$

where T_F is the cycle of the apical flower appearance. The 3D output is shown in Figure 3.12.

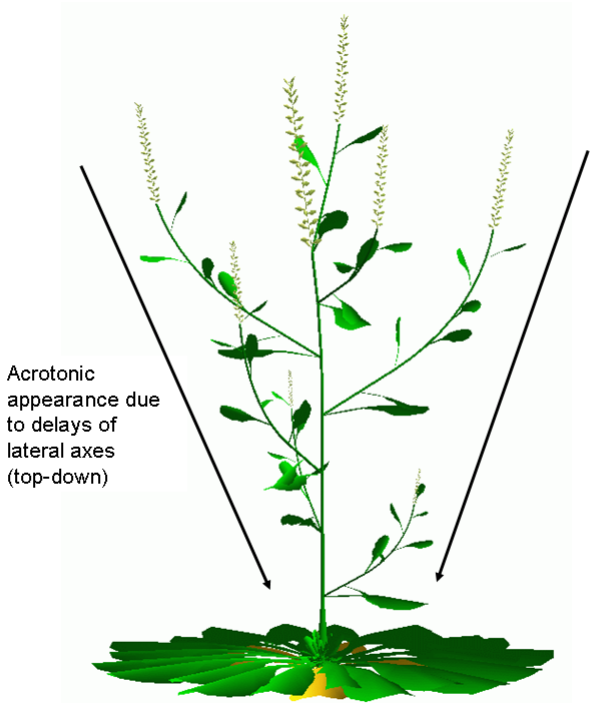


Figure 3.12: 3D visualization of *Arabidopsis thaliana* with delays of lateral axes functions of the Q/D ratio

The delay parameter can be set different according to the organ type. For example, delays can be shorter for leaves than for internodes. As noted by Wardlaw [1990], «stems and leaves of cereals have a distinct pattern of growth, with the stem below the point of insertion of the leaf only growing after the leaf is fully expanded ». Indeed, for some plants (e.g. radish, chrysanthemum), a few leaves (usually two leaves, called alpha and beta) appear first. Then they are pushed out when internodes elongate afterwards.

3.2.2 Organ geometrical shapes : Allometries

Organ volume is proportional to its weight. It is assumed that the volumic mass is constant for each organ type.

Leaf

Petiole volume is simply proportional to its biomass. Petiole dimensions are not considered. Blade area S_a is proportional to blade volume and therefore to blade weight q^a:

$$q^a = SBM \cdot S^a \tag{3.20}$$

The proportionality constant is called specific leaf weight (SLW) or Specific Blade Mass (SBM) and expressed in g.cm^{-2}. As only blade area is important in the model, blade length and width depend only the shape of the symbol chosen for the 3D visualisation. SBM is an input parameter that can be assessed from measurements on the plant. It can be determined as a function of the environmental conditions or of the growth cycle of leaf appearance.

Internode

Internode is considered as a perfect cylinder. Its shape is determined by the allometric relationship linking the internode length l to its transversal section s:

$$l = a \cdot s^b \tag{3.21}$$

Therefore the length and transversal section can be computed according to the internode weight:

$$\begin{cases} l = \sqrt{Be} \cdot q^{\frac{1+\beta}{2}} \\ s = \dfrac{1}{\sqrt{Be}} \cdot q^{\frac{1-\beta}{2}} \end{cases} \quad \text{with} \quad \begin{cases} Be = a^{\frac{2}{b+1}} \\ \beta = \dfrac{b-1}{b+1} \end{cases} \tag{3.22}$$

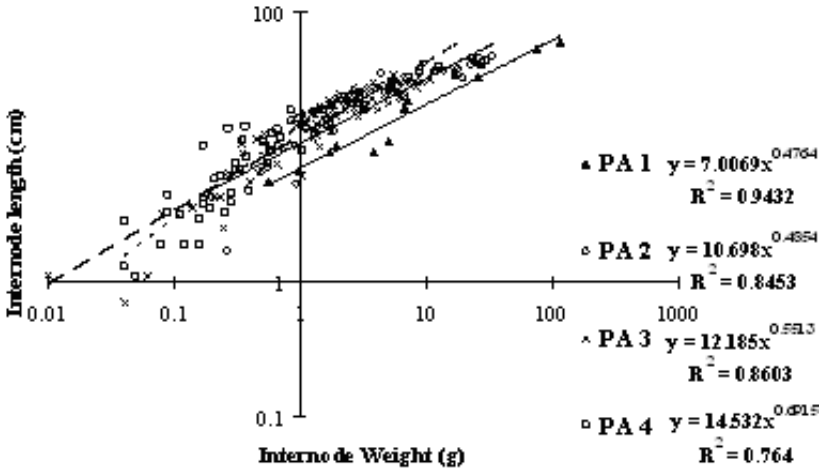


Figure 3.13: Allometric relationships between internode length and weight for young pines (data from Guo Hong, CAF Beijing)

Allometric relationships are determined in a preliminary step to any simulation from the measurement of internode dimensions (e.g. for young pines, Figure 3.13). These measurement are taken from axis tips only: for other metamers, several ring layers cover the pith and its dimensions are not observable. Internode allometric parameters can vary with the metamer rank.

3.2.3 Secondary growth

Specificities of cambial growth

For trees, cambial growth is a predominant factor for determining the quality of the log so modelling of allocation to the ring compartment is of crucial importance. There are numerous evidences in the literature that cambial growth is more sensitive to changes in the growth conditions than primary growth [Collet et al., 2001]. Several authors have reported the relative insensivity of height growth to spacing [Lanner, 1985]. By contrast, diameter growth can be influenced by many factors: competition, change in competition, tree size, age, length of suppression and genotype [Haywood, 2002]. Trees suffering from high competition can produce long primary shoots but usually narrow rings [Nicolini et al., 2001]. In bad light conditions, the growth in height is increased compared to the growth of other compartments [Mäkelä, 1986].
A pioneer modelling work was realized by Mitchell [1975] on Douglas-fir: he took into account crown competition through their projected surfaces and stipulated that ring growth increment is proportional to the number of leaves above the ring position in the tree. Then Houllier et al. [1995] adapted that model to spruce (CEP). In many models (e.g. [Mäkelä, 1990], [Perttunen et al., 1996], [Sterck et al., 2005]), secondary growth is based on the principles of pipe model theory, as introduced by Shinozaki et al. [1964]. It was formulated as follows:

> «A unit amount of leaves is provided with a pipe whose thickness or cross-sectional area is constant. The pipe serves both as the vascular passage and as the mechanical support and runs from the leaves to the stem base through all the intervening strata.»

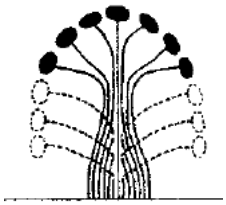


Figure 3.14: Pipe model theory, from [Shinozaki et al., 1964].

In those models, when a new leaf is created, a corresponding pipe unit is created at the same time, running from the leaf position to the stem base (see figure 3.14). The pipe model mechanistically generates the *Pressler* law (Pressler, 1865, in Assmann 1970 [Deleuze, 1996]):

> «The area increment on any part of the stem is proportional to the foliage capacity in the upper part of the tree, and therefore is nearly equal in all parts of the stem, which are free from branches.»

However, limitations of the Pressler rule have been also pointed out. Pouderoux et al. [2001] note that this rule was verified for coniferous but not for oaks where ring thickness (and not cross-sectional area) are related to the foliage volume. According to Deleuze and Houllier [2002], the Pressler rule is too rigid and would need adaptations to be applied with more flexibility. Indeed, the authors report that the stem profile is affected by changes in environmental conditions, by the social status of trees and by fertilization. Pressler rule does not account for those effects, nor for the formation of the buttress.
Recently, some attempts have been made to bridge the gap between process-based models and dendrometric models that output global relationships between radial growth and crown size ([Ottorini et al., 1996]; [Pouderoux et al., 2001]). Deleuze and Houllier [2002] used a reaction-diffusion model for radial growth but it was not related to crown structure and to assimilate allocation in the whole tree. But to consider specific applications such as studies of the biomechanical stresses in the main stem (e.g. in [Fourcaud et al., 2003]), the branching architecture is essential. Indeed, the weight and position of the main branches of order 2 influence the bending of the main stem and the formation of reaction wood.

First step: global allocation

The modelling of cambial growth in GreenLab was developed from those considerations. In a first step, biomass is allocated to a pool used for the cambial growth of the whole plant (step 1 in Figure 3.10). From that pool, biomass is then partitioned for the cambial growth of each metamer. The definition of that virtual pool for ring growth is not physiologically relevant, as no pool of this kind exists in the real plant. However, it allows conceptual simplifications by considering the cambial growth as a whole-plant phenomenon and it is consistent with the modelling point of view adopted for biomass production (common pool at the whole plant level). Thus the ring compartment plays a buffer role: the simulated tree invests more in secondary growth if the conditions are favorable.
Three methods are defined for the global allocation step.

- First method: *Number of leaves*. In that case, the demand for ring compartment is proportionally higher when the number of leaves is higher. The demand of the ring compartment is calculated according to the number of active leaves at cycle t following the equation:

$$D_{ring}(t) = \left(P_0^{rg} + P_1^{rg} \cdot \frac{Q(t)}{D(t)} \right) \cdot N^a(t) \tag{3.23}$$

  where P_0^{rg} is a constant demand (dimensionless) and P_1^{rg} is a mass sink (g^{-1}). As the demand is associated to the number of leaves $N^a(t)$, it is analogous to the pipe model approach but the pipe length is not taken into account (i.e. in the demand of a pipe is calculated regardless of its length).

- Second method: *Model Q/D*. Cambial growth is set to depend on the "vigour" of the plant, which is associated to the main driving variable of the model, the ratio of biomass production of a plant to its demand (Q/D) [Mathieu, 2006]. This ratio is representative of the level of trophic competition inside the plant. A low value of that variable means a high level of trophic competition between organs. In that case, the demand of the ring compartment is:

$$D_{ring}(t) = P_0^{rg} + P_1^{rg} \cdot \left(\frac{Q(t)}{D(t)}\right)^{\gamma} \tag{3.24}$$

  where P_0^{rg} and P_1^{rg} are defined above and γ is a model parameter driving the relative importance of the Q/D variable. The ring compartment acts as a buffer. In good conditions, relatively more biomass can be allocated to cambial growth whereas this compartment is the first to be sacrificed in case of bad conditions. That equation can also represent the ontogenetic changes in cambial growth investment. At young growth stages, a plant generally has low values of Q/D which implies in turn low allocation to the ring compartment by comparison to primary growth compartments. Note that in that case, the total plant demand D must be calculated as the solution of the equation:

$$D = D_0 + D_1 \cdot \left(\frac{Q}{D}\right) + D_2 \cdot \left(\frac{Q}{D}\right)^{\gamma} \tag{3.25}$$

  where D_0 is the constant term of the plant demand, D_1 quantifies the demand part that have a linear variation as function of Q/D (e.g. demand of root compartment or of new organs with sinks depending on Q/D) and D_2 account for variations as functions of power of Q/D. If $D_2 \neq 0$, this equation is solved numerically using the Newton method (see Appendix C).

- Third method : *Model Q*. The demand of the ring compartment is set according to the value of the biomass production at the previous cycle :

$$D_{ring}(t) = P_0^{rg} + P_1^{rg} \cdot Q(t) \tag{3.26}$$

  It means that the higher the production is, the more the plant invests in the ring compartment, relatively to other growth compartments. These three methods need to be tested against experimental data.

Second step: allocation to metamers

Let Q_{ring} denote the biomass allocated to the ring compartment, according to its demand and to the Q/D ratio as in Equation 3.11. This biomass is then partitioned between all metamers for their cambial growth depending on their positions in the plant topology. To overcome the limitations of the Pressler rule or of the Pipe model

theory, we define a flexible sub-model of allocation with two modes of partitioning. In the first mode ("*Pool*"), all metamers can receive the same biomass for their cambial growth: it only depends on their chronological and physiological ages. In the second mode ("*Pressler*"), the metamer position influences the biomass amount that it receives: it is a function of the number (or area) of leaves above the metamer in the plant topology. Those two modes can be mixed together with a coefficient λ in [0;1] (see figure 3.15).

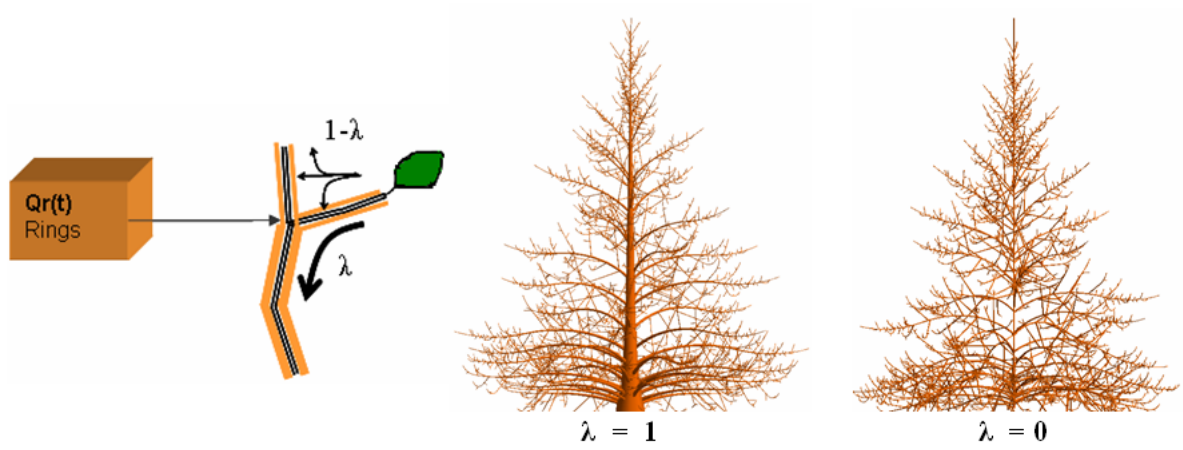


Figure 3.15: Influence of blade position on ring biomass partitioning: effect of parameter λ.

More precisely, the biomass q_{rg} allocated at growth cycle t to a metamer of physiological age p, chronological age n and ontogenetic[1] age (OA) m can be written as follows:

$$q_p^{rg}(n,m,t) = \left(\frac{1-\lambda}{D_{Pool}(t)} + \frac{\lambda \cdot N_p^{a,a}(n,m,t)}{D_{Pressler}(t)} \right) \cdot p_p^{rg} \cdot l_p(n,t) \cdot Q_{ring}(t) \qquad (3.27)$$

where

- $N_p^{a,a}(n,m,t)$ is the number of leaves located above the metamer of PA p, CA n and OA m at GC t (we recall that a list of abbreviations can be found in appendix A). If there is no reiteration (as defined in section 2.4.1), this number is unequivocally associated to each metamer, i.e. metamers characterized by these four indices "see" a fixed number of above leaves. In case of reiteration with a finite reiteration order, one extra index is required to identify every axis, as two axes with the same physiological and chronological ages can bear different number of axes.

[1] Age of the apical bud that initiated the metamer. In most of our simulation cases, the ontogenetic age of a phytomer can be considered as its rank along the branch

- p_p^{rg} are linear secondary sinks for cambial growth (in m^{-1}) that determine the relative demands of metamers of each physiological age. The secondary sink value is calculated by multiplying p^{rg} by the metamer length, $l_p(n,t)$. As they are parameters of an independent allocation sub-model, a reference value can be chosen. For instance, it can be assumed that the secondary sink of metamers of physiological age 1 is $p_1^{rg} = 1$.

- D_{Pool} is the demand of all metamers for their cambial growth calculated with the first mode. It only depends on the number of metamers $N_p^m(n,t)$ and not on their positions. So in that case, the biomass for ring growth of a metamer can be provided by leaves from the whole tree and not only by leaves located above it as in the Pressler mode.

$$D_{Pool}(t) = \sum_{p=1}^{P_m} \sum_{n=1}^{t} N_p^m(n,t) \cdot p_p^{rg} \cdot l_p(n,t) \tag{3.28}$$

- $D_{Pressler}$ is the demand of all metamers for their cambial growth calculated in keeping with the Pressler law: it depends on the relative position of leaves and metamers in the plant topology.

$$D_{Pressler}(t) = \sum_{p=1}^{P_m} \sum_{n=1}^{t} \sum_{m=1}^{t} N_p^m(n,m,t) \cdot N_p^{a,a}(n,m,t) \cdot p_p^{rg} \cdot l_p(n,t) \tag{3.29}$$

  where $N_p^m(n,m,t)$ is the number of metamers of PA p, CA n and OA m at GC t.

The coefficient λ is used to assess the level of influence of the number of leaves and their positions on the partitioning of ring biomass. Another interpretation (although not theoretically equivalent) is to consider λ as the proportion of leaf biomass production (for cambial growth) that can go up in the tree architecture. It is not an exact equivalence because in GreenLab, biomass production is computed at the whole plant level and not for each leaf independently; only global allocation is considered.
The Pressler law corresponds to the case $\lambda = 1$. Indeed, in that case, the sectional area of the annual ring $s_p(n,m,t)$ at cycle t is proportional to the number of leaves above the metamer position m in the plant architecture, as stated by the Pressler rule:

$$\begin{aligned} s_p(n,m,t) &= \frac{q_p^{rg}(n,m,t)}{l_p(n,t)} \\ &= N_p^{a,a}(n,m,t) \cdot \frac{p_p^{rg} \cdot Q^{rg}(t)}{D_{Pressler}(t)} \end{aligned} \tag{3.30}$$

3.2.4 Root system compartment

In the current version of GreenLab, roots are considered as a compartment. No topological structure is considered but this could be done based on the same principles as for the aerial part. However, the associated data are scarcely available.
Three kinds of modes are defined for the calculation of the root demand:

- *Proportional allocation.* At each growth cycle, the root system receives a fixed proportion of the biomass production. This option can be useful when few information is available concerning the root system. It is also a good approximation of the principle of functional balance in steady state: the root growth rate is proportional to that of the shoot [Canell and Dewar, 1994]. It was done for instance on young Chinese pines (joint work with Guo Hong, CAF, Beijing) but the proportional relationship was valid only in a limited range of growth stages (from 1 to 5 year-old, Figure 3.16).

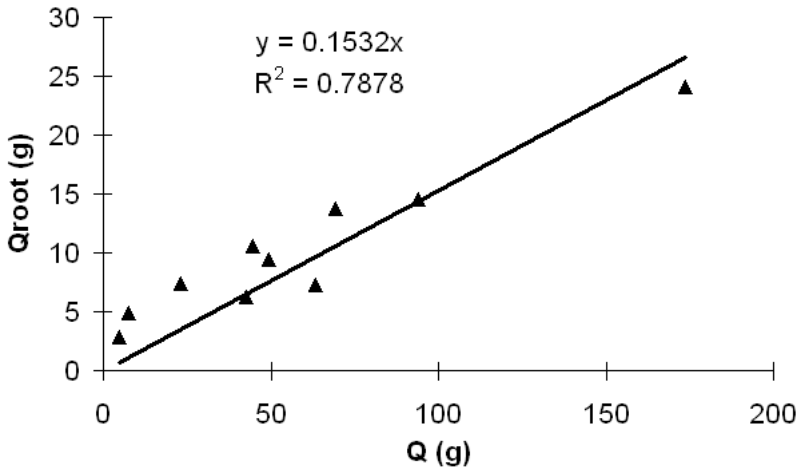


Figure 3.16: Root weight as a linear function of total plant weight for young pines (1 year-old to 5 year-old). Data from a joint work with Guo Hong and Hong LingXia, CAF.

- *Organ mode.* The root system has the same demand as a generic organ. It requires four parameters: its sink strength and the three parameters of its sink variation function (a, b, T_{exp}). This mode was adopted for instance in the fitting of *Arabidopsis thaliana* (joint work with A. Christophe, LEPSE Montpellier). The results for organ sink variations are presented in Figure 3.17. The expansion duration of the root system was set to last the whole plant life. Root sink is twice higher than that of the reference rosette leaf but as there are many leaves and only one root system, the leaf compartment attracts more biomass than the root compartment.

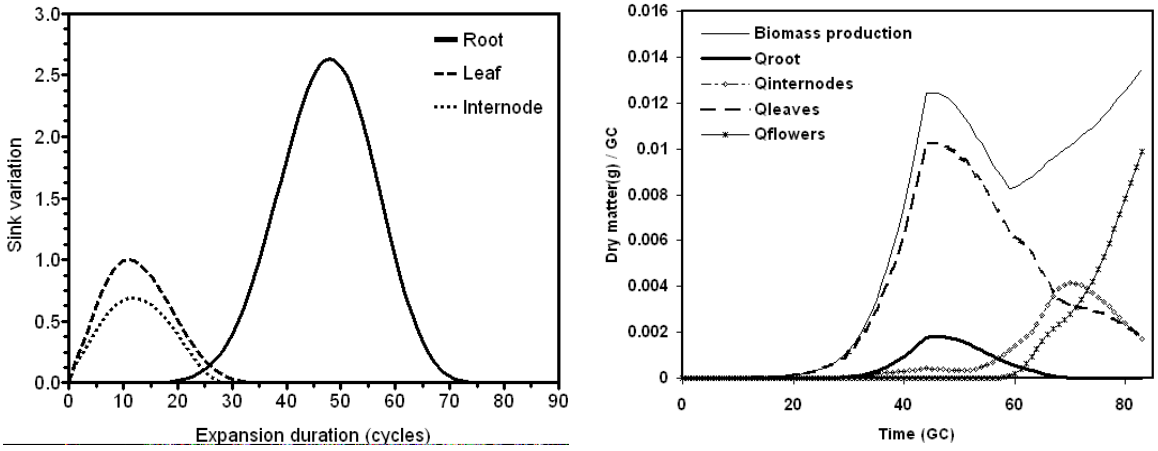


Figure 3.17: Sink variation for rosette leaf, inflorescence internode and root system of *Arabidopsis thaliana* (fruit sink is not represented) and biomass partitioning at each growth cycle into the plant compartments. The parameter values are:

Organ	Sink	T_{exp}	a	b
Leaf	1	40	3.1	5.6
Internode	0.7	30	2.6	3
Fruit	21.5	40	3.6	2.6
Roots	2.6	85	3.6	2.6

- *Mode Q/D or Q.* This mode is analogous to the demand of the ring compartment as defined in section 3.2.3. The root demand is dependent on the plant trophic state (Q/D) or on its production (Q):

$$D_{root}(t) = P_0^{root} + P_1^{root} \cdot \frac{Q(t)}{D(t)} \text{ or } D_{root}(t) = P_0^{root} + P_1^{root} \cdot Q(t) \qquad (3.31)$$

  where P_0^{root} and P_1^{root} are model parameters. However, this last mode is difficult to justify from a physiological point of view. In the current version of GreenLab, roots are considered as mere sinks that do not play any role in the source process. But when the uptake function of the root system is considered, its growth is more likely to depend on soil properties and on the nitrogen concentrations (principle of functional balance) than on a simple trophic balance.

Chapter 4

Model analysis

4.1 Recurrent formula of GreenLab

GreenLab can be written as a discrete dynamic model. Let N_m be the final considered growth cycle. With the notations of Mathieu [2006], it can be formulated as a recurrent system:

$$\begin{cases} Q(t+1) = F(Q(t), T(t+1), P, E(t)) \\ T(t+1) = G(Q(t), T(t), P, E(t)) \end{cases} \tag{4.1}$$

where:

- Q(t) is a vector of the past biomass productions (of size N_m).
- T(t) is a vector of the numbers of organs until cycle t (size N_m). It characterizes the topological development.
- P is the vector of parameters of the model. Its size depends on the number of parameters considered.
- E(t) is the set of control variables until growth cycle t (size N_m).

Note that P consists of two classes of parameters: observable parameters (e.g. expansion durations, specific leaf weight) and hidden parameters (e.g. sinks, resistance). We can also distinguish functional parameters from topological parameters (driving the plant architecture). Eventually, the set of parameters P can be written as :

$$P = (P_t, P'_t, P_f, P'_f)$$

where

- P_t: topological parameters to fit (e.g. GL2 probabilities, GL3 coefficients)
- P'_t: observed topological parameters (e.g. GL1 number of phytomers per growth unit)

- P_f: functional parameters to fit (e.g. source parameters, sink strengths)
- P'_f: observed functional parameters (e.g. expansion durations, SLW)

The first equation of the system 4.1 allows calculating the biomass production at each cycle as a function of the state variables at the previous growth cycles. Let us assume that the resistivities of internodes and petioles are negligible. In that case, the equation of biomass production is Equation 3.5 where the biomass production only depends on the plant total blade area. However, the plant development influences the value of $S(t)$, as blades have to compete with other organs to get their biomass. $S(t)$ is calculated as a function of the cumulated biomass attributed throughout their expansion to leaves that are still active, *i.e.* with a chronological age less than t_a:

$$S(t) = \frac{1}{e} \sum_{i=1}^{t_a} \sum_{k=1}^{P_m} N_k^a(t-i+1) \cdot P_k^a \cdot \sum_{j=1}^{i} \frac{\phi^a(j) \cdot Q_{t-(i-j)-1}}{D_{t-(i-j)-1}} \tag{4.2}$$

where e denotes specific blade mass (as defined in paragraph 3.2.2) and $N_k^a(t-i+1)$ is the number of leaves of PA k appearing at cycle $t-i+1$. Blade sinks are defined by their sink strengths P_k^a and their sink variation functions $\phi^a(j)$ (see paragraph 3.2.1).

Remark: The domain of definition of the sink variation function of every organ o, $\phi^o : j \mapsto \phi^o(j)$, initially defined on $\{0, \ldots, t_{exp}\}$, is extended to $\{0, \ldots, t_a\}$. The function takes the value 0 in the interval $\{t_{exp}, \ldots, t_a\}$.

By replacing $S(t)$ by its expression in equation 3.5, it gives:

$$\begin{aligned} &\forall t \geq t_a, \\ &Q(t) = E(t)\mu Sp \left[1 - \exp\left(\frac{-k}{eSp} \cdot \sum_{i=1}^{t_a} \sum_{k=1}^{P_m} N_k^a(t-i+1) P_k^a \cdot \sum_{j=1}^{i} \frac{\phi^a(j) \cdot Q_{t-(i-j)-1}}{D_{t-(i-j)-1}}\right)\right] \end{aligned} \tag{4.3}$$

This expression shows that the plant development influences the production through the demand and the numbers of organs.

4.2 Limit production

In this section, we study the limit production of the model, *i.e.* the production of the plant after an infinite number of growth cycles. We investigate under which conditions the plant production can reach a stable non-zero phase. Of course the notion of infinite number of growth cycles is purely theoretical but in practice the stable regime can be reached after a phase of growth establishment. In this stable regime, the plant

production is equal (or tends to) the limit production. If the limit production is zero, it means that the simulated plant does not survive. Therefore it is interesting to compare the performances of different architectural models.
Firstly, we notice that:

$$\forall t,\ Q(t) \leq E(t)\mu Sp \tag{4.4}$$

The production is always bounded. For further considerations, we distinguish two cases: (1) demand independent of Q and (2) demand dependent on Q. This last case occurs for instance when sinks of organs or of compartments depend on the Q/D ratio.

4.2.1 Case 1: D independent of Q

Equation verified by the limit production

Let us consider the ideal case of a plant with a fixed deterministic development ($GL1$) with no or negligible ring compartment growing under constant environmental conditions or at least under conditions that tend to stabilize ($\lim_{t\to\infty} E(t) = E_\infty$). If it exists, the limit production Q_∞ must be solution of the equation:

$$Q_\infty = E_\infty \mu Sp \left(1 - \exp\left(\frac{-k}{eSp} \sum_{i=1}^{t_a} \sum_{j=1}^{i} \phi^a(j) \cdot Q_\infty \cdot \sum_{k=1}^{P_m} P_k^a \cdot \left(\frac{N_k^a}{D}\right)_\infty\right)\right) \tag{4.5}$$

where $\left(\frac{N_k^a}{D}\right)_\infty = \lim_{t\to\infty} \frac{N_k^a(t)}{D(t)}$. Its value depends on the plant development. Let us detail the limit behaviour of this ratio.
The total demand of the plant aerial part can be written as follows:

$$\forall t \geq t_{exp},\ D(t) = \sum_{o\in\{a,i,f\}} \sum_{k=1}^{P_m} \sum_{j=1}^{t_{exp}} N_k^o(t-j+1) \cdot P_k^o \cdot \phi^o(j) \tag{4.6}$$

Indeed, the plant demand is equal to the sum of organ demands: there are $N_k^o(t-j+1)$ organs of kind o ($o \in \{a, i, f\}$ with the respective index a for leaf, i for internode, f for fruit or flower) of physiological age k and chronological age j (these organs were born at cycle $t-j+1$). Organs have a non-null demand if they are still in their active expansion phase: $j \leq t_{exp}$ where t_{exp} is the maximum of expansion durations of all kinds of organs. In that case, their demand is $P_k^o \cdot \phi^o(j)$ (see Paragraph 3.2.1).
Hence, we can show that the term in the exponential factor of Equation 4.5 does not increase to infinity. Indeed :

$$D_\infty \geq \underbrace{\sum_{j=1}^{t_{exp}} \phi^a(j)}_{=Cte} \cdot \sum_{k=1}^{P_m} P_k^a (N_k^a)_\infty \tag{4.7}$$

Thus the term $\sum_{k=1}^{P_m} P_k^a \cdot \left(\frac{N_k^a}{D}\right)_\infty$ is bounded. This assertion is verified whatever normalization mode is chosen for the sink variation function (Appendix B). Moreover, we can check easily that the limit of this sum is non null. Indeed, in the $GL1$ version, the number of organs per metamer is constant for a given physiological age. So, choosing $p = Arg\ max\{N_k^a, 1 \leq k \leq P_m\}$, we have:

$$\forall o \in \{a, i, f\},\ \forall k \in \{1..P_m\},\ \left(\frac{N_k^o}{N_p^a}\right)_\infty < K \tag{4.8}$$

where K is a constant depending on the proportionality constants of numbers of organs on metamers. It implies that:

$$\left(\frac{N_p^a}{D}\right)_\infty = \frac{1}{\sum_{o\in\{a,i,f\}} \sum_{k=1}^{P_m} \sum_{j=1}^{t_{exp}} \left(\frac{N_k^o}{N_p^a}\right)_\infty \cdot P_k^o \cdot \phi^o(j)} > 0 \tag{4.9}$$

At least one term of the sum is non null so, all terms being positive, the sum is non null.
We conclude that in the general case (t_a, t_{exp} and Sp taking finite values), the limit production is solution an equation of the following form:

$$Q_\infty = A(1 - e^{-B \cdot Q_\infty}) \tag{4.10}$$

where $A = E_\infty \cdot \mu \cdot Sp$ but B depends on the plant development and on the duration of leaf activity t_a. We have shown above that $0 < B < \infty$. Note that the limit production Q_∞ is strictly inferior to $E_\infty \mu Sp$: this maximal potential limit production is never reached. In practice, if for instance t_a is high, the limit production is close to the maximal one (see section 4.3).

4.2.2 Solutions

Let us define the function $f : \mathbb{R} \to \mathbb{R}$ as: $f(x) = A(1 - e^{-Bx})$. Solving Equation 4.10 amounts to finding the positive solutions of $f(x) = x$. It admits a non-null solution if and only if $f'(0) > 1$, i.e. $\underline{AB > 1}$:

$$\frac{E_\infty \mu k}{e} \sum_{i=1}^{t_a} \sum_{j=1}^{i} \phi^a(j) \cdot \sum_{k=1}^{P_m} P_k^a \cdot \left(\frac{N_k^a}{D}\right)_\infty > 1 \tag{4.11}$$

Grange [2006] (see also [Corless et al., 1996]) has proved that in that case, the solution of this equation is:

$$Q_\infty = A + \frac{1}{B} \cdot W\left(-AB \cdot e^{-AB}\right) \tag{4.12}$$

where W is the *Lambert function*. The Lambert function, also called the omega function, is the inverse function of h defined by $h(w) = we^w$. It means that $W(x)$ is solution of the equation:

$$W(x)e^{W(x)} = x \tag{4.13}$$

As h is non injective, W is multivalued: on the interval $[-\frac{1}{e}; 0)$, $W(x)$ takes two values, as shown in Figure 4.1. One can easily check that for $x = 0$ both solutions $W(x) = 0$ and $W(x) \to -\infty$ satisfy the equation.

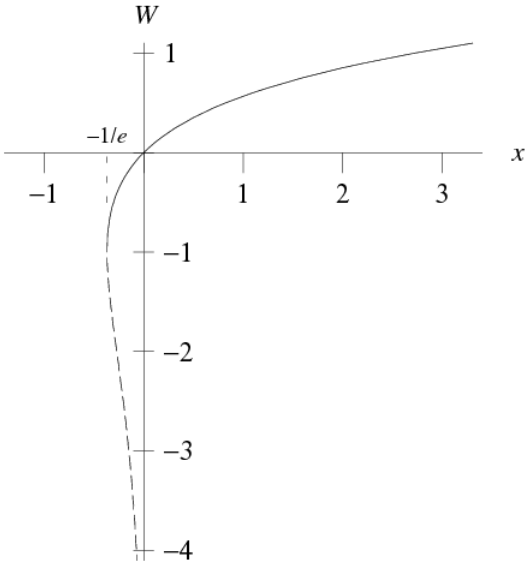


Figure 4.1: Lambert function for $x \geq -\frac{1}{e}$. There are two branches on the interval $[-\frac{1}{e}; 0)$ [Corless et al., 1996]. The continuous line represents the principal branch W_0.

As $AB > 1$, we are interested on the values of W on this interval: $-AB \cdot e^{-AB} \in [-\frac{1}{e}; 0)$. One value is higher than -1 and the other one is smaller. Fortunately, one of the solutions is known: it is $W(x) = -AB$ (that gives the solution $Q_\infty = 0$). Therefore the positive solution $f(x) = x$ is given by:

$$Q_\infty = A + \frac{1}{B} \cdot W_0\left(-AB \cdot e^{-AB}\right) \tag{4.14}$$

where W_0 is referred to as the *principal branch* of the W function [Corless et al., 1996]. It is the non-dotted part of the curve represented in Figure 4.1.
It implies the following comparisons:

$$0 < A - \frac{1}{B} < Q_\infty < A \tag{4.15}$$

The Taylor series for W_0, converging for $|x| < \frac{1}{e}$, is given by:

$$W_0(x) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} x^n \tag{4.16}$$

[Corless and Jeffrey, 1997]. A truncated sum provides good approximations of the values of $W_0(x)$.
Using this development of W_0, we can see that Q_∞ is an increasing function of B. Therefore calculating the value of B for different topological parameters provides a criterion to evaluate the performance of different architectural models. Q_∞ is also an increasing function of A which is coherent with the interpretation of A ($= E_\infty \mu S p$): increasing the environmental variable E or the coefficient of biomass conversion efficiency μ naturally results in increasing the limit production.

4.2.3 Application for some architectural models

In application, we give the expression of the constant B of equation 4.10 for two particular architectural models: Corner and Leeuwenberg models, as represented in Figure 4.2.

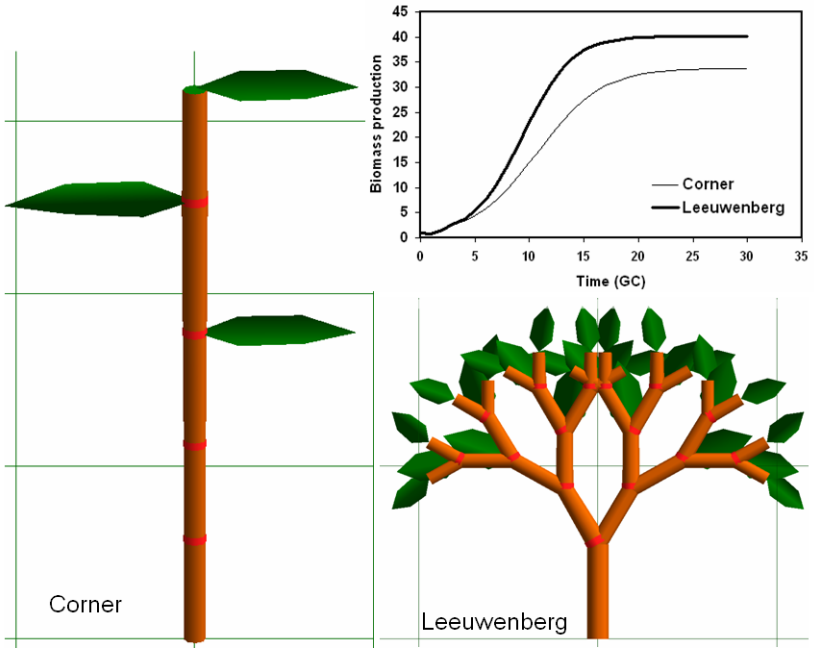


Figure 4.2: Corner and Leeuwenberg architectural models: comparison of biomass productions. The simulated values are consistent with the numerical results found after solving eq. 4.10 as shown in fig. 4.3.

The case of the Roux model can be found in [Grange, 2006]. No fruits are considered and all organs are of the same physiological ages. P^a and P^i represent respectively blade and internode sink strength while ϕ^a and ϕ^i are sink variation functions.

- Corner model: there is one more metamer at each growth cycle.

$$\begin{aligned} &\forall t \geq 1,\ N^a(t) = N^i(t) = 1 \\ &\forall t \geq t_{exp},\ D(t) = \sum_{j=1}^{t_{exp}} P^a \cdot \phi^a(j) + \cdot P^i \cdot \phi^i(j) \\ &B = \frac{kP^a}{eSp} \frac{\sum_{n=1}^{t_a} \sum_{j=1}^{n} \phi^a(j)}{\sum_{j=1}^{t_{exp}} P^a \phi^a(j) + P^i \phi^i(j)} \end{aligned} \tag{4.17}$$

  If $t_a = 1$, B does not depend on the sink variation coefficients and the condition for a strictly positive limit production ($AB > 1$) can be further simplified to:

$$\frac{E_\infty \mu k}{e} \frac{P^a}{P^a + P^i} > 1 \tag{4.18}$$

  A numerical value of the limit production is given in the example of Figure 4.3.

- In Leeuwenberg model, each apical bud gives birth to M new metamers:

$$\begin{aligned} &\forall t \geq 1,\ N^a(t) = N^i(t) = M^{t-1} \\ &\forall t \geq t_{exp},\ D(t) = \sum_{j=1}^{t_{exp}} M^{t-j+1} \cdot \left(P^a \cdot \phi^a(j) + \cdot P^i \cdot \phi^i(j)\right) \\ &B = \frac{kP^a}{eSp} \sum_{n=1}^{t_a} \sum_{j=1}^{n} \frac{\phi^a(j)}{\sum_{j'=1}^{t_{exp}} M^{j-j'} (P^a \phi^a(j') + P^i \phi^i(j'))} \end{aligned} \tag{4.19}$$

  In the illustration presented in Figures 4.3 and 4.2, the limit production for the Leeuwenberg model is higher than that of the Corner model. In that case ($T_{exp} = 3$), it can be shown that as soon as its number of buds per metamer is $M > 1$, the limit production of a Leeuwenberg model is necessary higher than that of a Corner model. As all functional parameters are chosen identical, this difference is purely related to the architectural model through the development. However, this trend would be likely to change if hydraulic resistances of axes were considered.

Note that $\phi^a(j) = 0$ if $j > t_{exp}$.

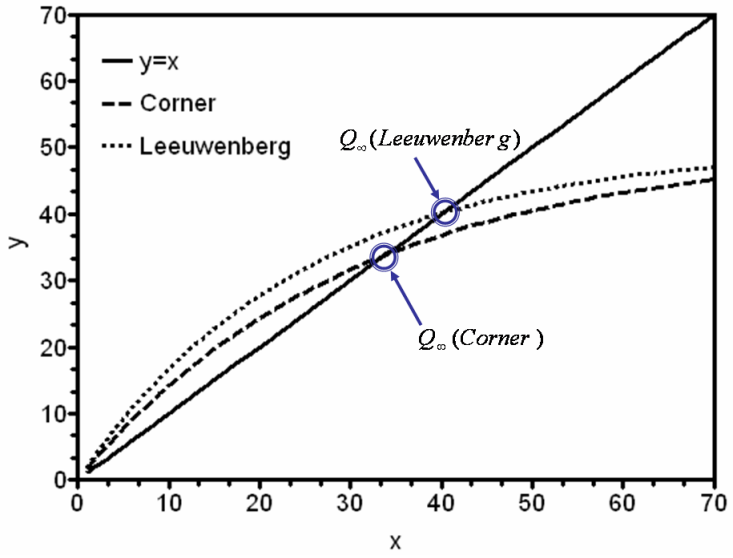


Figure 4.3: Graphical resolution of Eq.4.10: $Q_\infty = A(1 - e^{-B \cdot Q_\infty})$ for Corner and Leeuwenberg (M=3) architectural models. The parameter values are given in table 4.1. Numerically (eq. 4.16), the limit productions are: $Q_{\infty\ Corner} = 33.79$ and $Q_{\infty\ Leeuwenberg} = 40.16$

t_a=3 $Sp = 500$ E=1 k=1 e=0.06 μ=0.1
t_{exp}=3 $P^a,\ P^i = 1$ $\phi^a,\ \phi^i = 1$
It gives: $A = 50$
$B_{Corner} = \dfrac{k \cdot P^a}{e \cdot Sp} \dfrac{2}{P^a + P^i} = 0.033$
$B_{Lee} = \dfrac{kP^a}{e \cdot Sp(P^a + P^i)} \left(\dfrac{3M^2 + 2M + 1}{M^2 + M + 1} \right) = 0.04$

Table 4.1: Table of parameter values of simulations of Figure 4.3.

Remark 1: This case 1 also includes the cases where the demand of the ring compartment does not depend on Q/D. In particular, it includes the case of the model Number of leaves *(Eq 3.23) when* $P_1^{rg} = 0$.

Remark 2: The seed biomass has no effect neither on the limit production nor on the condition $(AB > 1)$ *to reach a stable positive value.*

4.2.4 Case 2: D dependent of Q

The demand depends on the production in cases where compartment demands are functions of the ratio Q/D as in Eqn. 3.24 or of the variable Q as in Eqn. 3.26

(especially for ring or root compartments, sections 3.2.3 and 3.2.4). The conditions to get a strictly positive limit production are calculated differently.
We assume for instance that ring demand follows the mode Q/D . To simplify the study, we assume here that $\gamma = 1$ and that only the ring compartment has a demand function of Q/D. Therefore the total plant demand can be written under the form (see Eqn. C.2 in Appendix C):

$$D(t) = K_1 + \sqrt{K_2 + K_3 Q(t)} \tag{4.20}$$

where:

$$\begin{cases} K_1 = \dfrac{1}{2}\left(\displaystyle\sum_{o\in\{a,i,f\}}\sum_{k=1}^{P_m}\sum_{j=1}^{t_{exp}} N_k^o(t-j+1)\cdot P_k^o\cdot\phi^o(j) + P_0^{rg}\right) \\ K_2 = \dfrac{1}{4}\left(\displaystyle\sum_{o\in\{a,i,f\}}\sum_{k=1}^{P_m}\sum_{j=1}^{t_{exp}} N_k^o(t-j+1)\cdot P_k^o\cdot\phi^o(j) + P_0^{rg}\right)^2 \\ K_3 = P_1^{rg} \end{cases}$$

Case A: The number of leaves increases to infinity

Firstly, we notice that if the number of leaves increases to infinity, the limit is the same as for case 1. The ring compartment has no influence on the long-term behaviour of the model. Indeed, the demand of new organs increases to infinity whereas the ring demand remains constant. The ratio Q/D decreases and the biomass allocated to the ring decreases to zero. However the convergence may be slower than when no ring compartment is present, as shown in Figure 4.4 for the Leeuwenberg model.

Case B: The number of leaves stabilizes

Let us assume now that the number of new metamers tends to a finite limit. For the sake of clarity, we study the case of the Corner model with immediate expansion and one growth cycle of functioning duration for leaves. The demand at GC t can be written as follows:

$$D(t) = P^a + P^i + D_{ring}(t) \tag{4.21}$$

We distinguish several cases according to the mode chosen to compute the ring demand.

- *Number of leaves* mode with constant sink (no influence of Q/D). As there is only one active leaf at each growth cycle, the demand is written $D(t) = P^a + P^i + P^r$ where P^r represents the ring sink. Thus this case comes down to the previous case and the limit production can be calculated using the Lambert function. There is a strictly positive limit production if $\dfrac{Ek\mu P^a}{e(P^a + P^i + P^r)} > 1$. The maximal potential limit value is $Q_\infty^{max} = E_\infty \mu S p$.

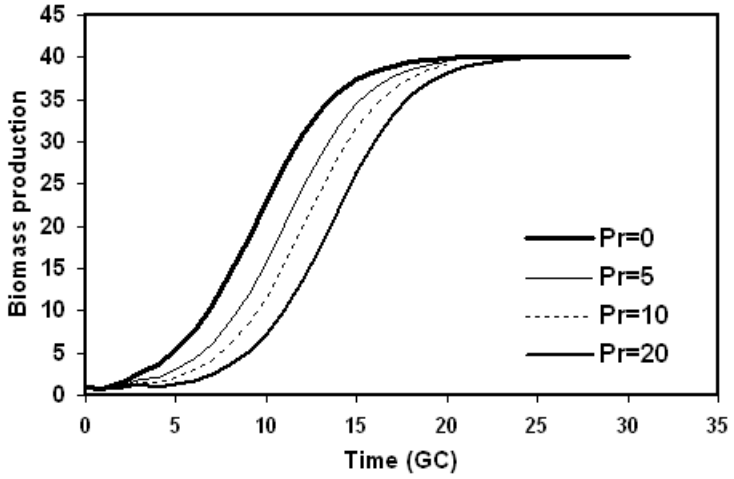


Figure 4.4: Biomass production for the Leeuwenberg model with parameters of table 4.1 with ring demand calculated as: $D_{ring}(t) = P^{rg}\left(1 + \frac{Q}{D}(t)\right)$. The limit production is the same but the largest the sink of rings P^{rg} is, the lowest the convergence is.

- *Q/D Mode.* If the ring demand is proportional to Q/D ratio, the total demand of the plant at GC t is:

$$D(t) = \frac{1}{2}\left(P^a + P^i + \sqrt{(P^a + P^i)^2 + 4P^r \cdot Q(t)}\right) \tag{4.22}$$

The limit production is solution of the following equation:

$$Q_\infty = E_\infty \mu Sp\left(1 - \exp\left(\frac{-kP^a}{eSp} \cdot \frac{2 \cdot Q_\infty}{P^a + P^i + \sqrt{(P^a + P^i)^2 + 4P^r \cdot Q_\infty}}\right)\right) \tag{4.23}$$

After differentiating this expression, we find the following condition for a strictly positive solution:

$$\frac{E_\infty \cdot k \cdot \mu \cdot P^a}{e \cdot (P^a + P^i)} > 1 \tag{4.24}$$

The maximal potential limit value is as in the previous case: $Q_\infty^{max} = E_\infty \mu Sp$.

- *Q Mode.* If the ring demand is proportional to the biomass production Q, we obtain:

$$Q_\infty = E_\infty \mu Sp\left(1 - \exp\left(\frac{-kP^a}{eSp} \cdot \frac{Q_\infty}{P^a + P^i + P^r \cdot Q_\infty}\right)\right) \tag{4.25}$$

The condition for a strictly positive limit production is the same as in the previous case:

$$\frac{E_\infty \cdot k \cdot \mu \cdot P^a}{e \cdot (P^a + P^i)} > 1 \tag{4.26}$$

But here the maximal potential limit value is reduced in comparison to the two previous cases: $Q_\infty^{max} = E_\infty \mu Sp \left(1 - \exp\left(\frac{-k}{e \cdot Sp \cdot P^r}\right)\right)$.

Note that for the modes Q/D and Q, the condition for a strictly positive solution does not depend on the sink of ring compartment P^r. It takes the same value as in the case of no ring demand (Eqn. 4.18, t_a=t_{exp} = 1). This is consistent with the assumption that the plant behaves like a self-regulating system that adapts its ring demand to the growth conditions. It generates an intrinsically stable system.
The maximal value that Q_∞ can take is inferior for the case of mode Q than for the two other cases. For Corner model, the mode Q for ring demand is the one that penalizes the most the growth. Figure 4.5 presents the graphical resolution for the three modes.

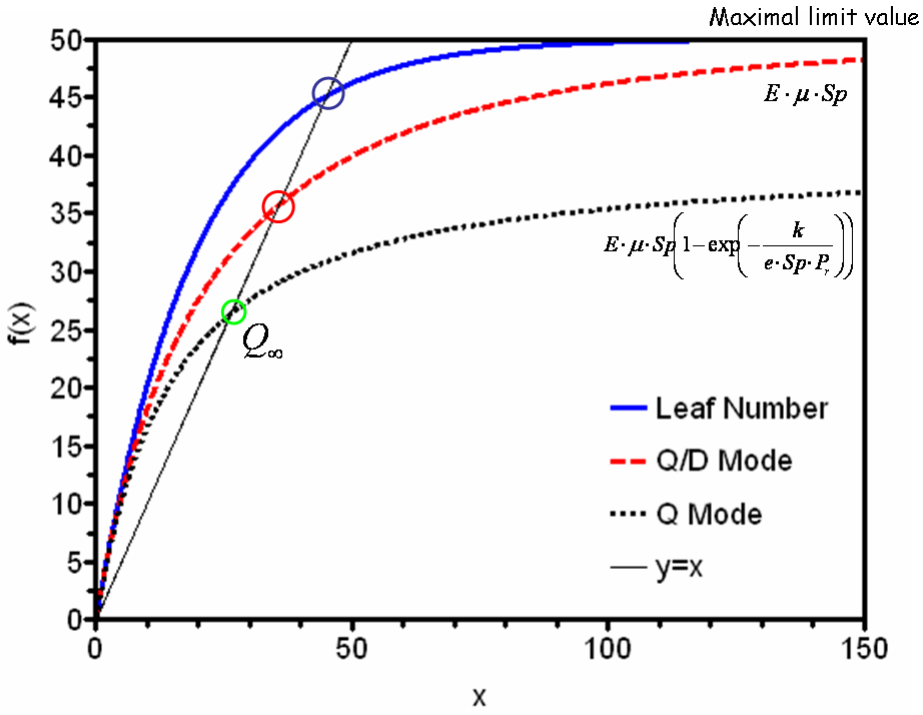


Figure 4.5: Limit production according to the method chosen for ring global allocation.

4.3 Influence of t_a values

4.3.1 Particular case: $t_a = \infty$

This particular case was excluded in the paragraphs above but the corresponding limit production is easy to determine. It means that the leaves are always active. The amount

of photosynthetic foliage increases indefinitely. In that case, the term in exponential factor of Equation 4.5 increases indefinitely and thus:

$$Q_\infty(t_a = \infty) = E_\infty \mu S p \quad (= A) \tag{4.27}$$

4.3.2 General case

Equations 4.17 and 4.19 show that t_a has a preponderant influence in the value of B and therefore on the value of the limit production. Figure 4.6 gives an illustration for a Roux architectural model.

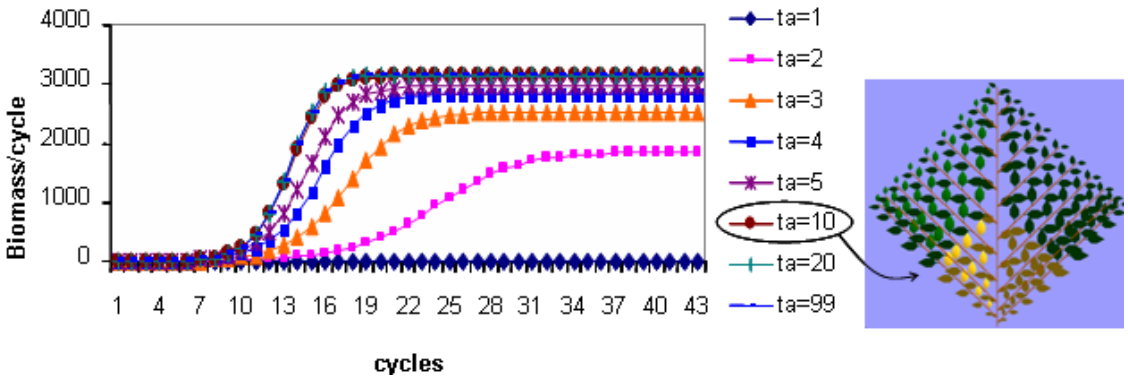


Figure 4.6: Biomass production per growth cycle in Roux model for different values of t_a. For $t_a > 10$, the curves are indissociable.

4.4 Invariances by topology changes

In this paragraph, we explore the consequences of the new formulation of the biomass production equation (Equation 3.5). In particular, we show that under some conditions, plants with different topologies can have the same biomass production at every cycle of their growth (not only the limit production of the steady state but also during the transitory phase !). The interest of that property is underlined in part II where equivalences between plant descriptions at different levels are studied.

4.4.1 Case 1: immediate expansion ($T_{exp} = 1$) of all organs with only one physiological age ($P_m = 1$) and no rings

In that case, there is proportionality between the numbers of organs and of leaves with the proportionality constant k^o. Thus the demand of the shoot part is the sum of the demand of new organs only (internodes with sink P^i and fruits or flowers with sink P^f):

$$D(t) = (P^a + k^i \cdot P^i + k^f \cdot P^f) \cdot N^a(t+1) \tag{4.28}$$

where $N^a(t+1)$ is the number of new leaves at GC $t+1$ (the demand is calculated at end of GC so $D(t)$ account for the new organs appearing at the following GC). Therefore the recurrent equation giving the biomass production at GC t (Eqn.(4.3)) can be simplified to:

$$\begin{aligned} Q(t) &= E\mu Sp\left[1-\exp\left(-\frac{kP^a}{eSp}\sum_{i=1}^{t_a} N^a(t-i+1)\frac{Q(t-i)}{D(t-i)}\right)\right] \\ &= E\mu Sp\left[1-\exp\left(-\frac{k}{eSp}\cdot\frac{P^a}{P^a+k^iP^i+k^fP^f}\sum_{i=1}^{t_a} Q(t-i)\right)\right] \end{aligned} \tag{4.29}$$

Thus, if the initial condition (biomass of the seed Q_0) is fixed, the sequence $(Q_t)_{t\in\mathbb{N}}$

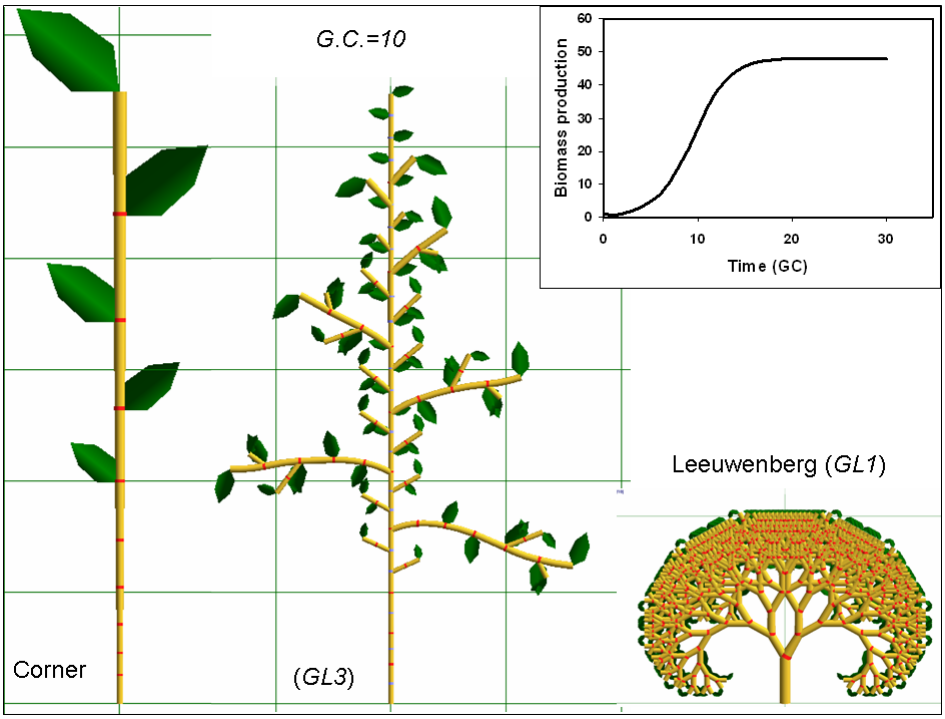


Figure 4.7: Simulation of plant growth with different topologies but same biomass production (Case 1: $t_{exp}=1$, $P_m=1$, $D_{ring}=0$). The functional parameters are: $t_a=5$, $P^a=1$, $P^i=1.5$, $\mu=0.1$, $e=0.06$, $Sp=500$.

is fixed. It means that two plants verifying these assumptions and having leaves with the same functioning duration produce exactly the same biomass amount at each GC, whatever their topology is. Figure 4.7 shows some simulation illustrations with Corner and Leeuwenberg models ($GL1$) and with a plant simulated with the $GL3$ version. The same arguments remain valid if there are several physiological ages in the plant($P_m>1$)

but with a constant proportionality between sinks of organs of each PA, *i.e.* under the following condition:

$$\exists k^o,\ \forall k \in 1 \dots P_m,\ P_k^o \cdot N_k^o(t) = k^o \cdot P_k^a \cdot N_k^a(t) \tag{4.30}$$

4.4.2 Case 2: immediate expansion ($T_{exp} = 1$) of all organs with only one physiological age ($P_m = 1$) and ring demand depending on the number of leaves

In that case, the demand of ring compartment also depends on the numbers of leaves. However, the ring demand is a function of the number of leaves at the current cycle whereas the demand of the other compartment is a function of the number of leaves that appear at the next growth cycle. Therefore it is possible to have similar production for plants with different topologies only if the number of leaves is constant during the plant life span. Such examples are shown in Figure 4.8. This case is of course purely theoretical.

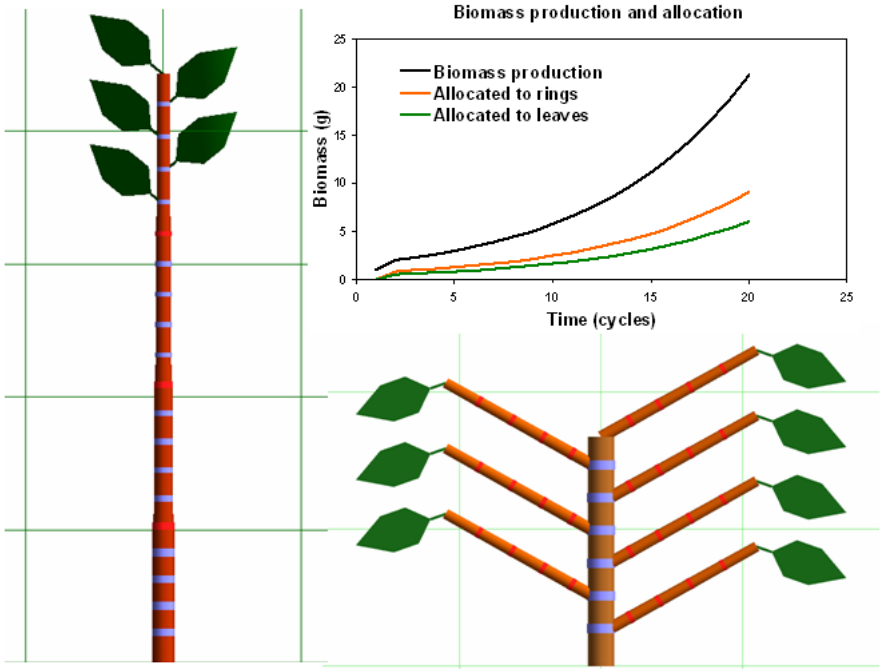


Figure 4.8: Example of two plants with different topologies but exactly same biomass production at all G.C. The number of leaves remains constant throughout each plant life span and thus ring demand is always proportional to the number of leaves.

4.4.3 Case 3: expansion duration of several G.C. ($t_{exp} \geq 1$) with same shape of sink variation function for all organs, one physiological age ($P_m = 1$), no ring compartment

Again, the number of internodes is proportional to the number of leaves so we can write:

$$D(t) = (P^a + k^i \cdot P^i) \sum_{i=1}^{Min(t,t_{exp})} N^a(t-i+1)\phi^a(i) \tag{4.31}$$

Therefore the recurrent equation for the biomass production is:

$$Q_t = E\mu Sp \left[1 - \exp\left(\frac{-kP^a}{eSp(P^a + k^i P^i)} \sum_{i=1}^{t_a} \sum_{j=1}^{i} \frac{N^a(t-i+1)\phi^a(j) Q_{t-(i-j)-1}}{\sum_{k=1}^{t_{exp}} N^a(t-(i-j)-k)\phi^a(k)} \right) \right] \tag{4.32}$$

Intuitively, it would seem possible to simplify this ratio, as the same sink variation coefficients and numbers of leaves appear in the ratio of the exponential factor. We will see that it is not always true (and this can be checked from the examples of Figure 4.2). More precisely, after several rearrangements, the above equation can be written under the form:

$$Q_t = E\mu Sp \left[1 - \exp\left(\frac{-kP^a}{eSp(P^a + k^i P^i)} \sum_{i=1}^{t_a} \frac{Q_{n-i} \sum_{j=1}^{t_a-i+1} N^a(t-i-j+2)\phi^a(j)}{\sum_{j=1}^{t_{exp}} N^a(t-i-j+2)\phi^a(j)} \right) \right] \tag{4.33}$$

One can see that the ratio does not simplify as the number of terms in the sums are different. The residual terms are due to the fact when a leaf dies, its whole biomass ceases activity regardless of the time it was allocated to the leaf.

Residual term due to leaf senescence

An alternative possibility would be to consider progressive senescence of leaf. The total amount of leaf biomass allocated at GC t would cease activity at GC $t + t_a$. In that

case, the recurrent expression simplifies to:

$$\begin{aligned} Q(t) &= E\mu Sp\left(1-\exp\left(\frac{-kP^a}{eSp}\sum_{i=1}^{t_a}\frac{Q(t-i)}{D(t-i)}\sum_{j=1}^{t-i+1}N^a(t-i-j+2)\phi^a(j)\right)\right) \\ &= E\mu Sp\left(1-\exp\left(\frac{-kP^a}{eSp(P^a+k^iP^i)}\sum_{i=1}^{t_a}Q(t-i)\right)\right) \end{aligned} \tag{4.34}$$

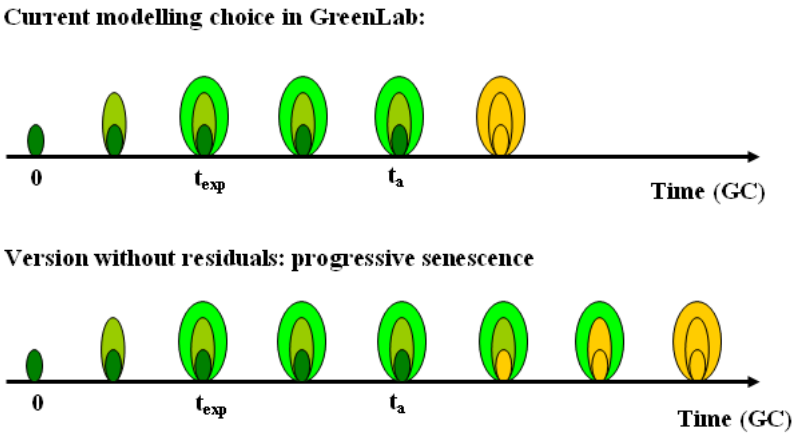


Figure 4.9: Leaf senescence: in the current version of GreenLab, leaf activity ceases abruptly: it engenders a residual term in the recurrent equation of biomass production. An alternative choice would be considering that leaf senescence is progressive: every unit of leaf biomass would remain active during t_a growth cycles.

Thus with this modelling of leaf senescence, it is possible to find plants with different topological development but similar sequences of biomass production. Note that this last equation is the same than Eqn. 4.29: expansion durations and sink variation shapes have no influence on the sequences of biomass production.

Case of infinite functioning duration of leaves

Apart from this question, the simplification can also be done when the functioning duration of leaves (t_a) is infinite. Indeed in this case, leaves never become senescent. Figure 4.10 presents some example of simulated plant growth with same biomass production and allocation to leaf and internode compartments.

Remark: in fact, this last case includes case 1: if there is immediate expansion, it implies obviously that sink variation shapes are identical for all organs. However as case 3 imposes more constraints, we choose to present it separately.

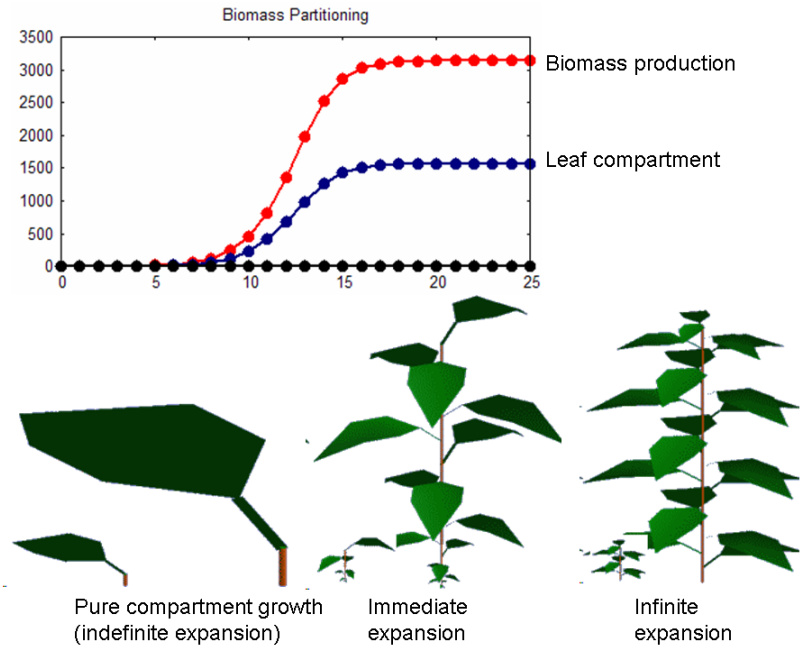


Figure 4.10: Examples of simulated plants with same biomass production and allocation to blade and internode compartments. Leaf functioning time is infinite. Expansion is immediate or infinite. For each plant, all organ kinds have the same shape of sink variation.

4.4.4 Conclusion

To conclude, similar production sequences can be found for plants with different topological development under the following conditions:

- Immediate expansion ($t_{exp} = 1$), one physiological age ($P_m = 1$), no rings (or ring demand proportional to the number of leaves and a constant number of leaves)
- Expansion duration for several cycles ($t_{exp} \geq 1$), one physiological age ($P_m = 1$) and same sink variation shape for all organs, provided that one of the following conditions is verified:
  - Leaf functioning duration is infinite (at least equal to the plant chronological age)
  - Leaf senescence is progressive: every unit of leaf biomass remains active during exactly t_a growth cycles.

Chapter 5

Introducing a genetic model and application to simulation of QTL detection

In this chapter we present an extension of GreenLab for applications in the field of plant genetics and breeding. More details can be found in [Letort et al., 2008b]. We model the chain from genotype to phenotype (set of physical traits) and simulate the inverse procedure (from phenotype to genotype). As illustrated in Figure 5.1, models can play the role of missing link between plant genotype and phenotype. Environment variables are considered as control variables: as they influence the model output variables but not its input parameters (ideally), models can help unraveling the complex genotype $\times$ environment interactions.

5.1 Interest of models for plant breeding

The main objective of plant genetic studies is to link chromosome loci to specific agricultural traits in the hope of increasing breeding efficiency for crop yield improvement. The recently developed marker-assisted selection strategies rely on attempts to identify and quantify the genetic contributions to the phenotype (set of physical traits). The identification of the number and position of loci or genes controlling the target quantitative traits is based on a population of individuals (called mapping population) segregating for the target traits and for molecular markers. Markers are "flags" regularly spaced on the whole genome map and representing intergenic (usually non-coding) short strands of DNA that can be hybridised with their counterparts on the target genome, thereby marking a certain location (see [Ribaut et al., 2001]). Thus it is possible to establish a statistical link between polymorphism at these markers and variability of the target quantitative traits in all individuals of the mapping population. The chromosomal segments, bordered by two adjacent significant markers, are called Quantitative Trait Loci

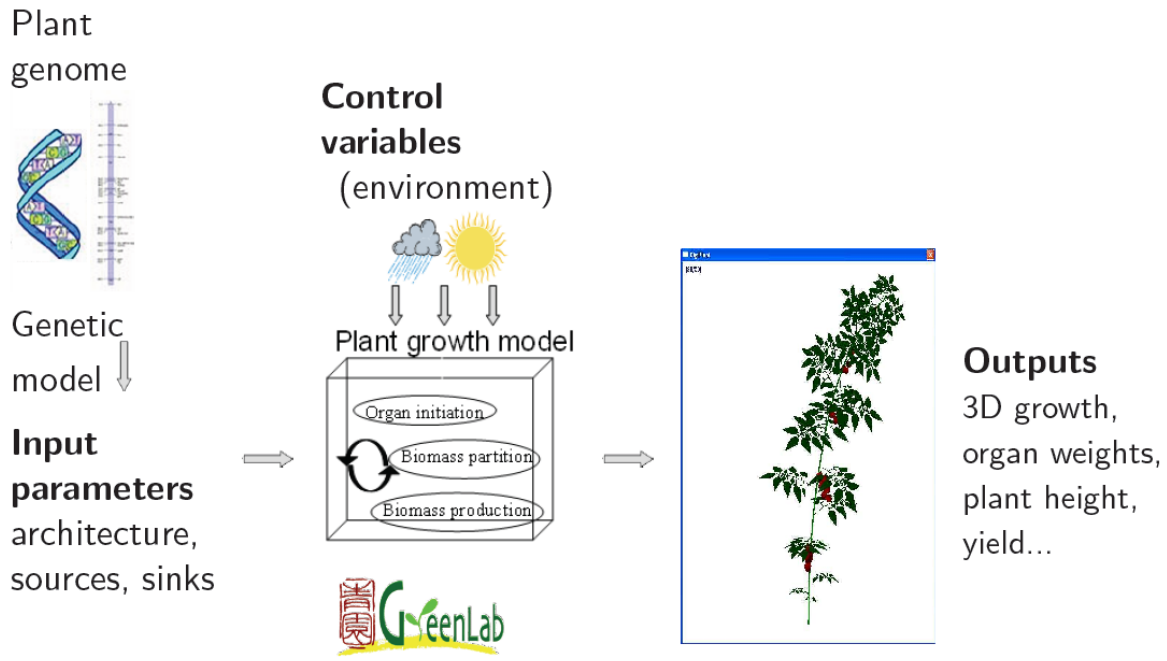


Figure 5.1: Global flowchart of simulation chain from plant genotype to phenotype.

(QTL). They contain the gene of interest but have a confidence interval largely overtaking the gene itself because of the limited power of the classical statistical detection methods.
The main phenotypic traits that are classically studied for crops are yield, duration, plant height, resistance to biotic and abiotic stresses, seedling vigour and quality ([de Vienne, 1998]). Although it has allowed significant advances in crop genetic improvement, there is nowadays a slowdown in yield potential increase for some crops such as rice ([Yin et al., 2003]). One major difficulty lies in the complex interactions between genotype and environment (G×E) since those traits integrate many physiological and biological phenomena and interactions with field and climatic conditions. To overcome this difficulty, a growing interest for the use of ecophysiological models is currently emerging through several studies that underlined the potential interest of building such a link ([Hammer et al., 2002], [Tardieu, 2003], [Yin et al., 2004], [Hammer et al., 2006]).

To deal with the gene level, it seems easier to make the linkage with low-level physiological phenomena at the molecular scale. Some attempts to reduce the gap between genetic and ecophysiological models are bottom-up, as in [Tomita et al., 1999]. But it requires identifying all the genes involved in each process and the multi-scale interactions resulting in the plant growth processes. Tardieu [2003] argued that using gene regulatory networks to simulate complex gene effects on phenotypic traits was not feasible, due to the large amount of unknown information concerning gene role and

regulation rules and to the high number of different genotypes that would have to be analysed.

The top-down approach, considering ecophysiological modelling at a higher organizational level, is more promising. The role of models is to provide helpful tools for the dissection of physiological traits into their constitutive components ([Yin et al., 2002]) and for unravelling the G×E interactions ([Hammer et al., 2005]). For example, Buck-Sorlin [2002] detected QTL for tillering and number of grains per ear in a winter barley population and he integrated them into a morphological growth model. Dingkuhn et al. [2005] linked a peach tree model with QTL but the predictive ability of the model decreased when linked with the genetic model. Despite that unconvincing result, their paper illustrates the interest to test further QTL detection for high level model parameters and emphasizes the necessary condition that those parameters should act independently from each other and be subjected to minimal G×E interactions. A successful work was achieved by Reymond et al. [2003] who focused on the equation linking leaf elongation rate (LER) to meristem temperature. The link between genetic and ecophysiological models was used to predict leaf elongation rate of non-tested combinations of genotypes and climatic conditions, with satisfactory success (the model explained 74% of the observed variability for LER).

The interest of this approach for breeding strategies is quantified in [Hammer et al., 2005] using gene-to-phenotype simulations of sorghum: they linked the yield to four basic traits (duration prior to floral initiation, osmotic adjustment, transpiration efficiency, stay-green), the values of which were simulated under three different environmental conditions according to a genetic model built from the relative information found in the bibliography. The simulation results showed that the predictive power and efficiency of marker-assisted selection was enhanced by the link with ecophysiological modelling. They finally discussed the pertinence of such an approach at the plant scale and the level of detail that may be required for the growth model.

Since the target traits, such as yield, are the results of the whole plant functioning, it is important to study them in association with all the other processes in the dynamic context of plant growth instead of considering them independently from each other. That is why models such as GreenLab can play a role, not pretending that their parameters are directly related to gene expression but assuming that they should, at least, allow detection of more stable QTLs than classically used phenotypic traits. Indeed, parameters for models at organ or plant level already integrate several interacting physiological processes but they are likely to be more stable under diverse environmental conditions than the phenotypic traits that they drive.

5.2 Genetic model: from chromosomes to model parameters

Some of the parameters were chosen to be considered as genetically determined and a simple genetic model was built to introduce a plant genotype into the growth model. To illustrate this study, the GreenLab parameters chosen for the simulations are taken from the calibration results of Guo et al. [2006] and Ma et al. [2007] on *Zea mays* L. The main endogenous parameters can be distinguished on the basis of the stability study made by Ma et al. [2007] but here, twelve parameters were arbitrarily chosen: photosynthetic resistance, blade SLA, sinks of sheaths, internodes and cob, parameters of sink variation function for blades, sheaths, internodes and cob, number of shorter internodes at the plant base, cob rank on the main stem, expansion time for all organs. For sake of clarity, it is assumed that the virtual genome consists of one pair of homologous chromosomes, although maize has in reality ten pairs of chromosomes; the general case is easily deduced. They are represented as vectors whose components are numbers that can take several values (as in [Buck-Sorlin and Bachmann, 2000]), called *alleles*. The rules driving the choice of allele values can be defined by the user in adequacy with the information available about the considered species, such as uneven distribution of genotype frequencies or the skewed distribution of alleles. From the two virtual chromosomes C_1 and C_2, an application f defines the rules of allele expression (dominance or additivity) and then the 'genetic' vector of parameters Y is calculated as a product of matrices:

$$Y = D \cdot A \cdot f(C_1, C_2) \tag{5.1}$$

The components of the vector Y are the endogenous parameters that are assumed to be genetically determined. A is a matrix defining the influence of genes on each parameter, including pleiotropic rules (one gene has an influence on several parameters) or combinations of several gene effects on one parameter. Here the matrix A is of size (12×15), i.e. 12 parameters were genetically determined by a set of 15 genes. D is a diagonal matrix whose coefficients are scaling factors to have range compatibility. Indeed, the i^{th} parameter is defined by its variation around its reference value $Y_r(j)$ (e.g. allele value of 0.9 induces a variation of -10% on the parameters it is related to) so the diagonal coefficients $D(i,i)$ of matrix D are defined as:

$$D(i,i) = \frac{Y_r(i)}{\sum_{k=1}^{n} A(i,k)} \tag{5.2}$$

Simulation of plant reproduction

The reproduction mechanisms are defined for a diploid plant, that is to say a plant having pairs of homologous chromosomes. For each pair of chromosomes, the 'child' inherits one chromosome from each of its parents. This inherited chromosome can

be the result of a crossing-over (exchange of two segments) between the homologous chromosomes of the corresponding parent. Within a population of chromosomes, the number of crossing-over between two markers determines the number of recombinants and is a function of the distance between the two markers. It is assumed here to follow a Poisson law and the points where the cutting occurs are chosen randomly.

5.3 Simulation of QTL detection

5.3.1 Procedure

The previous section introduced the matrix A that represents the effect of genes on the model parameters. For real experiments, determining the values of the coefficients of matrix A is analogous to QTL detection on model parameters since it relates to searching the associations between locations on genome and parameter values. In our study, the model is used to simulate the phenotypic values and the detection of QTL for the endogenous parameters of GreenLab. For application to real plants, a preliminary step would thus be the estimation of the hidden parameters of the model from the organ – or compartment– level experimental measurements on plants (as in [Ma et al., 2007]). Several software packages are used by geneticists to detect QTL, such as QTL Cartographer ([Basten et al., 2005]). In the simulation, the detection of QTL associated to given traits was done on a mapping population that was generated from recombinant inbred lines: the procedure can be represented as in Figure 5.2. First, two individuals are chosen to be the parents, generally with the criterion of being as different and complementary as possible for the considered traits. In the ideal case, those two parents are completely homozygous (i.e. same allele values for all genes) so that all individuals issued from their reproduction have the same genome: one chromosome from one parent line (noted 1111...) and one chromosome from the second parent line (noted 2222...). From that $F1$ generation, several selfings are done until a population whose individuals are homozygous for almost all their genes (97% for the $F6$ generation) is obtained. To study a real population, the measurements are done on that $F6$ generation: geneticists genotype each plant with molecular markers covering the whole genome, and measure the target quantitative traits such as cob weight (details can be found in [de Vienne, 1998]).

In our simulation, each QTL corresponded exactly to one virtual gene and it was placed at a marker location. Markers were regularly spaced all along the chromosome with a distance of 10 cM between two consecutive markers and three markers were intercalated between two successive QTL. Single marker analysis was sufficient to detect QTL, since in this virtual study, QTL were represented by the position of non zero components of the matrix A. The single marker analysis method uses linear regression to test the presence of a QTL at each marker by using a likelihood ratio test whose statistic can

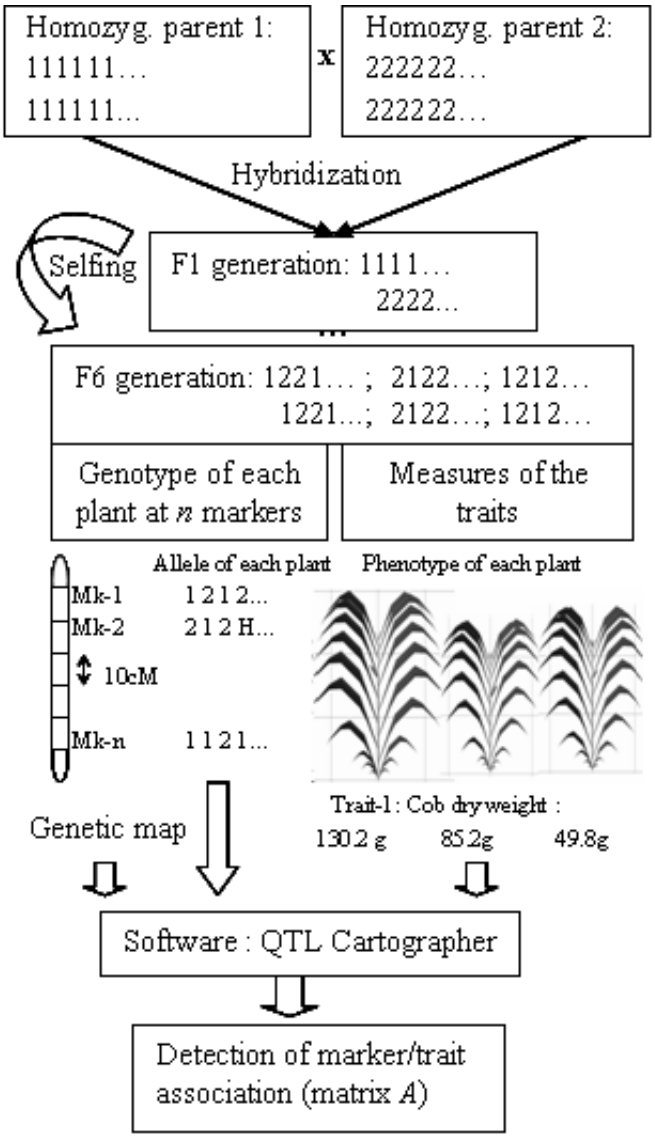


Figure 5.2: Procedure to build data for QTL detection using QTL Cartographer with recombinant inbred lines, from [Letort et al., 2008b].

be converted into a LOD (Logarithm of odds) score as:

$$LOD = -log\left(\frac{L0}{L1}\right) \tag{5.3}$$

where $L0/L1$ is the ratio of the likelihood under the null hypothesis (there is no QTL in the interval of markers) to the alternative hypothesis (there is a QTL in the interval).

5.3.2 Results

The first trait selected is the first parameter of the model, that is to say the first component of the vector Y. If the first line of the matrix A is: (0 0 1 0 0 0 0 1 0 0 0 0 0 0 0), then the LOD curve showing the probability of QTL presence at a marker is presented in Figure 5.3(A). The position of the two detected QTL is denoted by grey triangles. The LOD scores are very high because, in the genetic model presented in the previous section, alleles have a linear effect on the parameter values. When the trait is a parameter depending on three QTLs with different weights, like in the second line of the matrix A: (0 0 3 0 0 0 0 2 0 0 0 0 0 1 0), it gives the curve shown in Figure 5.3(B). Those examples illustrate that, as expected since virtual data are considered, QTL controlling the endogenous parameters of GreenLab are correctly detected by QTL Cartographer.

Use of a virtual population allows comparing easily the QTL detection on model parameters to QTL detection on the classical direct measurements of plant architecture, got from the growth simulation. Cob fresh weight was chosen as a classical phenotypic trait and the relationship between genes and model parameters (matrix A) is defined in Figure 5.3. Figure 5.3(E) gives the results of QTL detection for cob weight: only one major QTL can be detected. The coefficients of the matrix A reveal that its position corresponds to genes influencing blade resistance that have a very strong influence on cob weight in the model. However, in graphs 5.3(A-D), other QTL are detected when considering the parameters independently. It means that, for the classical criteria such as plant height, leaf surface or ear weight, only part of the QTL can be detected, even in the ideal case of our simulation. Indeed, those characteristics are the result of a step by step plant growth process where all the genetic parameters are interactively involved through complex equations.

5.4 Optimization and model-based ideotype

After solving the QTL×parameter interactions, the simulation tool can be used to determine the allelic combination that gives the optimal cob weight. Thus models provide new criteria to define plant ideotypes (defined in [Donald, 1968] as the set of desirable traits that a plant should present to enhance yield or any other objective trait under specified climatic conditions.). Genetic correlations can be added as constraints to represents the cases of gene influence on several parameter values (pleiotropy) with opposite variation or with opposite effects on the cob weight. A genetic algorithm has been implemented to calculate the set of parameter values and their associated genetic values that gives the highest cob weight [Letort et al., 2007b]. The results are gathered in Table 5.1. The variation range for each parameter was ±30% around the reference value. The case where matrix A is the identity matrix (one gene for one parameter) can

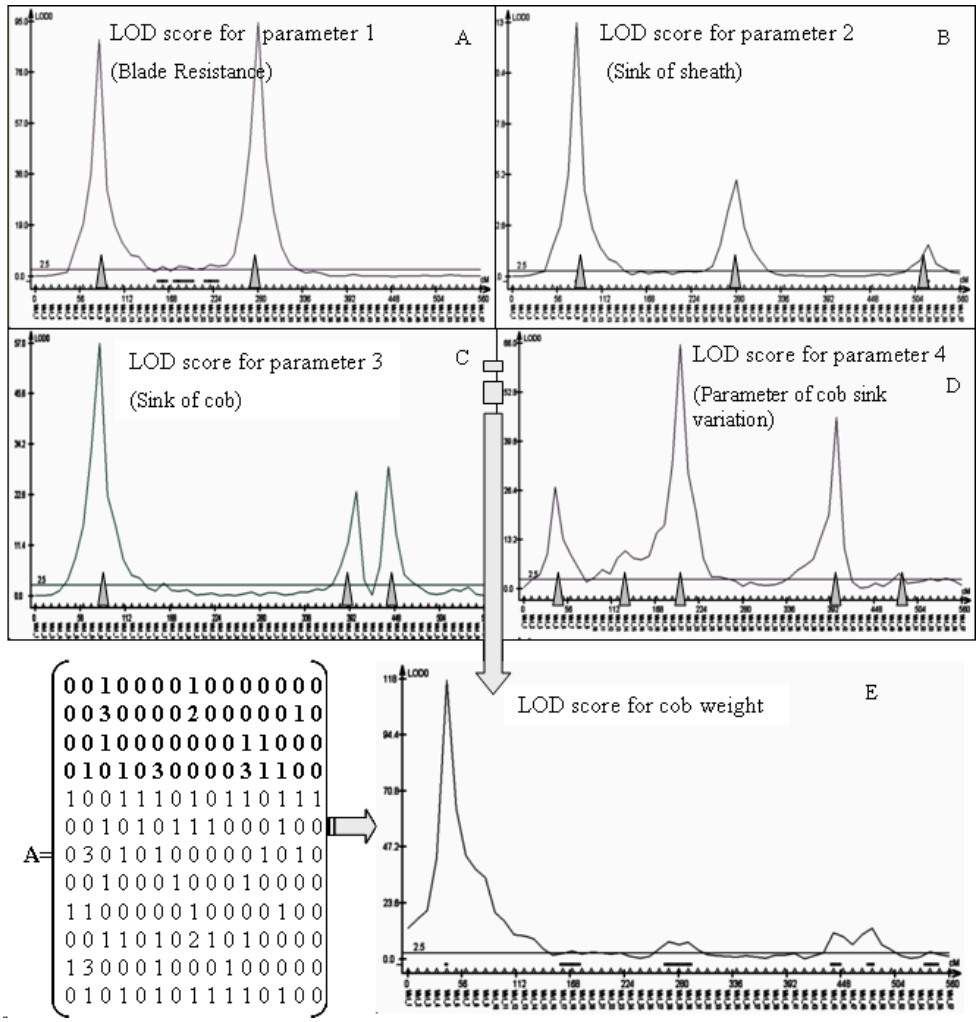


Figure 5.3: QTL detection on four model parameters ($Y(1)$, $Y(2)$, $Y(3)$, $Y(4)$) and on the corresponding cob weight. The curves show the probability of QTL presence at each marker position along the chromosome (X-axis represents marker positions in cM). The matrix A coefficients define the effect of each gene on the model parameters. Grey triangles indicate the most probable QTL positions.

be compared to the probably more realistic case where matrix A defines correlations between the parameters (arbitrarily chosen and represented by superscript indices in Table 5.1: parameters with the same index are controled by the same genes). When genetic correlations are introduced, the potential cob mass is reduced of 38% and the dynamics of biomass partitioning are radically different: the optimal coordinates are not anymore positioned at the search domain boundaries for parameters concerning cob sink and plant topology.

Parameter	Ref. values	Optim. values (A=Id)	Optimal values with correlations
SLA ($cm^2.g^{-1}$) [1]	35.7	46.4 (max)	25 (min)
Resistance [1]	354	248 (min)	248 (min)
Sheath sink [2]	0.7	0.49 (min)	0.91 (max)
Internode sink [3]	2.17	1.52 (min)	2.82 (max)
Cob sink [4]	202	222	161
Blade sink variation param. [2]	0.4	0.31 (min)	0.31 (min)
Sheath sink variation param. [2]	0.53	0.41 (min)	0.41 (min)
Internode sink variation param. [3]	0.79	0.61 (min)	0.61 (min)
Cob sink variation param.[4]	0.62	0.43 (min)	0.5
Nb of short base internodes[5]	6	8 (max)	6
Cob rank [5]	15	9 (min)	14
Expansion duration [6]	12	16 (max)	12
Cob mass (g)	750.2	2325	1428

Table 5.1: Results of genetic algorithm without or with genetic correlations. The superscript indices denote the parameters linked by the same genetic control. It is indicated whether the optimum value is found at the search interval boundary.

5.5 Conclusion

In this study, some important aspects of the chain from genetic model to plant growth model were simulated, ending with QTL detection. It illustrates the statement of Dingkuhn et al. [2005]: "Classical, descriptive phenotyping is based on traits that are too integrative or utilitarian (e.g. yield or leaf area index) and, therefore, insufficiently based on biological functioning to be directly related to gene level information." Indeed, in the simulation, better QTL detection was observed on model parameters than on classical phenotypic traits.

It gives the opportunity to discuss further the assets of functional-structural models, and in particular of GreenLab, as candidate plant growth models for QTL detection on their parameters. In the first papers exploring the possibility to link genetic models to plant growth models, the QTL were associated either to the parameters controlling specific physiological phenomena ([Rey, 2003], [Yin et al., 1999]) or to the parameters of crop models ([Hammer et al., 2005]). However, process-based models present several limitations that could restrict applications in genetics. Indeed, their main drawbacks are: a poor predictive ability of architectural response to environmental factors, such as tillering or organ abortion ([Dingkuhn, 1996], [Luquet et al., 2007]), difficulties to get reliable computation of leaf area index (LAI) which is mostly the main component of biomass production modules ([Marcelis et al., 1998], [Heuvelink, 1999]), an empirical control of environmental stresses at compartment level ([Jeuffroy et al., 2002]), difficul-

ties to deal with the inter-plant variability and to handle the often complex interactions between all the different physiological modules ([Heuvelink, 1999]). These drawbacks result from the fact that process-based models do not take into account plant morphogenesis: at compartment level, since all organs are mixed together, the memory of the growth process is lost and so is the architectural plasticity that reflects the feedbacks between growth and development processes. The endogenous parameters that control both plant development and plant growth are useful key components for yield prediction. Thus they provide new information to renew the breeding process. It provides an adequate strategy to measure plant morphogenesis and to analyze its dynamical biomass production and partitioning.

Several authors (Hammer et al. [2002], Chapman et al. [2003], Tardieu [2003], Hammer et al. [2006]) discussed the properties that growth models should have to expect reasonable chances of success when applied to genetics. Hammer et al. [2002] state that their main quality should be a good predictive ability under various environmental conditions. This property can be verified if the growth model parameters define the environmental control of growth phenomena at the different biological levels. Although further analysis still remains to be done, the predictive ability of GreenLab has been demonstrated in [Ma et al., 2007]. The authors found that parameters were stable along development stages and that the model could explain part of the inter-seasonal phenotypic variability. This paper confirmed the analysis of Dingkuhn et al. [2005] who discussed the use of GreenLab as a link to genetics. The main drawback they detected was the absence of detailed biological knowledge; however, they suggested that it was "worthwhile to test the GreenLab approach in a genetic context, despite its rudimentary physiology". Indeed, Hammer et al. [2002] also emphasized the point that gene-to-phenotype prediction did not require an increase in model complexity, as long as it allowed understanding some key processes so that various combinations of phenotypic responses could be generated through different G×E conditions. The stability analysis of GreenLab parameters tends to reinforce this conviction since it revealed that a small set of chosen rules was sufficient to reproduce plant response to environmental variations [Ma et al., 2007]. In the most recent development of GreenLab, it is possible to simulate the complex plasticity of plant architectural and functional responses to environmental factors [Mathieu, 2006]. Indeed, the parameters are driven by a state variable of the model: the ratio of global biomass supply Q to total plant demand D. The environmental conditions strongly affect the biomass supply and the genetic background of the plant intervenes in the determination of the demand at each growth cycle. That Q/D ratio can be considered as an index of plant vigour and can in particular reflect the environmental impact on plant growth, in interaction with its genome effect. Consequently, the model follows the rules defined by Chapman et al. [2003] that stated that a growth model should include "principles of responses and feedbacks" to "handle perturbations to any process and self-correct, as do plants under hormonal control when growing in the field" and to "express complex behaviour (...) even given simple

operational rules at a functional crop physiological level".

Another key point is that QTL detection implies heavy data processing on populations of high individual numbers. As in most models, some GreenLab parameters (e.g. organ sinks) cannot be directly measured on plants: those hidden parameters have to be estimated from experimental data collected with destructive measurements. The data collection process for each individual can seem tedious if done on complete measurement [Guo et al., 2006] but, as shown in [Ma et al., 2007], the number of needed data can be reduced by methods of aggregation or samplings at different levels. Also, the speed of the fitting procedure is a key factor for processing the large size populations required for QTL detection. Thanks to its mathematical formalism, the inverse problem can be computed. GreenLab is associated to a dedicated fitting tool for parameter estimation that relies on the generalized non linear least squares method [Zhan et al., 2003], which allows a very fast resolution (usually, ten iterations are sufficient and the computation time is generally a few seconds).

Finally, it is worthwhile to anticipate what could be the limitations in the use of GreenLab for QTL detection. First, the model's ability to discriminate genotypes with close allelic composition is an important issue [Tardieu, 2003] and depends on the accuracy of the fitting procedure. Also, the level of required accuracy still needs to be determined. Other criteria such as geometrical shape of organs might need to be taken into account, since it is one of the main features used by breeders to differentiate genotypes. In their generic framework for combining crop modelling and QTL mapping to select the best crop ideotype for a specific environment, Yin et al. [2003] particularly recommended to test the growth model under several environments: thus the G×E interactions would be analysed in a biological way and not only statistically as in classical genetic models. Concerning the GreenLab model, testing under several environments has been undertaken in [Ma et al., 2007] but this step should be further investigated. Moreover, the integrative scale of the growth model may be too large. The basic rules that drive plant growth would thus be unlikely to be the direct expression of independent genes, even if they proved stable in various environmental conditions. Indeed, Luquet et al. [2007] investigated the phenotypic impact of a single-gene mutation in the genome of the 'Nipponbare' rice cultivar. They used a model simulating phenotypic plasticity through resource allocation by introducing an internal competition index for the plant. Apart from detailed observations of differences between the growths of mutant and wild cultivars, the estimation of model parameters highlighted that many traits affected by the mutation closely interacted and it was difficult to reconstruct their causal chronology. It means that some traits can be artificially associated to the same QTL even though the underlying gene influences only one physiological function of the plant. Using growth model at the plant level can thus induce artificial pleiotropic effects since the determination of some parameters could be driven by common primary mechanisms [Yin et al., 2003].

A genetic algorithm was used to optimize the parameters in order to get the highest

cob weight for maize. One advantage of this kind of optimization algorithm is that it can take into account complex constraints (by defining the viability of individuals) and multi-objective criteria (with weighted fitness values, for example). Thus, if one single allele has combined effects on the phenotype, with positive influence on some traits and negative on others, the algorithm can help to find the best compromise. Here, the optimization procedure was realized on twelve parameters that were considered as genetically determined but in a complete study, more parameters, and their interacting effects, could be included. In the same way, new constraints should be added to have more realistic optimized values. Considering for example plant height, the biomechanical constraints in the internodes were not implemented, thus allometric relationships for internodes were also kept constant and the optimization algorithm gave a sink value for internodes as small as possible. Therefore, the optimization criteria should be adapted and made more complex to answer specific objectives on real species. But such optimization results are anyway an interesting contribution of modellers to breeders' work, even if the model relies on simplifying assumptions. The modeller can determine the best allelic combination of genes controlling a given trait through the model under specified conditions. Then the production of the genotype can be more or less difficult depending on the positions of the considered genes and the distances between them, but breeders have developed strategies to separate closely linked genes, involving large segregation populations to get and select the proper recombinant. In any case, it is extremely useful for genotype building to have an idea of the value of virtual ideal genotype without having really to build them, especially in case of pleiotropy when compromises have to be done. This approach could broaden the set of morphological, physiological, biochemical and phenological traits commonly used to characterize ideoptypes, as defined by Donald [1968] and Rasmusson [1987]. Using model parameters to build ideoptypes should help overcoming the limitations due to environment pressure on QTL detection [Beattie et al., 2003]. Their exploitation in breeding programs, however, is conditioned by their heritability, by the level of genetic variations in the populations and by the genetic correlations among them [Reynolds et al., 2007].

To conclude, GreenLab parameters should have higher heritability since they are expected to be less dependent on the environmental conditions and to be more direct gene expression than classical phenotypic measurements. Optimization algorithms allows determining the key parameters influencing the yield, even when complex genetic correlations are introduced, and providing new promising criteria for ideotype definition.

Part II

Model identification and multi-scale analysis

Chapter 6

Interest of simplified architecture for trees

In this part, we raise the problem of GreenLab parameter identification in the case of plants with complex topology. In this introduction, we discuss the use of GreenLab for specific applications in forestry field (adapted from Letort et al. [2008a]). In a second and third chapters, we present the fitting procedure and the adaptations required to deal with complex branched plants. In a last chapter, we study theoretically the different levels of simplifications of the model.

6.1 Characteristics of tree-level models for stand growth

Predictions of stand growth for forestry applications traditionally rely on models related to the theory of population dynamics and on statistical analysis of correlations between different descriptive variables. Those models are classically divided into three categories (Munro 1974): (i) stand-level models, (ii) distance-independent tree-level models and (iii) distance-dependent tree level models. Regarding the resolution scale, tree level seems a promising scale. Indeed, results obtained from models with lower resolution (stand level) can also be obtained by aggregation of models outputs with higher resolution (tree level). When the associated data are not available, a stand simulator can generate a coherent virtual stand structure as it is done for instance in the SILVA simulator [Pretzsch, 2002]. Tree-level models provide more flexibility to deal with heterogeneous stands and new management practices [Pretzsch, 2002]. Most of those models incorporate a simplified description of the crown shape and are based on definitions of site indices and competition indices. An usual way to simplify the crown shape is to consider only its exterior surface and to fit the equation of a species-dependent geometrical surface [Pretzsch, 2002].

6.2 Why introducing tree architecture

Parameterizing of these models usually requires a high number of observations. A possible reduction of the experimental work could be gained by inserting more physiologically based processes to improve the mechanistic ability of the models. In parallel, some of the descriptive variables (e.g. crown surface, competition index) could possibly be calculated with more accuracy if a less coarse description of the tree architecture was taken into account. The simplified representation of the crown may be sufficient to compute the average light intercepted by trees but the variations of the crown shapes under different growth conditions cannot be accurately represented. Indeed, crown shapes result from the combined influences of branch appearances, morphogenesis, positions and orientations. The competition index of a tree is likely to be dependent on its crown architecture, especially in heterogeneous stands where species-specific architectural patterns can play a predominant role in determining the interactions effects. Moreover, a major field of study concerns the biomechanical stresses in the trunk that contribute to determining the value of the log. To build predictive models of the biomechanical stresses, it is crucial to know the insertion points of branches, their orientations and weight dynamics ([Fourcaud et al., 2003], [Jirasek et al., 2000]). Besides, more precise information is provided by botanical analysis concerning the species-specific architectural trends, such as the 23 architectural models defined by Hallé and Oldemann (1970). Such knowledge could provide a useful framework to define simplified species-specific rules driving the topological development of trees [Barthélémy and Caraglio, 2007].
The objective of functional-structural models (FSMs) is precisely to simulate interactively the architectural development of trees and their physiological functioning ([Sievänen et al., 2000]; [Prusinkiewicz, 2004]). But despite the research effort involved in their development, the use of those models remains confined to research and teaching contexts and forestry applications are scarce [Le Roux et al., 2001]. Besides the problem of simulation efficiency, a reason could be that parameterization and validation of such models remain a critical point due to the complexity of tree architecture. Indeed, although many models can provide a very accurate description of the tree at organ scale, only global, aggregated or sampled measurements are reasonably expectable. For instance, LIGNUM has been parameterized for Scots pine [Perttunen et al., 1996] and Sugar maple [Perttunen et al., 2001] considering independently the physiological processes involved and comparing the model output to more aggregated data (e.g. taper curve, height, leaf area). Moreover, any model inversion procedure relies on a large number of direct model simulations so a time-consuming simulation algorithm is not suitable for an efficient calibration.
In this context, it is important to find a balance between the complexity of the simulation rules and the adequate level of simplifications. Similarly, the level of details chosen for the variables of interest (as outputs of the simulation) should be coherent with the experimental data collected on real trees. This is the objective of this part.

Chapter 7

Fitting with complete target

7.1 Target description

The classical set of data included in a GreenLab target consists of data at organ level. It is consistent with the simulation outputs that provide organ weights and dimensions.

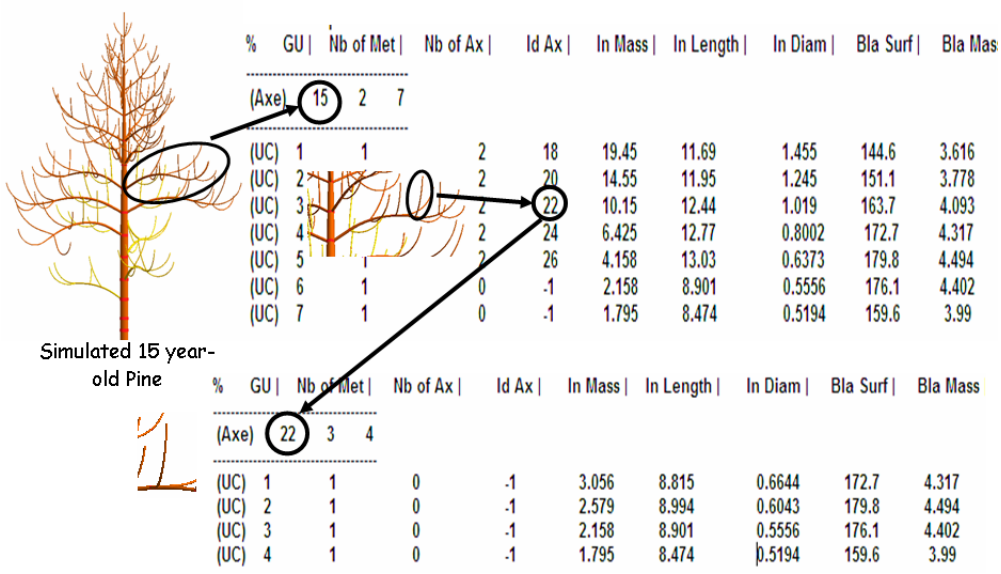


Figure 7.1: Complete target data on a 15-year-old virtual pine: every branch is described at G.U. level for all branching orders.

It consists of:

- A complete description of the plant topology: number of metamers per growth unit, number, position and physiological age of each axis. Axes are characterized by their physiological age (PA) and chronological age (CA).

- For each growth unit, the average weight of internodes, leaves and fruits; average length and diameter of metamers; average blade area.

Each axis of given PA and CA is decomposed into its constituent metamers and measured following a same procedure, as illustrated in Figure 7.1 with the example of a 15 year-old virtual pine. Every branch is described at G.U. level for all branching orders. With this level of details, even measurements of simple single-stem plants represent a tedious work. It takes two hours to two persons to measure single-stem maize at final stage for each existing metamer [Ma et al., 2006].

GreenLab parameters are considered species-specific. To reproduce the complete growth dynamics through time, several stages can be included in the target (*multi-date fitting*). As measurement are destructive, a given plant cannot be measured twice. If the target plants are of the same cultivar and grown in same conditions, they are considered as clones and fitted as a single plant at different stages of its life. If their growth conditions differ, the targets are considered as different plants (*multi-plant fitting*). In both cases, a single set of hidden parameters is obtained.

Comparison fresh/dry biomass Organ mass data consist of either fresh biomass or dry biomass. Indeed, the resulting sink values are simply proportional to the dry matter content of organs compared to that of leaves:

$$P^o(DW) = P^o(FW) \cdot \%DW^o / \%DW^{leaf}$$

Where $P^o(DW)$ is the sink strength value estimated from dry weights of organs, $P^o(FW)$ is the sink strength value estimated from fresh weights and $\%DW^o$ represents the dry matter content in organ o [Louarn et al., 2007]. That relation was con-

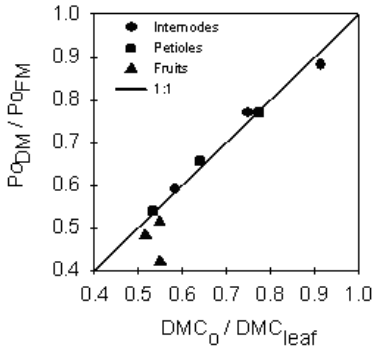


Figure 7.2: Comparison between sink ratios and dry matter content ratios, from [Louarn et al., 2007]

firmed by data on greenhouse tomatoes grown under different densities at the Chinese Agricultural University (Figure 7.2, Dong Qiaoxue and Gaetan Louarn).

7.2 Optimization procedure

Objective function. In section 4.1, we have presented the GreenLab recurrence equations under the generic form of a discrete dynamic system. Let us consider now the system representing the entire plant growth. Inputs are model parameters, P. Control parameters are the same as in Equation 4.1 (mainly environmental variables). Outputs Y are numbers of organs and their weights at each growth cycle. As we are concerned here by the problem of parameter identification, we distinguish observable parameters from hidden parameters. Observable parameters are for example those controlling the default topological rules (defining the physiological age of potential buds on the metamers) or physiological parameters that can be directly measured on plants, such as expansion duration of organs or specific blade mass (SBM). They are taken as model inputs. The values of the hidden parameters need to be estimated using the following procedure.
The model can be formally represented as a function F of the vector of hidden parameters $U \in \mathbb{R}^p$. Given a set of observed data $Y \in \mathbb{R}^n$, we search an estimator $\widehat{U}$ minimizing the following objective criterion:

$$J =^t (Y - F(U))\Omega(Y - F(U)) \tag{7.1}$$

where Ω is a symmetric positive-definite matrix of observation weights. For a linear model ($F(U) = A \cdot U$, A with constant coefficients), the best unbiased estimator is given for $\Omega = \Sigma^{-1}$ where Σ is the covariance matrix of the error vector, ϵ (assumed to be a vector of random centered variables). If Σ is unknown, an iterative procedure provides an estimator [Houllier, 1999].
Although our model is not linear, it can be locally approximated by a linear model and therefore the same method is applied with several iterations. In practice, the vector Y consists of n_i observations for each variable y_i, denoted $(y_{j,i})_{j=1..n_i}$. If errors are independent, Ω is diagonal and its coefficients can be chosen as:

$$\Omega_i = \frac{\sum_{j=1}^{n_i} y_{i,j} - \widehat{y_{i,j}}}{n_i} \cdot \frac{n}{n-p} \tag{7.2}$$

Generalized Least Squares Algorithm. The parameter estimator $\widehat{U}$ is computed after several iterations of the following relationship:

$$\widehat{U}_{n+1} = \widehat{U}_n + \left({}^tX\Omega X\right)^{-1} {}^tX\Omega \left(Y - F(\widehat{U}_n)\right) \tag{7.3}$$

where X is the Jacobian matrix of F (the partial derivative for each dimension are numerically computed using the finite difference method). As our model is not linear, several iterations are needed to converge to the solution.

This method can be applied for parameters with continuous variations and generating continuous variations of $F(U)$ on the search domain. In the deterministic version of the model ($GL1$), the plant development as function of thermal time is set as input. Other environmental variables have no effect on the development. This is reasonable for some plants such as maize [Ma et al., 2007] or *Cecropia* (see section 7.3.2) that have stable development rates. For these plants, only functional parameters need to be assessed. Their topology is fixed from observations. Therefore, the development (numbers of organs) is not considered as target data and the objective function is continuous for the functional parameters.
When $GL2$ or $GL3$ versions of the model are considered, plant development is not fixed. Therefore, development can be included in the set of target data. Moreover, plant development is a non-continuous function of the topological parameters. Indeed, the corresponding data are numbers of organs at different growth stages. These variables take only discrete values, even for continuous variations of their driving parameter. Equations of $GL3$ (e.g. Eq2.11) involve rounded values or integer parts. In that case, the generalized least square method cannot be applied as the objective function is not differentiable. We resort to heuristic algorithms.

Heuristic algorithms. Several algorithms have been implemented in the DigiPlante software: simulated annealing [Kirkpatrick et al., 1983], particle swarm optimization [Shi and Eberhart, 1998], genetic algorithm [Sastry et al., 2005], tree annealing [Bilbro and Snyder, 1991], tabu search [Cvijovic and Klinowski, 1995].
In practice, we use mainly simulated annealing and particle swarm optimization as they have reasonable convergence time and robustness.

7.3 Some applications

We present applications of parameter identification for plants with a branched structure but simple enough to be completely described at the organ level: young Chinese pines (*Pinus tabulaeformis*, joint work with Guo Hong and Hong LingXia from the Chinese Academy of Forestry (CAF) of Beijing) and *Cecropia sciadophylla* (joint work with Patrick Heuret and Camilo Zalamea with the help of Valentin Bellassen, Susanne Braun, Oyétoundé Djiwa, Valentin Le Tellier (CIRAD)).

7.3.1 Young Chinese pines

Modelling assumptions

Like for many temperate trees, the growth cycle duration was set to one year. Shoot elongation last only a few weeks of the first year after bud breakout. Therefore we made the assumption of immediate expansion in the model, i.e. organs primary growth lasts

only one growth cycle. From our observations and from the literature, needles remain active during three growth cycles. We consider that needle biomass and activity of a shoot are constant during these three years, although the shoots usually loose part of their needles and although the remaining needles are less efficient.
The growth unit is regarded as a single virtual metamer and the set of needle that it bears is considered as a leaf. A simple allometric relationship taken from Yang and Yang [2004] links the virtual blade area S^a (needed to compute photosynthesis) to the needle biomass q^a:

$$S^a = 12.4 \cdot q^a$$

Physiological ages were attributed according to branching order.

Measurements and targets. Data were collected on eight Chinese pine saplings. These included two one-year-old trees, two two-year-old trees, two three-year-old trees and two five-year-old trees. They were taken from the open-field nursery garden in Shisanling forest farm, Beijing (39°50'N;116°25'E) in the spring 2006.
The eight plants were gathered into three target files according to their branching topology. Plants with similar topology were considered as repetitions of the same plant (at different growth stages). The simulated topology was set identical to that defined from the measurements for each of the three virtual trees. We assumed that all trees were grown under the same conditions which is reasonable as they were taken from the same plantation.

Results

The needle-internode ratios (assumption of constant sink strengths). This study offered the opportunity to test a strong assumption of the model: sink strength of organs of the same type and belonging to the same PA-based class are constant. We compared internode mass to needle mass of every one-year-old growth units. Results are shown in Figure 7.3. We found that the assumption of constant sink ratios was

Param.	Meaning	Values			
		PA 1	PA 2	PA 3	PA 4
Be	Internode allometry	62	127	148	211
β	Internode allometry	-0.01	-0.12	0.1	0.38
P^i	Internode sink	0.8	0.38	0.11	0.02
P^a	Blade sink	1	0.71	0.28	0.07
e	Needle "SLW"		0.08 g.cm^{-2}		

Table 7.1: Sinks and allometry coefficients for young Chinese pines.

reasonable. Only for PA 1, we observed a decrease of the needle to internode mass ratio in time. For all PAs, the slope coefficients revealed that the average ratio of needle

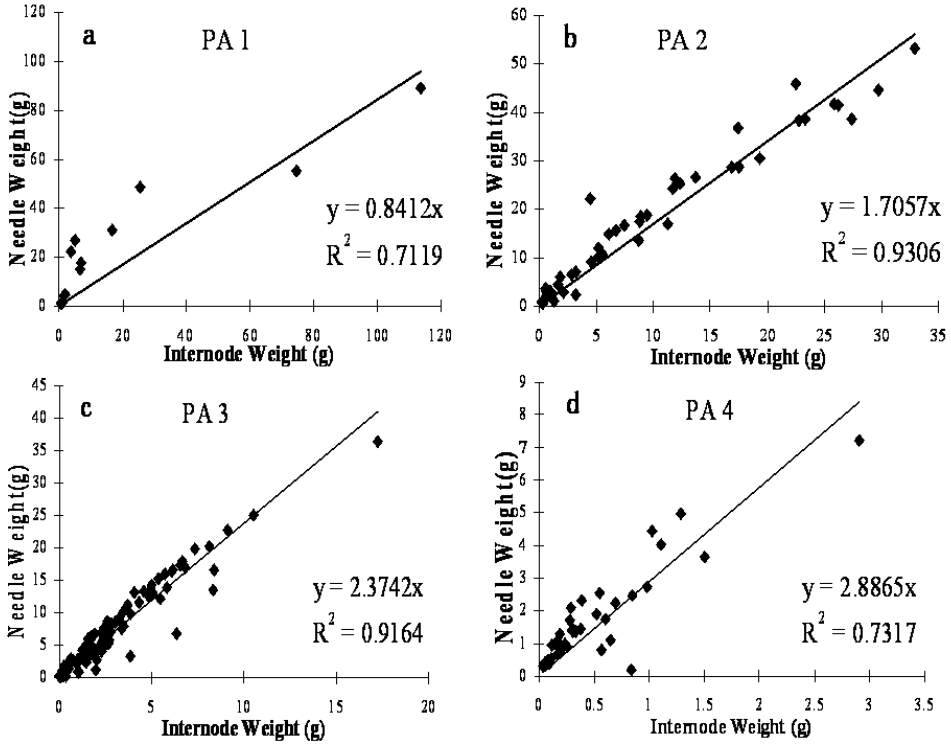


Figure 7.3: Internode-needle mass ratios on young pine shoots [Guo et al., 2007] for each branch order (physiological age, PA). Joint work with Guo Hong, CAF.

mass to internode mass increases with branching order. It means that for Chinese pine saplings, allocation to internodes decreases with branching order to the benefit of needles of the shoot. Sinks and allometry coefficients can be found in table 7.1.

Fitting results. The complete set of fitting graphs can be found in [Guo et al., 2007]. We present only the results on internodes in Figure 7.4. The fitted parameter values are gathered in table 7.2. The choice of the model for ring biomass global allocation (3.2.3) was based on successive trials with each method. The mode Q/D allowed the smallest final weighted mean square error. Concerning ring biomass partitioning at the metamer level (3.2.3), the fitted value of λ is $\lambda = 0.14$. It means that biomass partitioning for ring growth is mostly done regardless of leaf positions. This result is reasonable because the trees are very young (no more than five year-old), so biomass allocation is likely to be managed at the whole-plant scale.

To conclude, this study revealed that it is possible to reproduce the growth of several young pines with a single set of parameter values. These young pines can be considered as mere repetitions of the same tree even if their topologies are different, since they were grown in similar conditions.

Param.	Meaning	Value	CV(%)
r=$1/\mu$	Needle resistance	0.64	3.5
P_0^{rg}	Sink of ring pool (g)	2.64	47.5
P_1^{rg}	Slope for ring sink variation	0.77	41.6
λ	Parameter for ring partitioning	0.14	74.8
p_2^{rg}	Linear density of PA 2 ring sinks	0.11	39.7

Table 7.2: Fitted parameter values for young Chinese pines. P_0^{rg} and P_1^{rg} concern biomass allocation to the ring pool (global step) that conforms the second mode (Q/D). λ and p_2^{rg} concern biomass partitioning to cambial growth of each metamer. The linear density of sinks of PA 3 and PA 4 branches was close to zero.

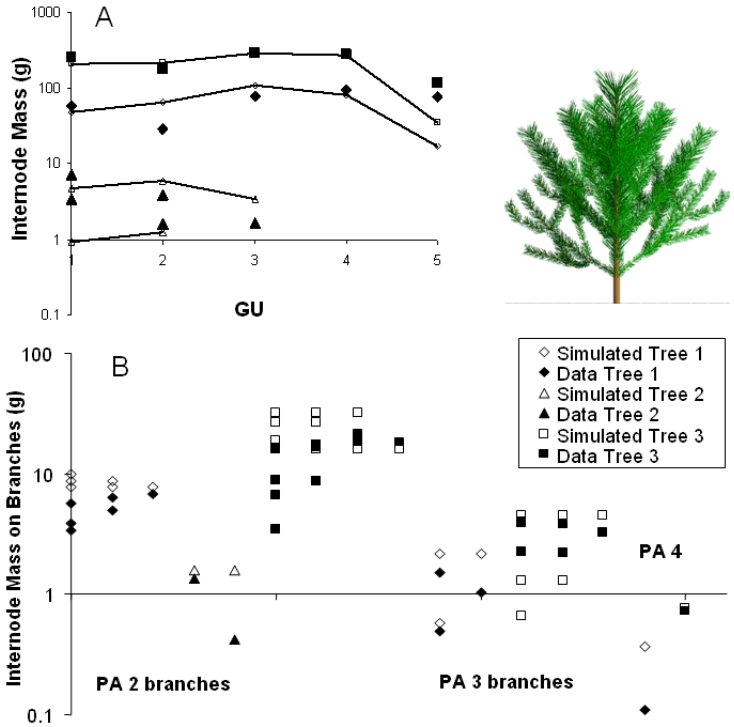


Figure 7.4: Fitting results for internodes of young Chinese pines (Data CAF) for the main stem (PA 1: graph A) and for branches (PA 2, 3, 4: graph B). In graph B, data are gathered per branch: all simulated (empty symbols) and measured (black symbols) internode mass of a branch are represented with the same x-absciss value.

7.3.2 *Cecropia sciadophylla*

Cecropia is a pioneer species that appears preferentially in forest gaps with high light levels.[Zalamea et al., 2008] have studied the phenology of *Cecropia sciadophylla* and have developed methods to assess tree age which provides also accurate ways to estimate

the age of the disturbance that caused the initial forest gap. It also revealed strong rhythmic patterns: synchronous branching periods cyclically alternate with flowering periods. These patterns might be linked to the dynamics of trophic competition during plant growth: [Mathieu et al., 2008] has shown that similar cyclic patterns can be generated as emergent properties in the *GL*3 version of GreenLab. Therefore a study was undertaken to assess the parameter values of this species. Eleven young individuals

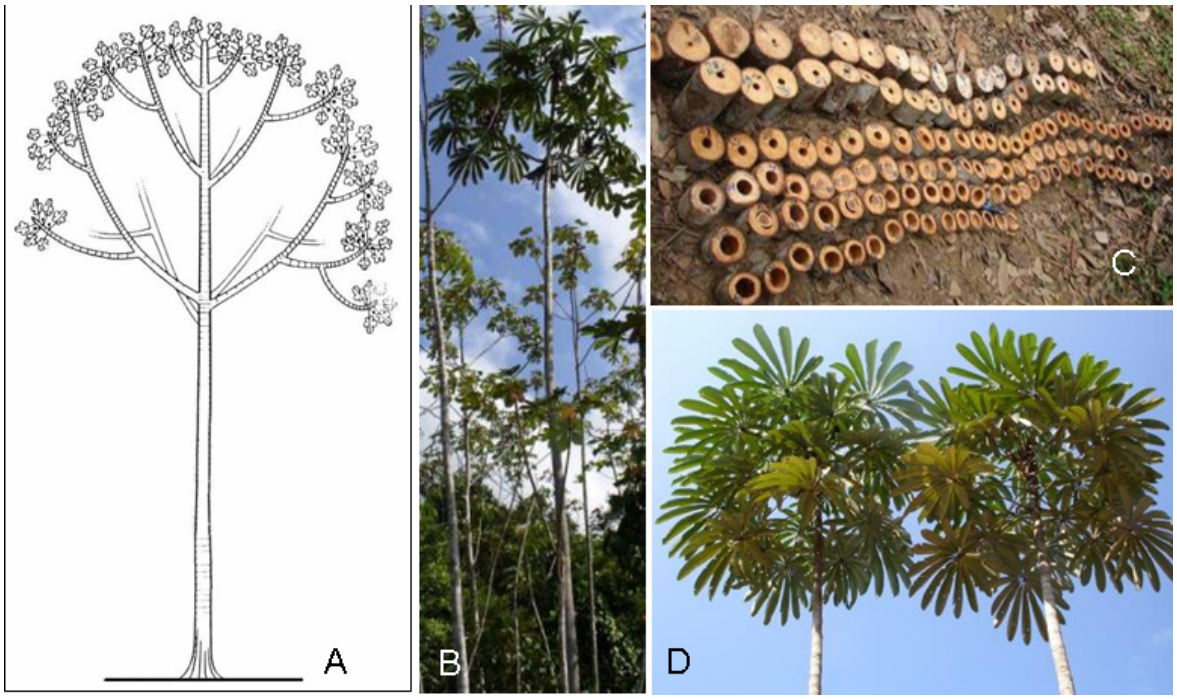


Figure 7.5: *Cecropia sciadophylla*. A: Rauh architectural model; B: A nine-year-old individual; C: phtytomer succession; D: leaf organization (light interception surface)

of *Cecropia sciadophylla* were measured in September 2007 in French Guiana near Saint-Elie road (5° 30' N, 53° W). This is a joint work with Patrick Heuret and Camilo Zalamea with the help of students from the "Forêt Tropical Humide 2007" module. *Cecropia* has a simple architecture with orthotropic axes (Rauh's model [Heuret et al., 2000], see Figure 7.5A) made of a succession of nodes and internodes whose length and associated lateral productions remain visible and measurable over years (Figure 7.5C). In average 25 phytomers are emitted each year. Each node bears three lateral buds that potentially give rise to a branch (central bud) and two inflorescences [Zalamea et al., 2008].
All phytomers have the same physiological ages (organ weights are identical whatever the branching order, see figure 7.6). However, at branch emergence, a "transitory phase" can be observed during which phytomer characteristics vary. Then there is stabilization to a "steady state". From our observations, we can hypothesize that during the transitory phase, the following characteristics are affected:

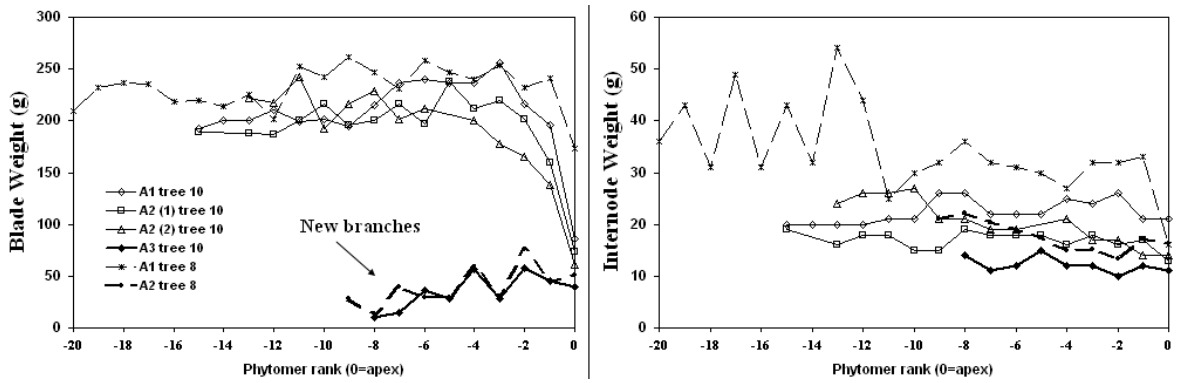


Figure 7.6: *Cecropia sciadophylla*: tips of branches. In their transitory phase, branches invest less biomass in leaves than they do in their steady state. Therefore creating new branches induces a decrease of the trophic balance that reaches back its steady state after a while.

- internode allometries vary: just after branch emergence, internodes are longer and with smaller diameter. Then the internode shape progressively tends to that of the steady state

- pith diameters increase linearly (the slope being presumably related to environmental conditions: a quick increase was observed for individuals grown in favorable environment). This phenomenon can be observed on Figure 7.5C.

- just after branch emergence, internode mass is about 50% of that of the main stem, blade and petiol mass are about 15% of that of the main stem, as shown in Figure 7.6.

- leaf functioning duration is about smaller than that of leaves on the main stem (need to be confirmed with further measurement)

Most of these characteristics might be related to meristem size: its diameter is smaller at branch emergence. It can also be analyzed as an efficient strategy in terms of light interception: as *Cecropia* bears large leaves, the space around the stem is already saturated so newly created leaves would insignificantly increase the light interception surface (see Figure 7.5D). Thus it is a good strategy to privilege long internodes even at the expense of blades: these blades would not be strong contributors to the plant photosynthetic activity. This also implies that new phytomers are probably more sink components than sources. Therefore setting in place a new branch is a costly operation for trees.
These qualitative hypotheses can be investigated at the lights of the fitting results on branched individuals. During the transitory phase, simulated sink strengths vary with

rank as:

$$P^o(k) = P^o \cdot (1 - a_o^k) \tag{7.4}$$

where $P^o(k)$ is the sink strength of organ o at rank k and a_o ($a_o \in [0;1]$) is a parameter depending on the organ type. In the fitted tree shown in Figure 7.7, the maximal Q/D ratio reached during a period decreases after the first branch appearance (at GC=229). It is consistent with the assumption that branch appearance induces an increase in the plant demand which penalizes the tree growth for a while before it can recover its previous trophic balance.

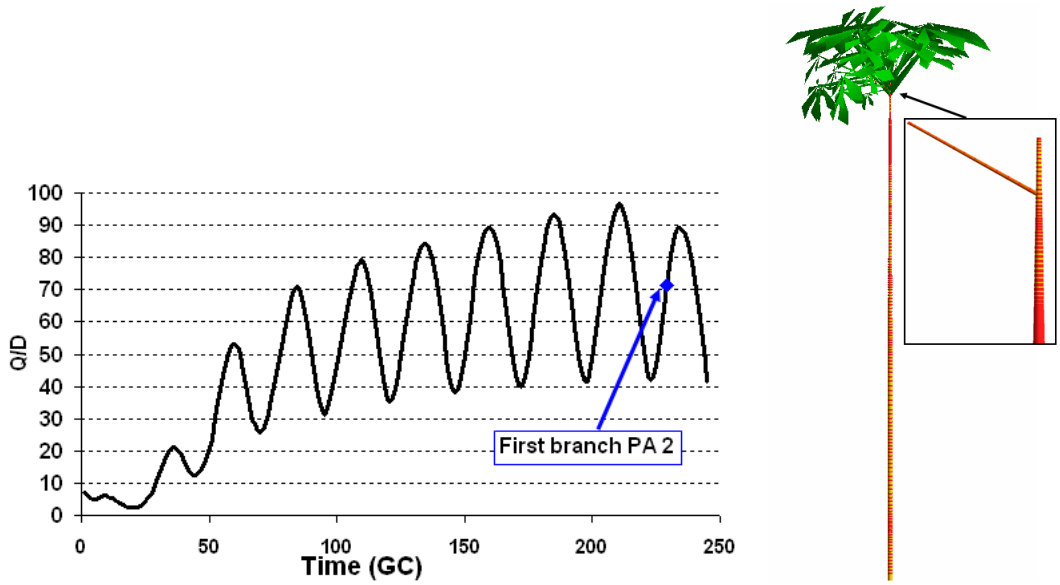


Figure 7.7: *Cecropia sciadophylla*: variations of the Q/D ratio and influence of appearance of a new branch (from fitting results of tree nb 8, 10 year-old). Joint work with Patrick Heuret and Camilo Zalamea, CIRAD.

Chapter 8

Fitting with simplified targets

Ma et al. [2007] pointed out that even for plants as simple as maize, «the use of many complete target files (describing observations on all metamers at a specific date) results in an unreasonably large experimental effort ». The measurement protocol followed for the previous fitting exercises of GreenLab (e.g. [Zhan et al., 2000], [Guo et al., 2006],[Kang et al., 2007b]) is simply not applicable for plants with complicated topology and large numbers of organs such as bushes or trees. Moreover, due to complex stochastic events interacting during plant growth, such a detailed representation would not be relevant. Reproducing the exact growth of a particular plant has no interest. These remarks have led our efforts to define more adequate target formats, fitting procedures and experimental protocols.

8.1 Different levels of simplified targets

This section only describes the different kinds of targets that have been considered in the present work. Some of them have been applied to real plants. Target files consist of information used to describe the real plant to analyze. It is also the set of possible criteria used to validate the simulation by comparing the virtual plant and the real one. We classify the data into three main categories according to their nature:

- *Biomass data*. This denomination covers measurements of organ (or compartment) weights and sizes (geometrical dimensions).
- *Topology data*. It consists of information on numbers of organs and their hierarchical organization (physiological age, topological connections, architecture...)
- *Ring data*. This is a specific category for trees where ring widths can be measured.

The set of data used to describe a plant for GreenLab analysis usually requires one target of type *biomass* and one target of type*topology*. If available, measurements on rings can be included in addition.

8.1.1 Simplified measurements of biomass

Three levels of aggregation of biomass data have been defined. The choice of these definitions have been led by considerations on the measurement practicability and on their potential applications (forestry and biomechanical stresses). They are consistent with the level of details of inputs and outputs of forestry models.

Cumulated target (B0)

The simplest level of simplification consists in considering the plant as a mere pool of organs of different kinds. This is a classical representation of the plant in process-based models, especially for crops. A compartment is denoted as the set of organs of the same kind (leaves, internodes, fruit...). No topology is considered: organs of different chronological and physiological ages are mixed together. Thus the cumulated target B0 ("level 0 for Biomass") contains weights of the different compartments of the plant: wood, leaves, fruits, flowers.
That kind of target was used to reduce the experimental work needed for the fitting procedure on Maize by Ma et al. [2006]. Their target included several growth stages. The young stages were described as cumulated target while only the last stage was described at organ level. This simplification implies a significant gain in experimental work: it takes two hours to two persons to measure a maize at maturity stage at metamer level while it takes only 20 minutes for measurements at compartment level. The authors show that this kind of fitting provide the same accuracy for parameter values as a fitting on complete targets for all growth stages. The same method was applied in [Mathieu, 2006] for fitting of rice with tillers.

"Sparse" target. Another kind of simplification was introduced in Ma et al. [2006]. For one of the measured growth stages, the target contains *sparse data*: only three out of the 21 phytomers are measured. These measurement were sufficient to retrace the entire morphological dynamics of maize growth. Note that including data collected at phytomer level, even partially, is necessary to follow the evolution of individual organ weights and to assess the values of sink variation functions.

Target "Lollipop" (B1)

This level of simplification was designed for trees. In forestry models, one of the products of interest is the trunk. So we define a particular target including detailed measurement of the trunk and including only compartment data for the crown (Figure 8.1). More precisely, this target consists of:

- mass, length and diameter of internodes of the trunk growth units; mass of blades and flowers or fruits, if they exist, on the trunk growth units; internode numbers of each growth units and their number of lateral branches.

- for the whole crown, total mass of internodes, leaves, fruits and flowers (if they exist)

An interest of this level of target is that it is consistent with the level of details available from forestry inventories. It allows incorporating data that would be generated by classical forestry models.

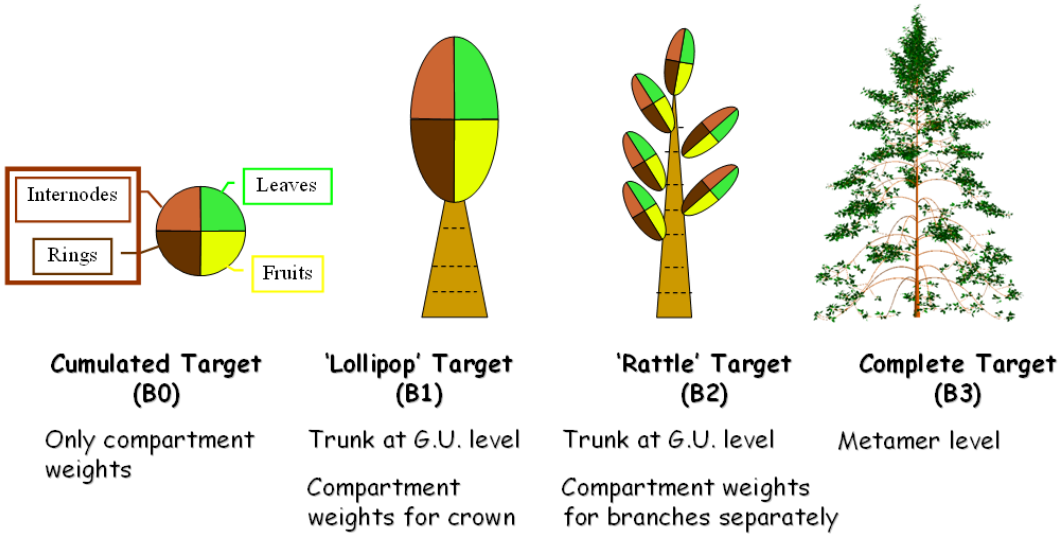


Figure 8.1: Schematic representation of the three level of biomass data aggregation.

Target "Rattle" (B2)

This target is an extension of the previous one. Indeed, an important potential application field of our work concerns the study of biomechanical stresses in the trunk. Models for their prediction require simulating the weight and shape dynamics of branches of order 2 (Figure 8.1). The set of data included in this target consists of:

- a complete description of the trunk, as in the Lollipop target: mass, length and diameter of internodes of the trunk growth units; mass of blades and flowers or fruits, if they exist, on the trunk growth units; internode numbers of each growth units and their number of lateral branches.

- for each branch of order 2 (born on the trunk): total mass of internodes, leaves, fruits and flowers. Additional data can be also included at this level such as the length of the branch main axis. It allows defining criteria for determining its physiological age and can help assessing the position of its gravity center.

8.1.2 Ring data

When available, measurement of ring diameters is particularly important as it gives access to information about the past growth of trees (Figure 8.2). This information is especially precious since it is the only source of information to get the growth temporal dynamics, as multi-date fitting is generally impossible for trees: it would require clones growing in exactly the same conditions.

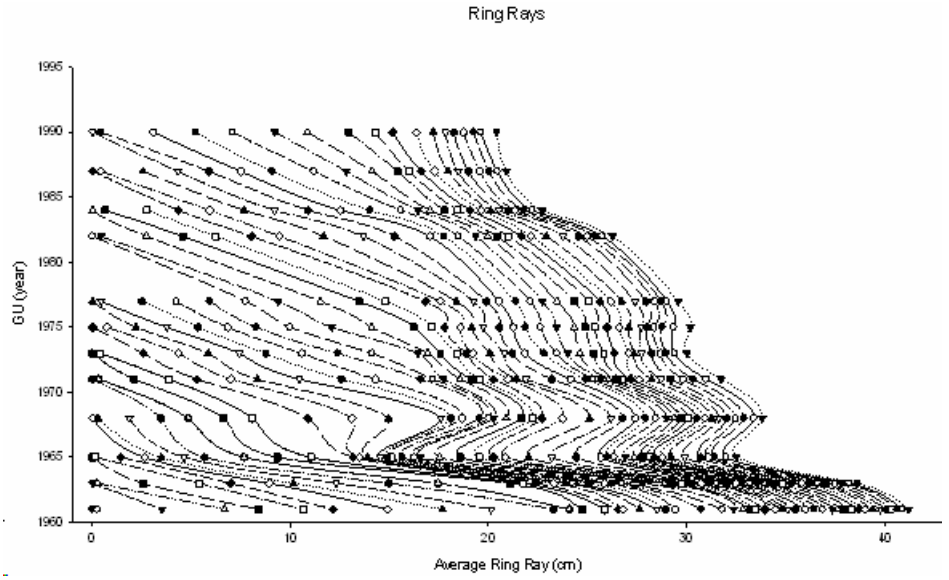


Figure 8.2: Target ring profile for Beech tree: longitudinal cross-section (data from joint work with Thiéry Constant, INRA Champenoux). The lines link the layers that appeared the same year.

The modelling approach presented in 3.2.3 provides an analytical relationship between ring widths of an axis and the amount of leaves that was present in the plant at different heights along this axis. If $\lambda = 1$ (case of Pressler law), the ring sectional area at GC t is:

$$s_p(n,t) = \frac{q_p(n,t)}{l_p(n,t)} = p_p^{rg} \cdot N_p^{a,a}(n,t) \cdot \frac{Q_{ring}(t)}{D_{Pressler}(t)} \tag{8.1}$$

As only the number of leaves depends on index n (chronological age of the underlying metamer), it implies that the ratio between ring sectional areas at different positions of a given axis gives access to the ratio of the numbers of leaves above ($N_{a,a}$) those positions:

$$\forall (n,n') \in \{1..t\}^2,\ \frac{s_p(n,t)}{s_p(n',t)} = \frac{N_p^{a,a}(n,t)}{N_p^{a,a}(n',t)} \tag{8.2}$$

This relationships involves two metamers of chronological ages n and n' that can be arbitrarily chosen as long as they have the same physiological age p. Thus, it is possible

to trace back the evolution of numbers of leaves from information on the corresponding ring sectional areas.
If $\lambda \neq 1$, the Pressler rule is not verified. However, a relationship between the ring sectional area of different metamers and the number of leaves above their positions can still be extracted. Given four different metamers of the same axis and characterized by their chronological ages n, n', $\widetilde{n}$, $\widetilde{n'}$, we have:

$$\frac{s_p(n,t) - s_p(n',t)}{s_p(\widetilde{n},t) - s_p(\widetilde{n}',t)} = \frac{N_p^{a,a}(n,t) - N_p^{a,a}(n',t)}{N_p^{a,a}(\widetilde{n},t) - N_p^{a,a}(\widetilde{n}',t)} \tag{8.3}$$

If we choose the metamers so that $n' = n-1$ and $\widetilde{n}' = \widetilde{n}' - 1$, we get:

$$\frac{s_p(n,t) - s_p(n-1,t)}{s_p(\widetilde{n},t) - s_p(\widetilde{n}-1,t)} = \frac{N_p^{a}(n-1,t)}{N_p^{a}(\widetilde{n}-1,t)} \tag{8.4}$$

where $N_p^a(n-1,t)$ is the number of leaves at GC t on the metamer of chronological age n and its lateral structures. It would correspond to the number of leaves located in the horizontal layer of the metamer (provided that branches are included in horizontal planes). Note we could write the above equations with leaf areas instead of leaf numbers, which would be more relevant if leaves are of different sizes.
Unfortunately, we did not have any data to test this last relationship. It would require the surfaces of blades of year t and rings of year $t+1$ that are fed by these leaves. These data are not accessible from destructive measurements but it would be possible to use sampling ratios or allometric relationships to estimate these values.

8.1.3 Simplified measurements of topology

Sample data (T0)

For adult trees, complete topological descriptions would be too tedious. Only sampling information is reasonable to expect. These sample data can consist for example in numbers of phytomers per growth units of different physiological and chronological classes or in numbers of lateral axes that they bear. Although the number of individuals per class are not large enough to provide significant results, figure 8.3 shows an example of such target data. It consists in numbers of phytomers per growth unit for a 21 year-old Beech tree (tree number one): this is a joint work with Thiéry Constant, LerFob, see section 8.4.1.

Numbers of organs for each physiological age-based category (T1)

This target is dedicated to shrubs. It is possible to count the organs but the topology is too complex to be accurately described. As organ counting is a non destructive measurement, it is possible to measure a large number of plants. For each measurement stages, the data consists of:

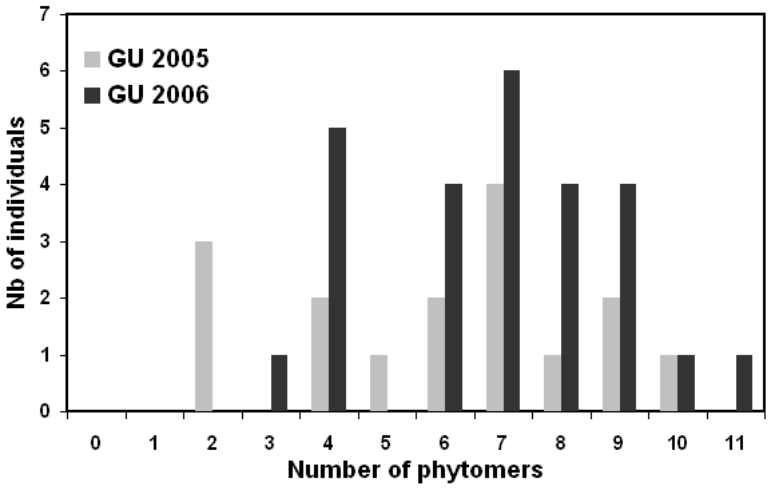


Figure 8.3: Example of target T0: sample data for the number of phytomers per growth unit in Beech tree (Thiéry Constant, LerFob).

- mean and variance of numbers of metamers for each physiological class
- mean and variance of numbers of fruits (if any) for each physiological class

Figure 8.4 shows an example of such target file.

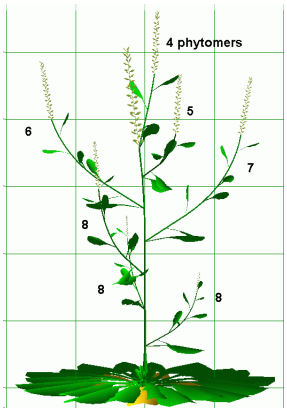
```
(Target_Type) 10   (Age) 27

Metamers Numbers:                Means                                  Variances
                  PA 1    PA 2    PA 3    PA 4    PA 5     PA 1    PA 2    PA 3    PA 4    PA 5
 (t) 6            5.44    0       0       0       0        0.61    0       0       0       0
 (t) 8            7.89    0       0       0       0        0.81    0       0       0       0
 (t) 17           9.56    6.89    3.02    9.50    0.78     1.44    3.28    8.73    25.32   1.61
 (t) 20           9.56    9.83    4.40    20.67   3.52     1.44    3.68    18.10   82.82   15.87
 (t) 23           9.56    11.94   5.54    34.41   7.26     1.44    2.80    28.07   153.51  64.27
 (t) 25           9.56    14.33   6.71    50.28   12.43    1.44    2.35    38.70   247.98  171.13
 (t) 27           9.56    16.17   7.71    67.28   18.77    1.44    2.38    50.81   209.98  357.65
```

Figure 8.4: Format of Target T1. For each growth stages (t), mean and variance of numbers of metamers and fruits are given. From data on 18 cotton plants collected by LI Dong (CAU, Beijing).

Mean (regular) automaton rules, measured or arbitrarily chosen (T2)

In many cases, pieces of information about the plant architecture are available prior to the fitting procedure. From direct observations or from botanical knowledge on some species, it is possible to determine the main rules of the topology of the target species. At this level, the data are not included in target file but input directly in the simulation. We illustrate this kind of description with two very different plants: *Arabidopsis thaliana* L. and *Fagus Sylvatica* L.

***Arabidopsis thaliana* L.** For instance, from the numbers of phytomers on lateral axes of *Arabidopsis* inflorescence, we implemented a regular topology consistent with our observations: four phytomers on the top lateral axis, one more phytomer on the axis below and so on. Regarding lateral axes emerged from the basal rosette, the number of phytomers was more difficult to count precisely and was bounded to eight in the simulation. Because of expansion delays, not all the phytomers could elongate. Only leaves belonging to these basal lateral axes were visible. Although not exact, these topological rules provide reasonable numbers of phytomers and adequate development sequences that allow reproducing the measured data on lateral compartments (if for instance target B2 is associated).



Axis	Numbers of phytomers		
	av. observed nb	std dev	simulated nb
Axis 1 (top)	4.1	0.3	4
Axis 2	4.8	1.0	5
Axis 3	6.2	0.4	6
Axis 4	6.5	1.0	7
Axis 5	6.4	1.0	8

Figure 8.5: Topology of *Arabidopsis thaliana* (joint work with Angélique Christophe, LEPSE). The average number and standard deviation (std. dev.) of phytomers are calculated from data collected on 10 plants (including six plants grown under moderated hydric stress).

***Fagus Sylvatica* L.** For beech trees, the topology of the trunk (PA 1) was measured and set as input of the simulation. The growth unit of physiological age 1 presented in Figure 8.6 is a virtual growth unit exhibiting the maximal numbers of phytomers per zone that were observed on the two beech trees. For branches, the choice was based on our observations (as for instance Fig. 8.3) and on the botanical studies reported about that species.

Beech tree follows the rules of a Troll architectural model (Hallé and Oldeman [1970]). In our observations, polycylic shoots (definition in Table 2.1) were scarce (or sometimes difficult to identify) so that phenomenon was not considered, which is reasonable since the trees were grown in understorey.

Beech development rely on two kinds of shoots: short shoots composed of a few metamers with negligible internodes and long shoots with elongated internodes. Long shoots are dedicated to space exploration whereas short shoots are high contributors

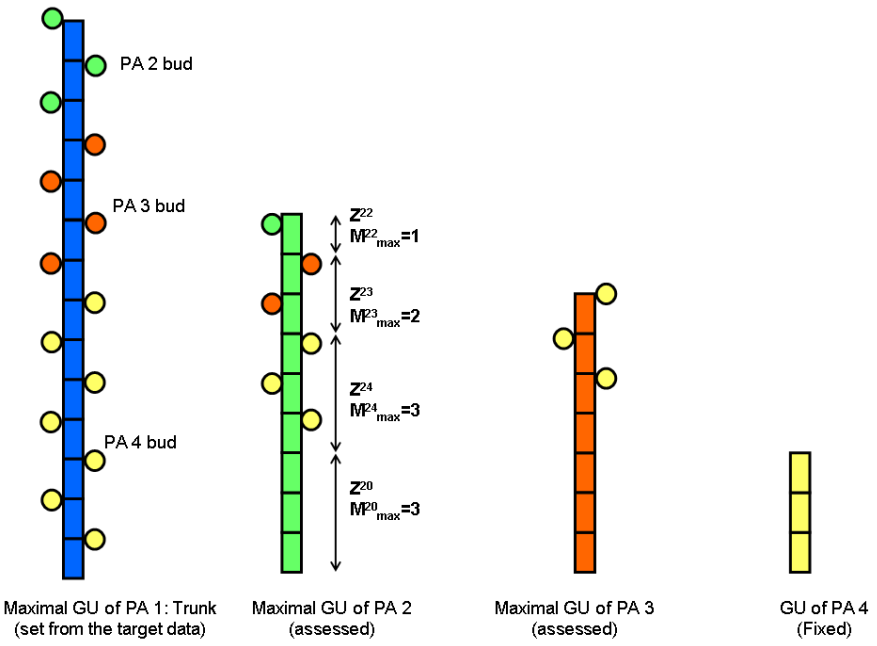


Figure 8.6: T2 level for description of topology.

to the plant production (Thiebaut and Puech [1984]). Besides the trunk, we sorted the shoots into three PA-based categories.

Following Thiebaut and Puech [1984], we distinguished these three classes according to their elongation and ramification characteristics: short shoots bearing no branches (PA 4), long shoots bearing only short shoots (PA 3) and long ramified shoots (PA 2). The default topology for those three kinds of growth units are defined in Figure 8.6. Growth units of PA 2 consist of four zones made of metamers that can bear respectively no branches, axes of PA 4, PA 3 and even PA 2 (partial reiterations). Branches of PA 2 and PA 3 are both composed of long shoots with similar physiological characteristics but differ on their ramification parameters.

For short shoots of PA 4, the number of metamers is relatively stable [Thiebaut and Puech, 1984]: from our observations, it can be 3 or 4. The value of 3 was chosen and kept constant. The sequence of zones along growth units follows the botanical principle of acrotony, as described by Rauh for beech trees (Nicolini [1998]). Note that figure 8.6 only gives the default topology. The *GL3* version was used which implies that at each growth cycle, the parameters driving the development are updated according to the Q/D ratio (see section 8.4.1).

8.2 Fitting the parameters of plant development and growth from simplified target: case of deterministic development ($GL1$)

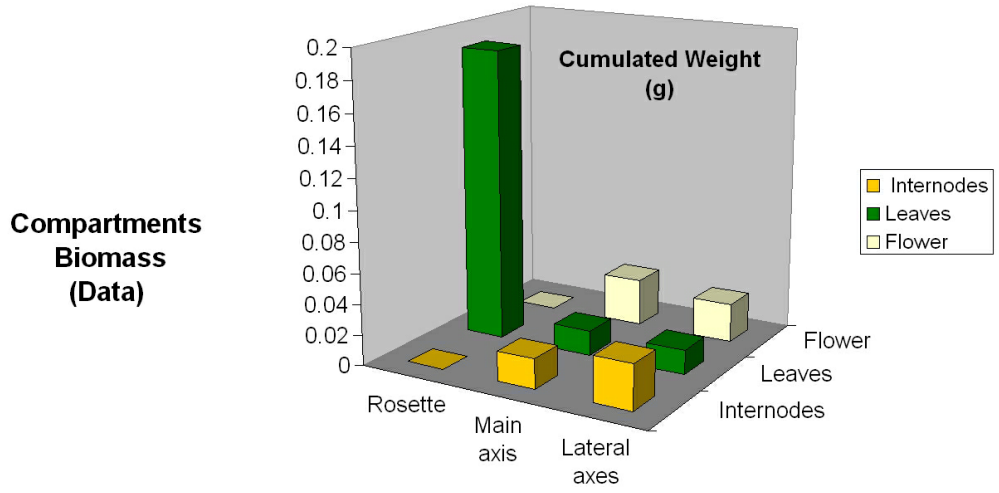
Depending on the level of description available for topology (*i.e.* target T0 to T2), simulations account for the plant topology in different ways. The objective is to find the rules of an automaton that would be consistent with both the biomass and the topology measurements. These topological rules can be deterministic or stochastic depending on the GL version considered. In the following sections, we present the fitting procedure and results for different plant species using the target types defined above. When the $GL1$ version is considered, the development is considered as deterministic so we can generally input the plant topology directly from the observations. It has been done so for *Arabidopsis thaliana* and *Pinus tabulaeformis*.

8.2.1 *A priori* choice of a mean topological automaton (Arabidopsis)

Arabidopsis has been defined as a model plant to study the molecular genetics of plant development and to build gene-based models of growth processes [White and Hoogenboom, 2003]. It has a rapid generation time (six weeks) and a small stature (around 20 cm high), allowing making it easily grow in controlled conditions. But few models exist that consider *Arabidopsis* at plant scale and all of them only simulate its geometrical and topological development [Mündermann et al., 2005]. A model such as GreenLab, linking architectural and functional development of the plant, allows testing the plasticity of the plant response to environmental factors such as hydric stress and allows studying the performances of different genotypes depending on the environmental conditions[Christophe et al., 2008]. A preliminary step is parameter identification of the model and simplification of the experimental protocol so that many plants can be considered in a further step. *Arabidopsis thaliana*(L.) Heynh. plants were grown in a growth chamber under non-limiting conditions. At five growth stages (characterized according to Boyes et al. [2001]), four plants were sampled and measured following the "rattle" description (B2): we measured dry weights of organs for the main stem and dry weights of compartments for the lateral axes of the inflorescence (last growth stage only). Root dry weight, individual blade area and projected rosette area were also measured (see section 3.1.2 about RUE variations). Phytomers were counted to assess the phyllochron. A change in phyllochron was found during the vegetative phase (presumably at inflorescence induction) and was modelled using rhythm ratio as presented in section 2.3. The topological structure was recorded and an average number of phytomers was chosen on each lateral axis for the simulation, as described in section 8.1.3. The PA-based categories were chosen according to Doerner [2001]: the basal

rosette was set PA 1 and the inflorescence was set PA 2 [Christophe et al., 2008].

Physiological age of inflorescence axes. In a prior version, the inflorescence lateral axes were attributed different PAs than that of the main axis. Indeed, biomass partitioning and organ mass of lateral axes were apparently different from those of the main axis. Figure 8.7 shows the comparison between compartment weights on the main axis and the cumulated data on the lateral axes.



	Rosette	Main axis	Lateral axes
Internodes	0	0.01957	0.02934
Leaves	0.189079993	0.01611	0.0158
Flower	0	0.03054	0.02522

Figure 8.7: *Arabidopsis thaliana* L.: Biomass partitioning in rosette and inflorescence at stage 6.00.

However, the fitting results revealed that the same sink strength values could generate these observations. The differences in dry weights are simply induced by expansion delays (top-down flowering, see section 3.2.1). Thus the fitting results presented in figure 8.8 were obtained with a single set of parameter values for organs on the main stem and for organs on lateral axes (*i.e.* lateral axes can be considered as mere reiterations of the main stem).

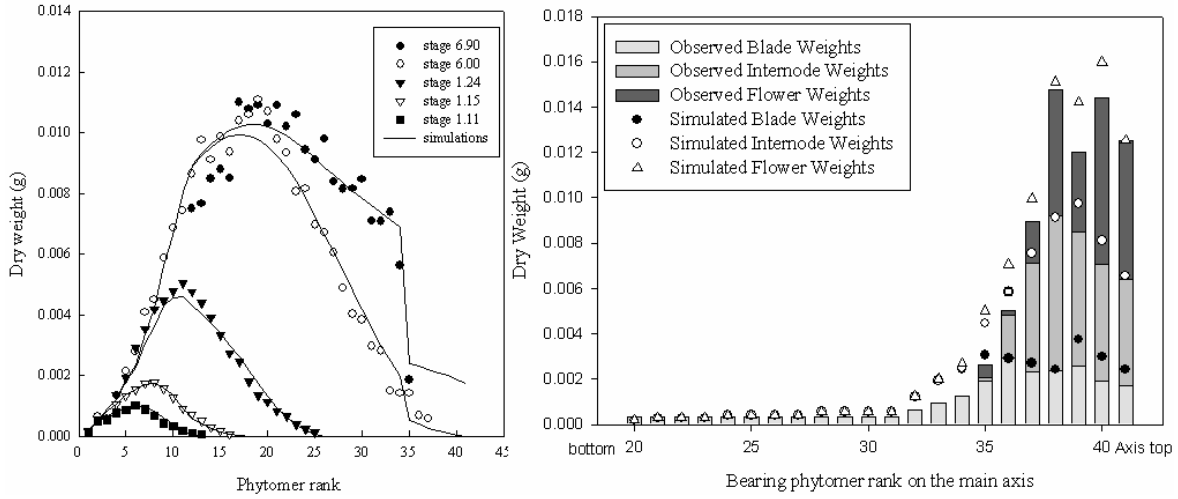


Figure 8.8: *Arabidopsis thaliana* L.: fitting of leaf dry weights on the rosette and of biomass compartments on lateral axes (joint work with Angélique Christophe, LEPSE).

8.2.2 Creation of a regular target consistent with the structure factorization (adult Chinese pine)

From measured data to target data

To test the effects of data aggregation on fitting results, data were collected on an adult pine tree by Guo Hong and Hong LingXia (Chinese Academy of Forestry, Beijing). From growth units counting, the tree was estimated to be 41-year-old. Due to the large number of data to measure, the following experimental procedure was adopted: at each node, one average branch was selected and measured in details. For the rest of the whorl branches, needles were separated from woody parts and each compartment was weighted. The measured data can be represented as in figure 8.9 where the semi-ellipses represent compartment mass (wood or needle) and the coloured lines represent detailed measures (mass and dimensions of each growth unit). From this semi-complete set of measurement, we built two kinds of target data.

In the first one, we created a virtual complete tree with a full branched structure. That branched structure was generated by sticking at each node several repetitions of the selected branch. The number of repetitions in the whorl was chosen so that the global mass of needle and internode compartments were conserved (see "average" tree in figure 8.9). Then an average topology (axes numbers per whorl, GU numbers per axis) was defined so that the tree structure would be consistent with the structure factorization of the tree. It implies that two structures that appeared at the same time with the same physiological age are supposed to be identical. We sorted the branches according to their physiological and chronological ages and we created an average representative branch

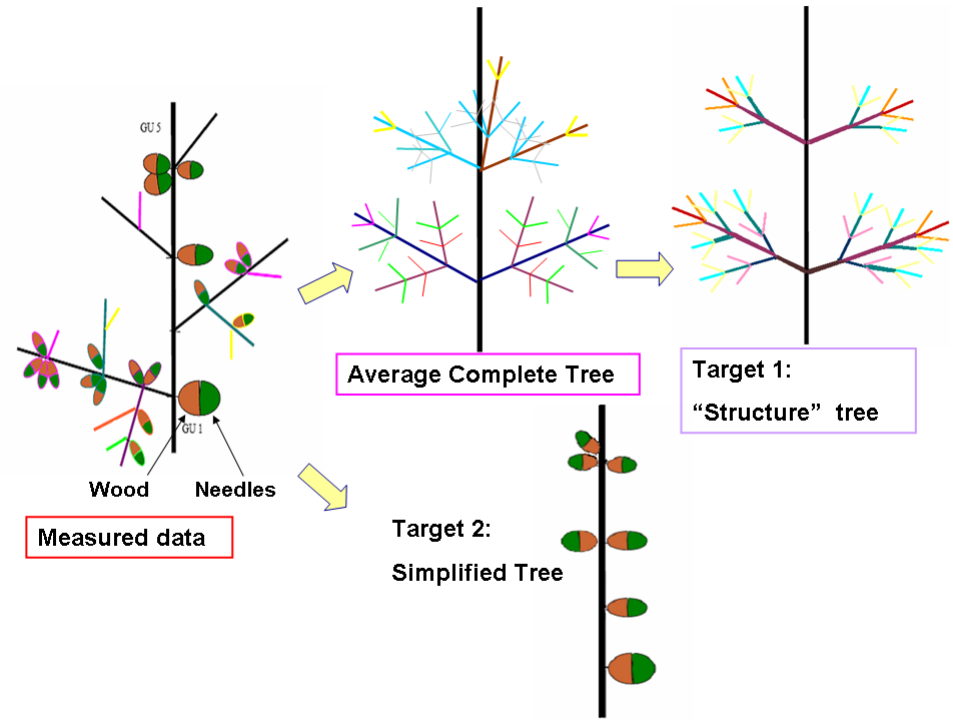


Figure 8.9: Chinese pine (*Pinus tabulaeformis* Carr.): from measurement data to target data.

for each class. Concerning the morphological data, metamer fresh weights (needles and internodes) were averaged for each axis class; length and diameter were kept from the measured axis that was the closest from the average one in terms of biomass. At each node, the average number of axes was calculated with the same principle as described above (see "structure" tree in figure 8.9). At each step, we checked that the total compartment weights were conserved (cumulated data shown in graph of figure 8.11). The second kind of target was formatted as described in section 8.1.1: Rattle target for biomass (B2). The topology was input in the simulation directly from the observed one (level T2 for topology, section 8.1.3). The set of target data can be represented as shown in figure 8.9 ("simplified" tree).

Modelling assumptions and fitting results

The modelling assumptions were set identical to those of young pine trees (section 7.3.1): physiological age varies with branching order; growth units consist of one single virtual metamer; organ growth lasts one cycle; sink ratios are constant (although with different values than for young pines, see table 7.1) The assumption of constant sink ratios for organs of each physiological age (branching order) was tested and validated

on the data, except for organs of physiological age 1, the sink of which seems to vary during the tree growth.

Param.	Meaning	Values				
		PA 1	PA 2	PA 3	PA 4	PA 5
Be	Internode allometry	-	27	28	4.6	19.8
β	Internode allometry	-	-0.02	0.7	0.5	0.4
P^i	Internode sink	0.0097	0.043	0.013	0.024	0.011
P^a	Blade sink	1	0.48	0.13	0.3	0.15

Table 8.1: Sinks and allometry coefficients for the adult Chinese pine.

Note that there is a relatively constant proportion between internode and needle sinks. Whereas it varies from 0.8 (PA 1) to 2.9 (PA 4) with branching order for young pines (*i.e.* a relative difference of more than 100%), the needle to internode weight ratio only varies from 8.7 to 11.7 for the adult pine (*i.e.* the relative difference is around 30%). It means that the internode to needle weight ratio becomes more homogeneous for the different branching orders when the tree grows older.

Param.	Meaning	Values	
		Complete target	Rattle target
$R_1 = 1/\mu$	Needle resistance	7.4	7.6
P_1^{rg}	Slope for sink variation	0.037	0.038
λ	Leaf influence	0.57	0.59
	on ring biomass partitioning		

Table 8.2: Comparison of fitting results for complete and "rattle" target format.

The fitting parameters are presented in table 8.2 and figures 8.10 and 8.11 shows the comparison between observed and simulated data for data collected at G.U. level or at compartment level. Close parameter values are found with both target types ("structure" or "simplified" at rattle level). This is reasonable since the topology was input from the observations and thus kept the same in both cases. Surprisingly enough, the model structure for ring global allocation that allowed getting the smallest final error was the third method (mode Q) which was thus selected for these simulations. The value found for parameter λ was close to 0.6. It means that the amount of leaf biomass located above the metamer influences for around 60% of the biomass partitioning for cambial growth of metamers.
The illustrations of 8.12 show some 3D simulation outputs. It shows that it is possible to mimic relatively well the growth of real trees with a simulated tree having a regular topology. The simulated tree and the real one have similar mass, length and diameter for internodes of the main stem and similar mass for needle and wood compartments on order 2 branches. The fitting accuracy was reduced for the detailed measurements

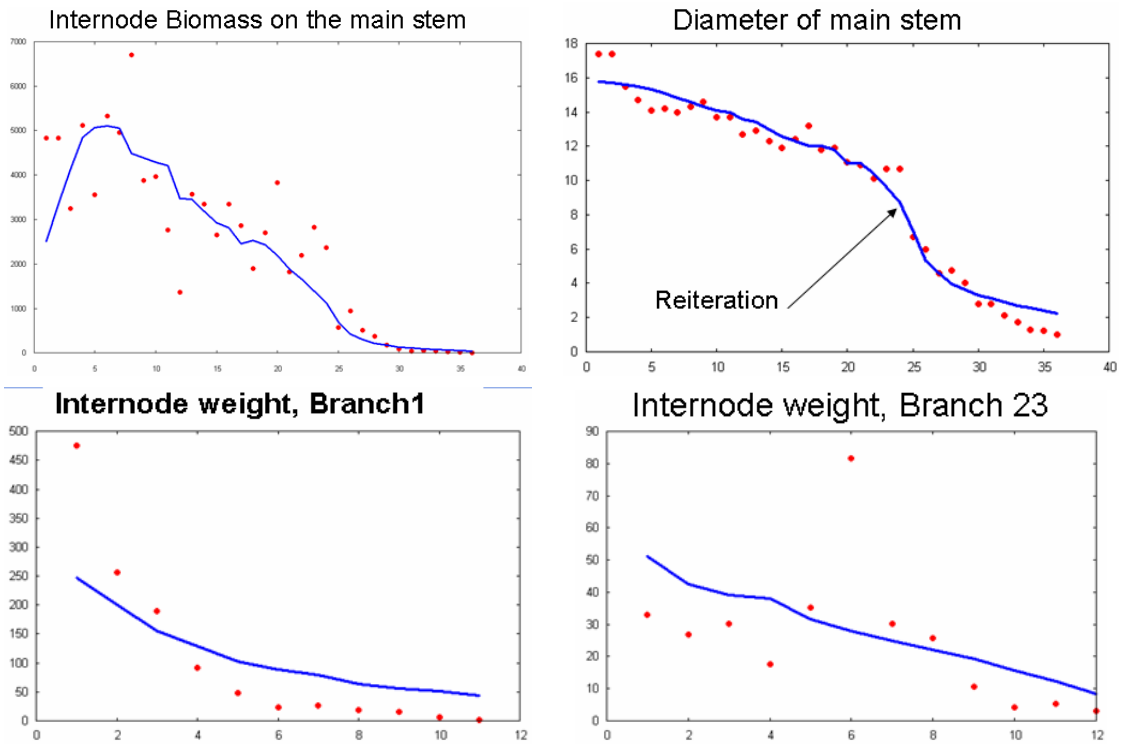


Figure 8.10: Chinese pine (*Pinus tabulaeformis* Carr.): fitting results at G.U. scale: data (dots) vs simulated (lines) G.U. weights and diameters of the trunk and G.U. weights of two branches (branch 1 and branch 23)

on higher order branches (branch 1 and branch 23 presented in figure 8.10 were among the best ones) but their global masses were conserved.
We can observe strong differences when comparing these results with the values found for young pines (section 7.3.1). The most important one is that the best model structure found for global allocation to secondary growth is different: mode Q/D for young pines and mode Q for the adult pine. These differences can be due to several effects that are difficult to unravel:

- ontogenetic changes: some parameters such as sink coefficients may vary with the tree development
- environmental influences: the young and adult pines were grown in different locations and hence did not undergo the same climatic conditions. Moreover, the adult tree was probably subjected to high competition from its neighbours.
- for young pines: the effect of transplanting from the nursery to the open field certainly had an impact on the growth, that we did not take into account in the simulation

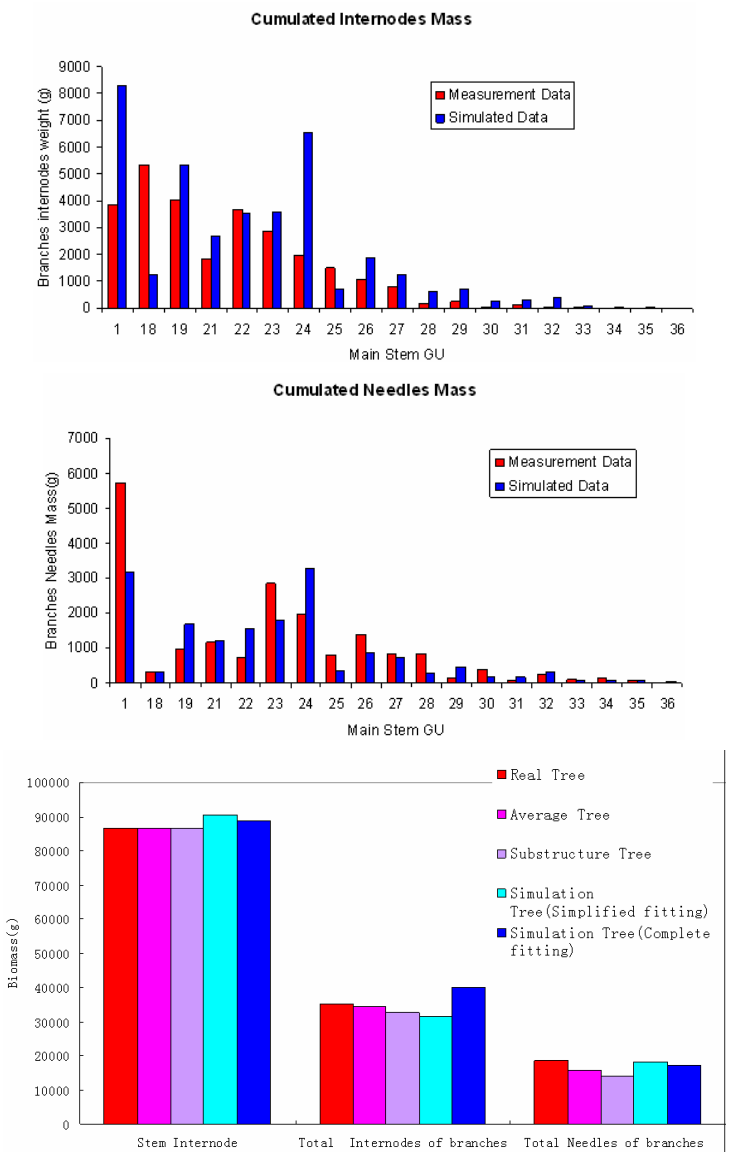


Figure 8.11: Chinese pine (*Pinus tabulaeformis* Carr.): fitting results at compartment scale for internode weight and needle weight on branches. The last graph represents the respective mass of (i) wood of the trunk, (ii) cumulated wood of the crown and (iii) cumulated needles of the crown. The red bars are the measured data, the pink bars are the "average" tree data, the purple bars are the "structure" tree data, the blue bars are data of the simulated tree after fitting on simplified target and the dark blue bars are data of simulated tree after fitting on "structure" target (see Figure 8.9.

- for the adult pine: determination of the tree age was difficult and the result was uncertain. Above the position of the reiteration, counting the growth units on

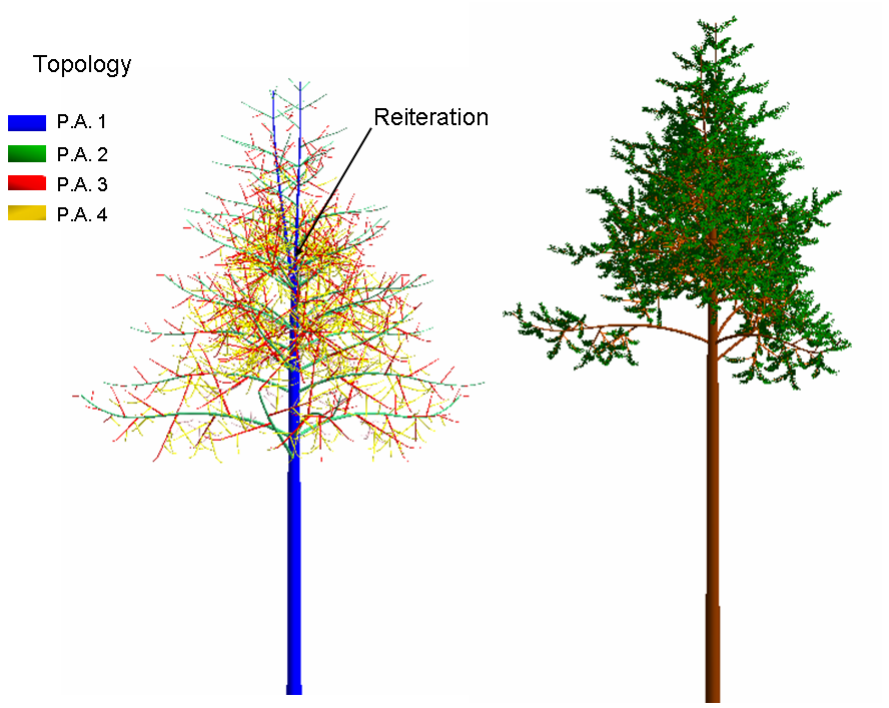


Figure 8.12: Simulation output of pine tree.

different branches provided different values. It seems that the some branches developed while others observed pauses during their growth, the causes of these pauses being unknown.

- introduction of errors with several potential sources:
  - measurement errors, especially for the adult pine (more than 3000 data measured: some errors were detected during the analysis but other may remain)
  - errors introduced when building the average target and the "structure" targets
  - residual errors from the fitting procedure
- modelling choices: regarding resistances, the value is much higher for the adult tree. It would be interesting to include hydraulic resistivity of branches: as the tree is old, it may have an influence on biomass production. Regarding allocation to ring compartment, we chose the method that reduced the most the final weight mean square error which may not be the most physiologically relevant strategy.

However, a positive conclusion is that the same parameter values can be found when a "rattle" or a complete target format are used. Note that here the topology of the

simulated tree was set from observations and therefore was the same in both cases. Further tests would be needed to confirm it and to analyze more precisely the effect of data aggregation.

8.3 Fitting the topological parameters of stochastic development ($GL2$)

This section is oriented to applications on plants for which it is possible to count the organs but maybe too fastidious to describe their topological connexions (typically for shrubs). We assume in this section that there is at most one phytomer per growth unit and that the probability of apical death is zero.

8.3.1 Analysis of axis development

Kang et al. [2007a] take advantage of some useful results from the theory of differential statistics. If X, Y and Z are stochastic variables so that $Z = f(X, Y)$ then a first-order approximation of the expectation of Z is:

$$M_Z = f(M_X, M_Y)$$

and similarly, an approximation of the variance is:

$$V_Z = f'_X(M_X, M_Y)^2 V_X + f'_Y(M_X, M_Y)^2 V_Y + 2 f'_X(M_X, M_Y) f'_Y(M_X, M_Y) Cov_{X,Y}$$

These formula are easily extended to the case of more than two stochastic variables. In $GL2$, the production at growth cycle t depends on the demand and production at previous growth cycles, $Q(t')$ and $D(t')$ with $t - t_a < t' < t$ where t_a is the life span of leaves (in GC). It implies that *computing the plant production from the theoretical demand expectation (Equation* (2.9)*) provides an approximation of the expectation of the production*. The same arguments prove that computing individual organ weights from theoretical demand expectation at each growth cycle provides correct approximations of organ weight expectations. Thus we can build a theoretical plant which is an approximation of the 'average plant' of a population. Note that this theoretical plant has non-integer numbers of phytomers in each PA-based category. If such a plant needs to be represented, all potential phytomers are represented but their sink strengths are weighted by their probabilities of presence.

Linking phytomer rank to its chronological age: effect of apical growth probability

Let us consider a single-stem plant (or a branch) growing with a probability $a < 1$ of phytomer appearance at each growth cycle. Using the notations of 2.5.2, we have

in fact $P_A(p) = a$ where p is the branch physiological age. The number of phytomers follows a Bernouilli process: it is impossible to determine the exact age of the observed phytomers (Figure 8.13).

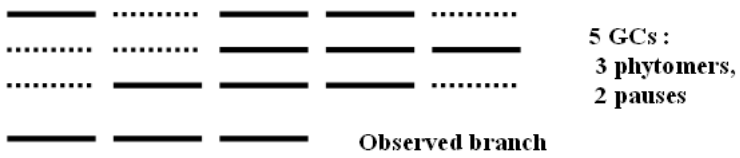


Figure 8.13: Several sequences of activity/pause of the apical bud can produce the same branch.

However we can get a relationship between the observed rank of a metamer and its probability to be of a certain age. Let X_i be the variable representing the number of growth cycles needed to see the i^{th} metamer of the axis. Its probability distribution follows the Pascal law:

$$P(X_i = k) = C_{k-1}^{i-1} a^i (1-a)^{k-i} \tag{8.5}$$

and thus the expectation of the chronological age of a metamer of rank i in the plant at growth cycle t is $t - \frac{i}{a} + 1$. If Y_i is the variable representing the weight of an organ of rank i, the expectation of Y_i can be calculated as a function of the weights of organs appeared at growth cycles k with $i \leq k \leq t$:

$$\begin{aligned} E(Y_i) &= \sum_{k=i}^{t} P(X_i = k) \cdot W_k(t) \\ &= \sum_{k=i}^{t} C_{k-1}^{i-1} a^i (1-a)^{k-i} \cdot W_k(t) \end{aligned} \tag{8.6}$$

This expression allows analyzing the metamers' weights according to their respective ranks (which we can measure, contrary to their exact chronological ages). But it requires the preliminary determination of the probability value a.
The probability a possibly interacts with the branch rhythm ratio, w (section 2.3). In that case, the sequence of growth cycles of pause and activity of the apical bud is periodic. The rhythm ratio determines the potential positions where phytomers can appear. The effective appearance is driven by the apical growth probability. Thus it is possible to generate a deterministic branch with the same average number of phytomers by setting a rhythm ratio $w' = a \cdot w$ (but of course the variance is null).

From these considerations we draw the following conclusions:

- From the topological parameters, namely development probabilities P_A, P_B, etc and rhythm ratio w, of the stochastic model GL2, it is possible to simulate a

deterministic plant that approximates the development and growth of the average plant of a population whose individuals share these same parameter values. The development is calculated from the expressions of mean numbers of organs for each physiological age (see 2.5.2):

$$D_{th}(t) = \sum_{o\ organ} \sum_{p=1}^{P_m} \sum_{n=1}^{t} \underbrace{(M_{S_1^p(n,n)} - M_{S_1^p(n-1,n-1)})}_{\substack{\text{Nb of PA } p \text{ organs} \\ \text{appeared at GC } n}} \cdot P_p^o(t-n+1) \tag{8.7}$$

where $P_p^o(t-n+1)$ is the sink value of organ o (o representing respectively all kinds of existing organs: leaf, internode, fruit...) of physiological age p and chronological age $t-n+1$. $M_{S_1^p(n,n)}$ is the number of metamers of PA p in the structure $S_1(n,n)$, i.e. the whole plant at growth cycle n. The number of organs appearing at cycle n is simply the difference of the numbers of organs in the plant between cycle n and cycle $n-1$. Note that this formula is valid only if the number of organs is equal to the number of metamers. The opposite case can be generated by two kinds of factors:

- each metamer bears several organs of the same kind (e.g. two leaves)
- flower abortion is possible. In that case, the probability of fruit appearance needs to be taken into account.

In these two cases, it is sufficient to multiply the expression by the adequate constant factor.

- We can take advantage of these properties to fit the development parameters of the model. From mean and variance of numbers of organs in a population, we can find the parameters generating the corresponding theoretical mean and variance of the virtual population. The procedure is as follows: from target data consisting of mean and variance of numbers of organs, it is possible to use the theoretical expressions for means and variances to fit the parameters of development, namely apical growth probability, branching probability, apical death probability, fruit abortion probability. These parameter values being set, we can calculate the theoretical demand expectation. It provides expectation of biomass production, the parameter of which (efficiency, reference surface) can be fitted using a second set of data consisting of compartment mass at different growth cycles. Parameters driving allocation to individual organs (sinks and sink variation functions) are fitted using a third data set consisting of organ weights for branch tips only. As organs are sorted according to their rank, equation 8.6 is used to assess their average chronological age (which is important for determining their individual demand and hence their sink variation). This method was applied to tillering wheat as described in the following paragraphs.

8.3.2 Tillering wheat

This section presents the results of the first application of the method described above. It has been applied on wheat data from Wageningen University (joint work with KANG MengZhen) for plants grown at a density of 100 plants per m^2. Although wheat topology is relatively simple, analysis of repetition measurements have revealed some variability in the development and thus the stochastic version of GreenLab has been chosen. The

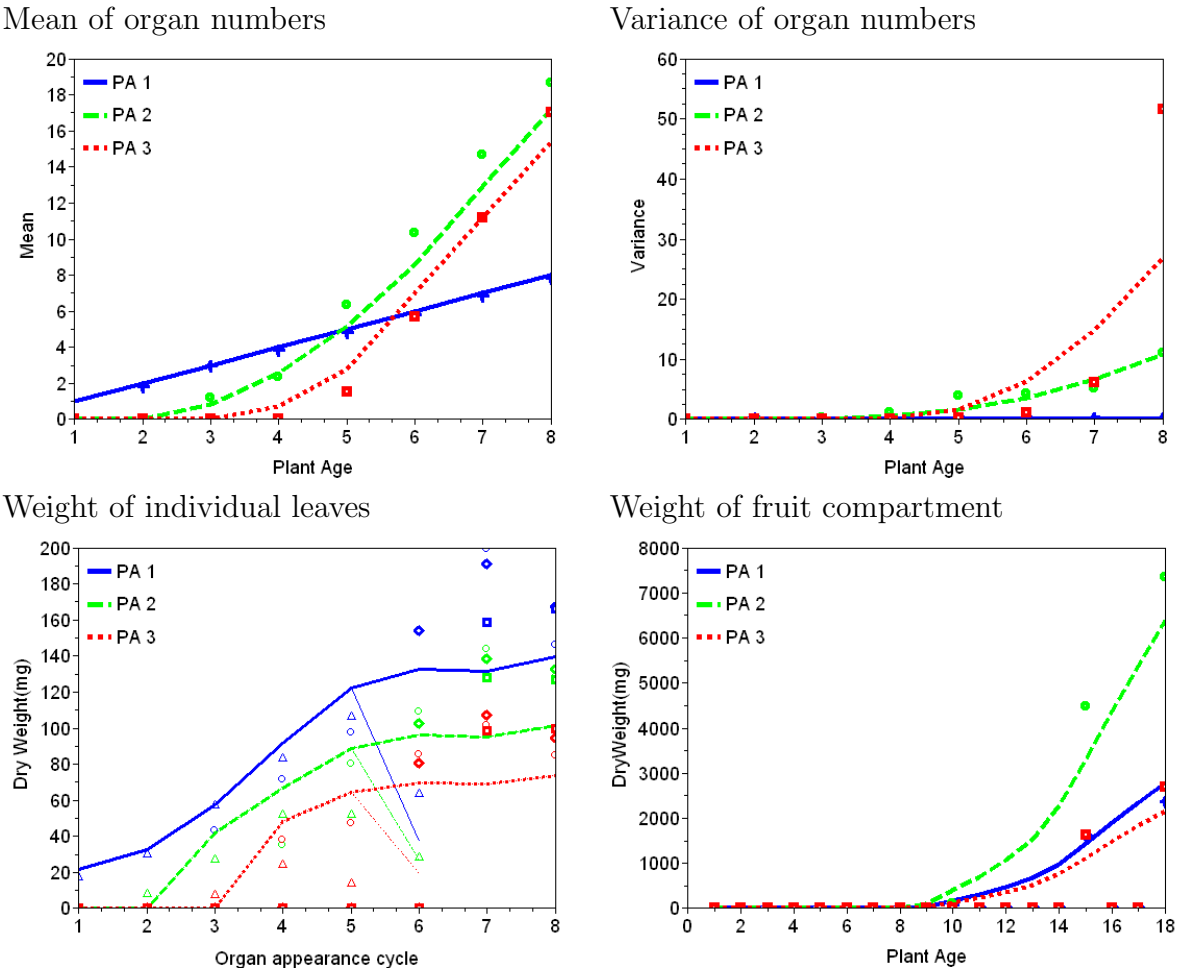


Figure 8.14: Some fitting results for stochastic modelling of wheat with tillers. Dots represent measured data and lines are simulated values. Data are from Wageningen University (Kang MengZhen).

wheat topology was set as follows: main stem is of physiological age 1, first-order tillers are of PA 2 and second order tillers are of PA 3. For each measurement growth stage, the target file consists of:

- a cumulated target B0 (weights of biomass compartments)

- a "sparse" target B3 *i.e.* description at phytomer level but with data of axis tips only
- a target T1 (mean and variance of numbers of organs for each PA category)

It was assumed that apical death probability was zero and that growth probability was equal to one (as all fruits appeared at the same growth cycle and with the same rank (depending on physiological age). Thus only branching probability (for tiller appearance) and fruit appearance probability remained to be assessed. If the fruit does not appear, the tiller dies. Figure 8.14 presents some fitting graphs for development (mean and variance of numbers of organs) and growth (mass of organs and compartments). A study on the influence of density on the development parameters is currently in progress.

8.4 Deterministic development driven by the trophic state ($GL3$)

8.4.1 Beech trees

Modelling choices

Beech trees are known for their architectural plasticity in response to ontogenetic and environmental changes [Nicolini, 2000]. Owing to the feedback control of the level of trophic competition on topological and functional variables of the model ($GL3$ version), it is possible to simulate these responses. The simulated trees reproduced qualitatively the behaviour of observed trees such as the progressive appearance of branches of higher vigour (called base effect) that characterizes the young growth phase of a tree [Mathieu, 2006]. However those parameters remained to be quantified.
We selected two of the main $GL3$ variables in this study [Letort et al., 2008a]: number of new axes of each physiological age initiated and number of metamers per growth unit zone. At growth cycle n, the total number of new axes $a_{pq}(n)$ appearing on zones Z_{pq} in the tree (i.e. the number of branches of PA q borne by metamers of PA p) depends on the number of existing positions potentially bearing that kind of axes $N_{pq}(n)$ and on the value of Q/D. It is given by the following equation:

$$a_{pq}(n) = \left[N_{pq}(n) \cdot \left(A_{pq,1} + A_{pq,2} \cdot \frac{Q}{D}(n) \right) \right] \tag{8.8}$$

where $[x]$ represents the rounded value of x. The $a_{pq}(n)$ axes are distributed on the $N_{pq}(n)$ positions under the following constraints:

- growth units that appeared at the same growth cycle, with same physiological age and same rank on their mother axis bear the same number of axes

- axes distribution begins from the oldest branches to the youngest ones (less ramified)
- the distribution is as uniform as possible (all existing positions receive one axis before one position receives one more axis).

Similarly, the number of metamers $u_{pq}(n)$ per growth unit zone is determined according to the default topological rules and modulated in function of the value of Q/D:

$$u_{pq}(n) = \left\lfloor U_{pq,1} + U_{pq,2} \cdot \frac{Q}{D}(n) \right\rfloor \tag{8.9}$$

Note that the number of metamers per zone is calculated at growth unit level whereas the number of axes is determined at whole-plant level.

Measurement procedure and data

Two common beech trees (*Fagus sylvatica* L.) grown in understorey were measured in May 2006 from the natural stand of Champenoux, near Nancy (North-Eastern part of France). The tree ages were estimated from data on growth unit numbers and ring counting. One tree ("tree 1") was 21 year-old and the second ("tree 2") was 46 year-old.

Measurement of simulation input data The default topology was set as described in paragraph 8.1.3. The position and number of metamers of growth units on the main stem were recorded while the separation markers were still visible (not possible at the stem basis). Branches numbers and positions were noted, including scars indicating the positions of dead branches. To allow comparing the biomass compartment data, the branch numbers and positions on the trunk were input directly from the measurement. It means that the variations driven by the topological parameters (namely $A_{pq,1}$, $A_{pq,2}$, $U_{pq,1}$ and $U_{pq,2}$) were considered only for physiological ages higher than 1 (p=2, 3). At random locations, blade area was measured on samplings in order to have an estimation of the specific leaf weight (SLW). It ranges from 0.0072 g.cm^{-2} for tree 1 to 0.0093 g.cm^{-2} for tree 2. In GreenLab, the variations of SLW according to leaf position in the crown were not considered, although several studies have revealed that a gradient can be observed in relation to light exposure [Beaudet and Messier, 1998]. However, no visible gradient was observed from our sampling data. As our trees were grown in the understorey and were still covered by higher neighbor tree shading when they were measured, the light they received was mainly diffuse light. So the light environment was likely to be relatively isotropic and similar for both trees. Nicolini and Chanson [1999] report that SLW also varies with the tree age because of ontogenic modifications of cell structures. Consequently, the same SLW evolution was set in the simulation for the two trees according to their age. From sampling data collected, allometric ratios for biomass partitioning inside the new growth units were found in coherence with the

literature [Comps et al., 1994]. The following values were taken: the average long shoot to short shoot weight ratio is 5.25; the ratio of internode weight over leaf weight is 0.065 in a new short shoot and 0.7 in a new long shoot.

Measurement of target data The target data chosen as criteria to evaluate the accuracy of the simulation were the complete profile of the trunk and compartment biomass on each branch of order two. For each growth unit of the main stem, fresh mass, mean diameter and length were measured. At regular intervals, annual ring widths were recorded along four directions. For each branch of order 2, its total fresh weigh, its main axis length and basal diameter were measured. For PA 4 and PA 3 branches, leaves were separated from internodes and weighted. Concerning PA 2 branches, the number of leaves was too high to get complete measurements. The following procedure was adopted to get an estimation of the wood and leaf weights: the basal part of the branch, bearing no leaves, was weighted separately and recorded as a woody part. For the rest of the branch, a representative branchlet was chosen. Its leaf and wood weights were measured. The same wood-leaf ratio was kept to get the compartment weights from the total weight of that part of the branch. If several branches of same PA are located on the same GU, the average values for the leaf and internode compartments are put in the target. Of course, the cumulated value remains the same when summed on all branches. To summarize, the following data were included in the target file: average internode length, diameter and fresh mass for each growth unit on the trunk, average ring diameters for some of those growth units, weights of leaf and wood compartments for each branch of order 2.

Fitting procedure These data were fitted using the generalized non linear least square method for the functional parameters and the simulated annealing algorithm for topological parameters (see section 7.2). The particle swarm optimization was also tested and gives similar results. As no quantitative topological information was recorded concerning the branching structure of order 2 branches, the topological parameters were fitted on aggregated data of biomass compartments only. The model parameters were fitted to create an average structure with a global demand at each growth cycle similar to the real one. The average simulated weights of blade and internode were compared to the measured data to ensure that the total number of metamers is globally coherent with that of the target. Data of the two trees were fitted simultaneously with the same set of parameter values, to test if the model was generic and robust enough to represent the growth of two different trees. Only a single parameter representing the environmental conditions (chosen as a constant) was set different between the two trees.

Topological parameters					
Bearing PA	Axill. PA	Nb of metamers Param		Nb of axes Param	
(p)	(q)	$U_{pq,1}$	$U_{pq,2}$	$A_{pq,1}$	$A_{pq,2}$
2	0	1*	0.4	-	-
	4	1*	4.35	0*	2.05
	3	1*	0.1	0*	0.1
	2	0*	0.2	0*	0.09
3	0	1*	1	-	-
	4	1*	2.8	0*	0.6
Physiological parameters					
$R_1 = 1/\mu$	Resistance to biomass conversion (3.5)			85.5	
Sp_0	Reference surface (3.6)			0.0046	
α	Variations of the reference surface area (3.6)			0.75	
P_1^{rg}	Sink for ring compartment (3.24)			2.3	
γ	Exponent for ring demand (3.24)			2.95	
p^{rg}	Linear sink density for rings (3.27)			PA2:0.1; PA3: 0.05; PA4:0.01	
λ	Blade influence on ring partitioning (3.27)			0.13	
$E(1)$	Environment value for tree 1 (3.5)			5.6	

Table 8.3: Parameter values after simultaneous fitting on data measured on tree 1 and tree 2. The topological parameters are estimated for each zone Z_{pq} (PA p with axillary buds of PA q). * denotes a fixed parameter (value defined a priori). Into brackets are the references of equations where the parameters intervene.

Fitting results

The two trees were fitted together with a single set of parameters (Table 8.3). However, their environment may be different and thus the variable representing the environmental conditions was fitted for tree 1. As it is a relative value, tree 2 was given the reference value 10. A smaller value of environment was found for tree 1 (5.6) which is coherent with the difference measured between SLW of the two trees. As a higher SLW is generally correlated to favorable light conditions, tree 2 that has a higher SLW, has probably grown in better conditions.
Fitting results are presented in graphs 8.15 and 8.16. Although data from tree 1 and tree 2 are of different ranges, they could be fitted in parallel with satisfactory success. The residual differences between simulated and measured data are mainly due to our modelling assumptions. Indeed, only three PA-based classes of branches were distinguished and it was assumed that all branches of a class have similar characteristics. If this assumptions may be relevant for short axes, it is not the case for long axes that exhibits more variability. Similarly, residual differences may also come from the kind of rules chosen to drive the plant topological development. For instance, numbers of new axes are piecewise constant functions of the Q/D ratio. This choice may not provide

enough flexibility to get a perfect fitting of the real tree lateral structures.

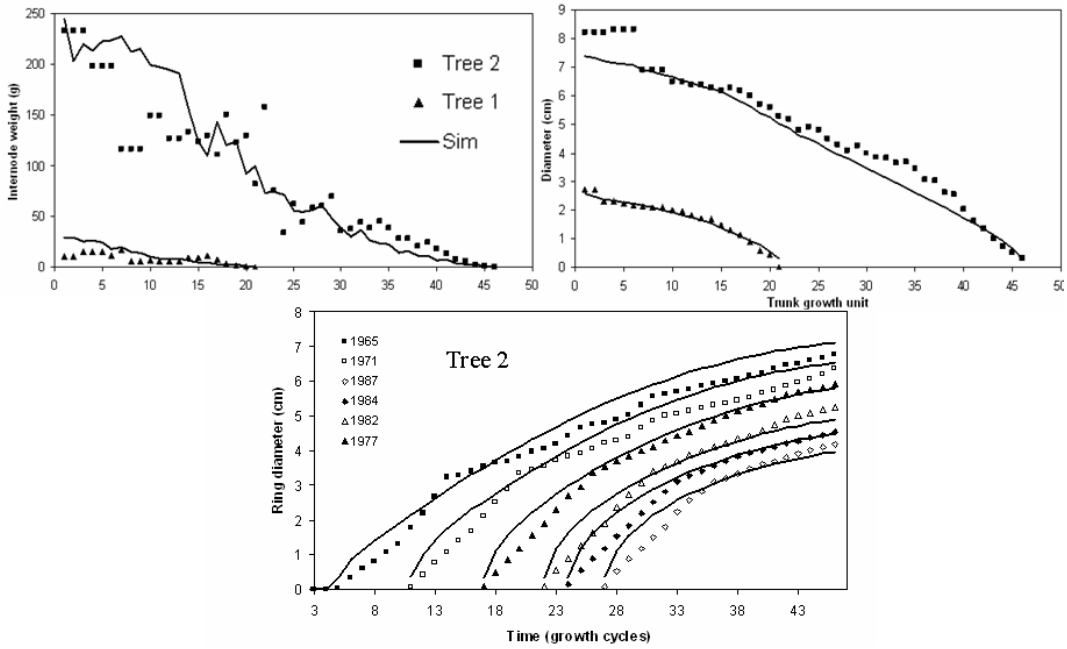


Figure 8.15: Fitting results: observed and simulated G.U. weights and diameters on the trunk; ring profile for tree 2. Series represent the successive ring diameters of growth units according to the year of their appearance.

Topology. Concerning topological parameters, several optima are possible, in particular because the equations 8.8 and 8.9 include transformations from real values to integers. Thus the set of values presented in that table is only one possible solution, not the unique one. The topological parameters were not fitted for branches on the trunk: to allow for comparison with the data, the exact position was recorded and simulated. This approach can be useful for biomechanics where branch positions can be measured precisely on a target tree. When other applications are concerned, such as growth of a simulated beech stand, such detailed information may not be available. In that case, it would require defining the values of branching parameters for the trunk also.
The topological structures and 3D outputs of the fitted trees are shown in Figures 8.17 and 8.18. The numbers of axes are not the same than in the target tree but their demands are similar enough to reproduce the biomass allocation to each compartment and to each growth unit of the stem.

Rings The target data of tree 2 included sequences of ring width for 12 growth units located at different heights near the bottom of the trunk (see Figure 8.2 in section 8.1.2). The fitting results are shown in figure 8.15. There is a slight over-estimation of

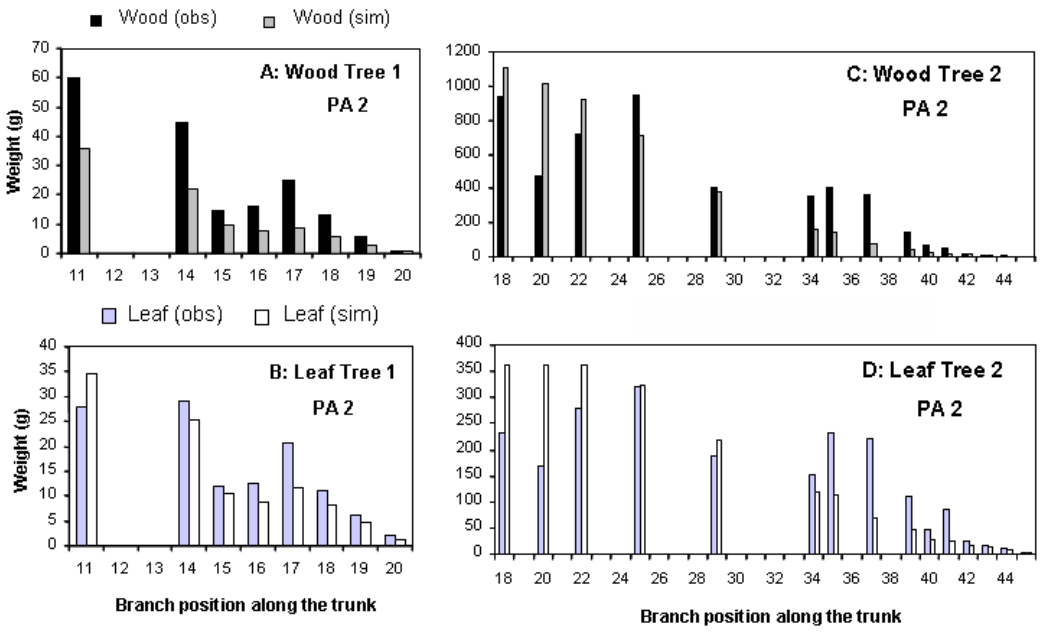


Figure 8.16: Fitting results: Simulated and measured compartment weight on branches of PA 2 for tree 1 and tree 2.

internode pith mass but the global trends are respected. The accuracy of simulation of ring diameters is particularly important as it gives access to information about the past growth of the tree. Thus, not only the last stage but the complete source-sink evolution in time is likely to be correct in the simulated tree.

The value of the exponent γ for ring demand is found to be nearly 3. It is coherent with the common observation that ring width is particularly sensitive to changes of growth conditions. As in the simulation, the environmental conditions were considered as constant, this parameter is likely to be over-estimated to account for higher variations than those generated by source-sink dynamics (Q/D). Concerning the partitioning sub-model for rings to each metamer, the influence of the blade surface located above the considered metamer is limited. The value of the parameter driving the proportion between the two allocation modes is λ=0.13. It means that the ring diameter profile in association with the simulated leaf areas does not follow the Pressler rule. In compensation, a strong effect of branching order was found in the simulation, through the secondary sinks for ring partitioning. Indeed, the value of p^{rg} is much smaller for branches (PA=2, 3, 4) than for the main stem (PA=1). It means that assimilate propagation from leaves to cambial sinks is nearly uniform in axes and preferentially directed into the trunk.

Although further study is needed, these first fitting results show that a simulated tree with a simplified structure can mimic the morphological plasticity of a target tree when the branching structure is controlled by a dynamic feedback of its trophic state on

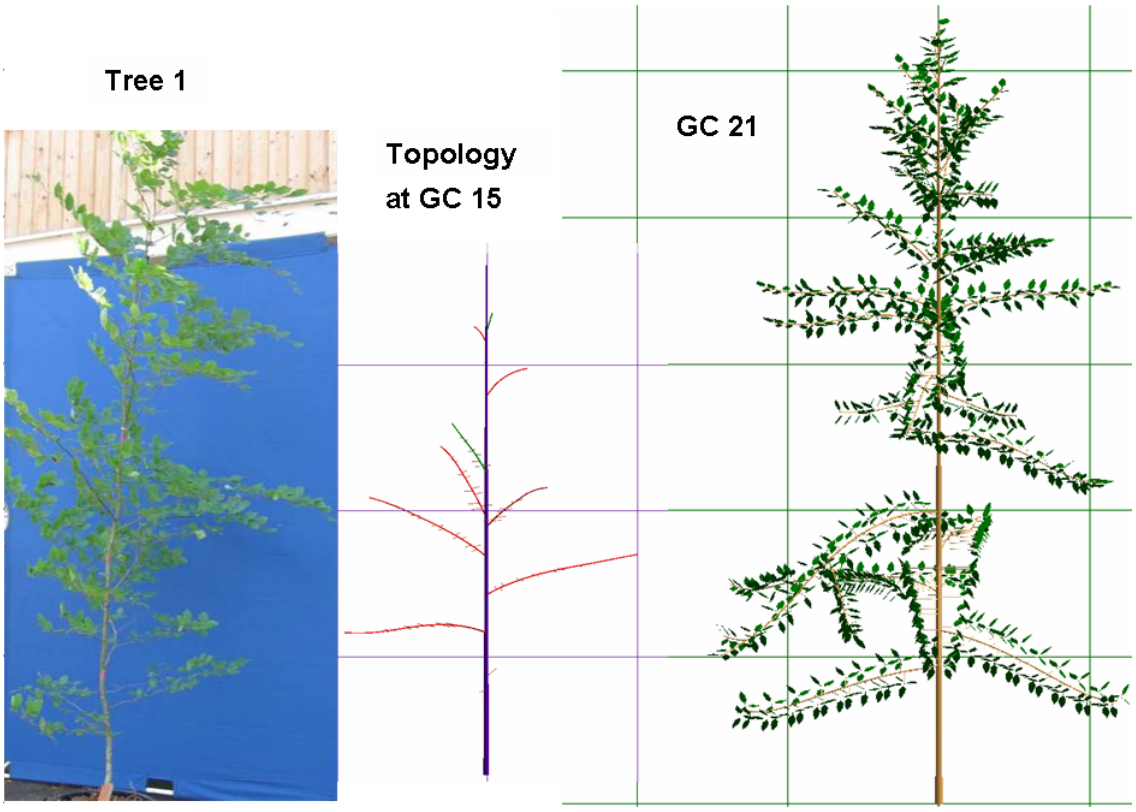


Figure 8.17: Tree 1: 3D output.

its organogenesis processes ($GL3$). The structural factorization results in a natural simplification of the topology: it is no longer based on a detailed description of each branch growth, which would entail facing an inextricable variability. For this reason, GreenLab focuses on the average behavior of the plant and tries to define global rules from botanical observations and estimations of empirical parameters. These first results show that it is possible to reproduce the source-sink dynamics of order 2 branches even if the topology of the plant is not accurately described, as soon as the average development is equivalent.

8.5 Conclusion and plant typology

As presented in section 4.4, several topological structures can give the same variations of biomass compartments. So in some particular cases, the solution for fitting the development may not be unique. That is why it is interesting to take into account the botanical knowledge that have been developed on species-specific trends for tree architecture and development. It can add constraints by defining a maximal or a default topology. A refinement of this methodology would consist in including more detailed

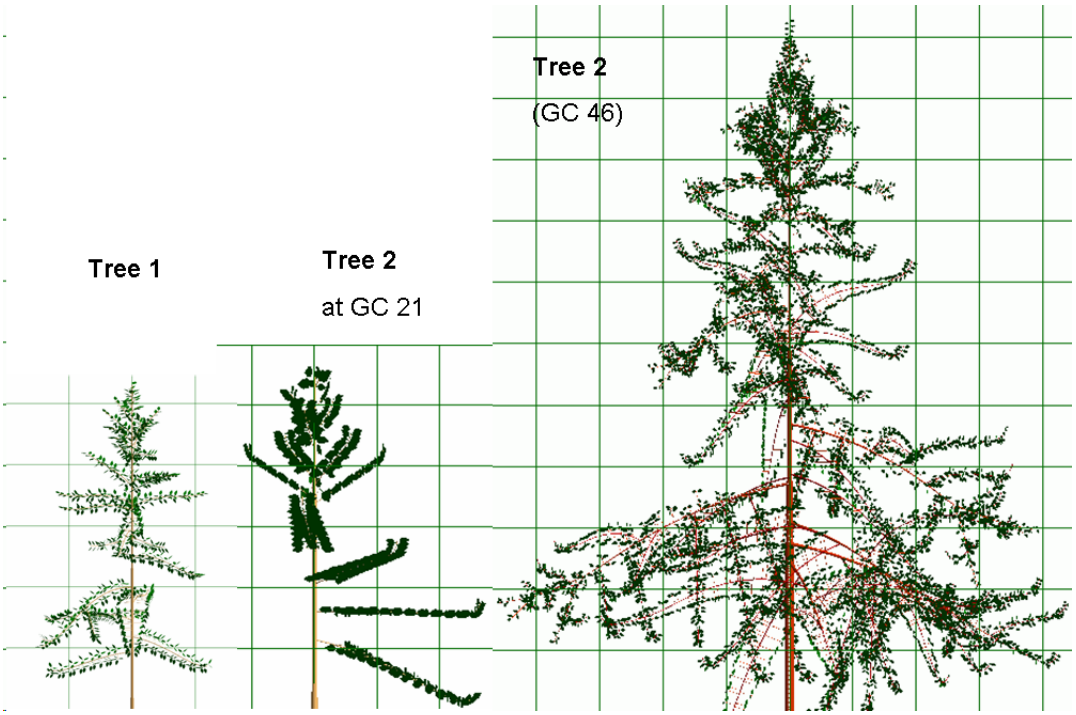


Figure 8.18: Comparison between Tree 1 and Tree 2.

criteria to compare tree architectures. Based on the formalism introduced by Godin and Caraglio [1998] that represents the tree as a rooted multi-scale graph, methods have been developed to define the distance between two tree structures [Ferraro and Godin, 2000] that could be incorporated in the definition of the objective function to minimize. It could provide a precious indicator about the reliability of the simulated architecture which is a weak point of our work. However in our context, the topological structures to compare should be carefully chosen. Ideally, the distance between the whole crown architectures of the simulated and the real trees would be taken into account. But the applicability of that method is questionable. First, except maybe from heavy measurement processes involving 3D digitizing, the complete architecture of a tree cannot be accurately measured. Secondly, in our modeling approach, the simplifications imposed by the complex structure of a tree crown imply that the exact replication of its topology is nearly impossible. Thirdly, the crown architecture is the result of a quasi-infinite number of interactions between unknown events (weather, local competition, insect attacks...) creating the stochastic aspect of the structure, especially for beech trees. The simulation cannot retrace the historic succession of those events and it thus intrinsically restricted to reproducing only its average or global dynamics. So the distance should be computed at an intermediate level to compare structures of branch samples from the target tree to the equivalent average structure of the simulated tree. The comparison procedure is maybe more straightforward when a stochastic simulation of the branching

structure is chosen, since distributions of descriptive variables can be compared. For our experiments on Beech trees, our goal was to find a set of parameters that allows reproducing the growth of two given trees for their trunk dimensions and compartment weights on order 2 branches so the deterministic approach was sufficient. But it opens access to further work including the interactions between stochastic processes driving the topological development and physiological variables.

Figure 8.19 classifies some of the plants studied in applications according to the level of details for their measurements. Another dimension could be added to this graph to account for the GreenLab version considered.

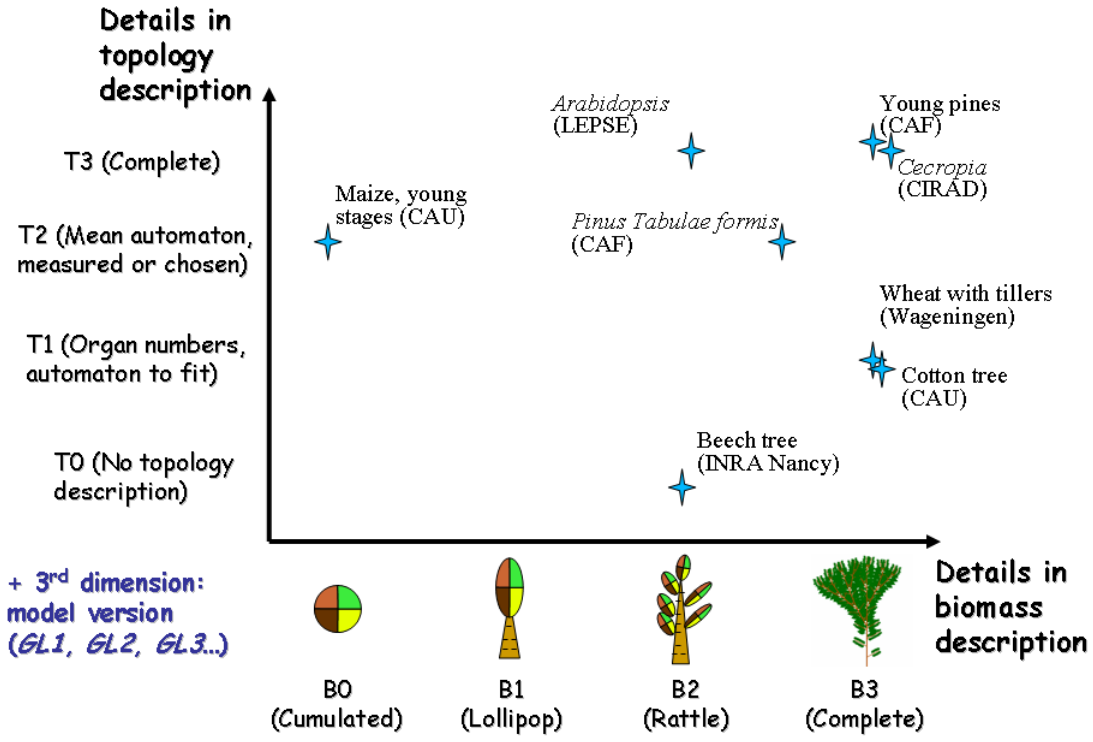


Figure 8.19: Typology of some plants according to the kind of description considered for biomass (X-axis) and topology (Y-axis). A third dimension would account for the model version used for the simulation (not represented).

In this chapter, we have considered the problem of model fitting using simplified targets of different kinds. The simulations took into account the whole topology of the plant with the "complete" model (see definition below). It is useful for all kinds of

applications where the plant architecture needs to be considered, as we have argued in chapter 6. For other applications such as visualization of functional landscapes (e.g. [Le Chevalier et al., 2007]), a lower level of details is needed in particular to reduce the computational cost of tree simulations. But the correct biomass weights and their dynamics are still required. In that context, defining simplified models that have the same global behaviour as the complete one for some aggregated variables has some interest. It is the objective of the next chapter to study the possibility of building such a simplified model from the equations of GreenLab.

Chapter 9

A new perspective for multi-scale analysis

Confrontations of the model to data from branched plants brought the idea that only key-variables of the model could be taken into account so that simplified equations with aggregated variables could be defined. Introducing simplified equations implies reducing the number of parameters and hence the number of degrees of freedom compared to the complete model. Therefore it was important to first test the procedure of parameter identification on the complete model, as done in the previous chapter (chapter 8). The problems solved in this latter step helped designing the new approaches presented here. Note that this part remains a theoretical study and still needs to be applied to real data.

9.1 Preliminary remarks and definitions

Let us give some precisions about the vocabulary and assumptions used in this chapter.

- The *complete model* refers to the model versions that were described in part I. It can be declined in $GL1$, $GL2$ or $GL3$ versions. It means that the whole structure is described up to the highest branching orders. By extension, we call the corresponding simulated tree *complete tree*.

  We set several assumptions concerning the complete tree, although not all trees and shrubs follow these rules:

  - Expansion is immediate for all organs ($t_{exp} = 1$)
  - Leaf functioning time is one growth cycle ($t_a = 1$)
  - SBM (specific blade mass) does not depend on leaf position in trees

- By contrast, the *simplified model* refers to a model version where some variables are aggregated. It can be described at different levels of simplifications. Of course,

the simplification levels of the models are chosen in adequacy with the level of data aggregation for the targets as described in section 8.1. The corresponding simulated tree is referred to as *simplified tree*.

Remark: Regarding branching orders, we remind that the trunk is of order 1, its lateral branches are of order 2 and so on.

9.1.1 "Meta-organ"

We introduce the term "meta-organ" to define an organ that is used to replace a pool of organs when switching from the complete model to simplified ones. In the general case, it consists of four biomass compartments: internode, rings, leaf and fruit (and/or flower). A petiole compartment could be distinguished if needed. It has hopefully the same general behaviour as the structure it stands for (named the *equivalent structure*). At a given time, two structures, or a meta-organ and a structure, are said to be *equivalent* if they have same compartment biomass and same production. This is an equivalence relation in the mathematical sense of the term.
Let Ω_t denote the set of meta-organs at growth cycle t. $\widetilde{S}_t$ is the set of plant structures at growth cycle t generated from the alphabet $\widetilde{A}(t)$. We define $\omega : \widetilde{S}_t \rightarrow \Omega_t$ as the application that associates to a structure its equivalent meta-organ. In section 4.4, we proved that ω is non-injective, *i.e.* that two different structures can generate the same meta-organ as soon as they have identical compartment weights and biomass production. By definition, the physiological age of the meta-organ $\omega(S_p(n,t))$ is the physiological age of the basis of the structure $S_p(n,t)$, namely p.
In the simplest case, meta-organs consist only of wood and blade compartments. Meta-organs can thus be represented as growing axes with one more internode and one blade added at each growth cycle. The problem is to determine the demand of these compartments and the rules defining their variations and biomass allocation. Hereafter we use the following abbreviation: if P is a parameter driving the meta-organ functioning: P_ω stands for $P(\omega(S_p(n,t)))$ whenever there is no ambiguity on the structure to consider. For example, P^i_ω is the sink of the internode compartment of the meta-organ $\omega(S_p(n,t))$. If there is possible ambiguity, we adopt the notation $P_\omega(S_p(n,t))$.

9.2 Objective equivalences

We present the equivalences that we want to keep when switching from simplified to complete models. In particular, at all scales, we impose conservation of the key variable Q/D. The chosen set of equivalences is consistent with the target data formats presented in section 8.1 (See Figure 8.1).

9.2.1 "Cumulated" level

The whole tree is transformed into a meta-organ: $\omega(S_1(t,t))$. The objective criteria that we impose to the Cumulated model in reference to the complete model are:

- Equality of the total biomass production of the plant at every growth cycle
- Equality of biomass compartments of the tree: total wood weight, blade weight, ring weight, fruit or flower weight.

9.2.2 "Lollipop" level

The whole crown is considered as a meta-organ but the trunk is detailed. The objective criteria that we impose to the Lollipop model in reference to the complete model are:

- Equality of the total biomass production of the plant at every growth cycle
- Equality of biomass compartment of the crown
- Equality of internode masses and dimensions (length, diameter) on the trunk

9.2.3 "Rattle" level

Branches of order 2 are transformed into meta-organs whereas the trunk is modeled at the level of phytomers as in the complete model. The objective criteria that we impose to the Rattle model in reference to the complete model are:

- Equality of the total biomass production of the plant at every growth cycle
- Equality of biomass compartment on branches of order 2
- Equality of internode masses and dimensions (length, diameter) on the trunk

9.3 Theoretical equivalences: aggregation of the complete model variables to build simplified models

In this part, we assume that the parameters of the complete model are known and we try to deduce information at aggregated levels, i.e. for the simplified models. We build a simplified model by aggregating the variables of the complete model so that we may keep some key-variables constant from one scale to another one. We express the parameters of simplified models as functions of that of the complete model. For the sake of clarity, we focus on the "Rattle" level; all results can be easily transposed to the other two levels. Similarly, we do not consider fruit compartments of meta-organs: they follow the same equations as internode compartments.

Notations: Hereafter we write vectors with square brackets ($[V]$) and $\cdot$ denotes the usual definition of the scalar product in $\mathbb{R}^n$. Wherever it is not specified, the notation $[V][k]$ will represent the k-th component of $[V]$ with $k \in \{1..P_m\}$.

9.3.1 Equation of biomass production

Formalism based on hydraulic resistances

If the hydraulic structure of the plant needs to be taken into account, the topology influences biomass production. Calculating the biomass production of the simplified model requires defining the equivalent resistance of the structures that are replaced by meta-organs. Let us consider for instance the particular case of negligible hydraulic resistance of axes: the tree is considered as a network of leaves connected in parallel. We study the conditions to get the same production from a structure $S_p(n,t)$ as from its meta-organ $\omega(S_p(n,t))$.
We remind that the biomass production of a blade with area $S(t)$ at growth cycle t is (paragraph 3.1.2) :

$$q(t) = \frac{E(t)}{\frac{R_1}{S(t)} + R_2} \tag{9.1}$$

The meta-organ production is determined by:

$$q_\omega(t) = \frac{E(t)}{R_\omega} \tag{9.2}$$

where R_ω denotes the resistance of the meta-organ. So, to get similar production at every growth cycle, we impose that

$$\forall t,\ 1 \leq t \leq N_m,\ \frac{1}{R_\omega} = \sum_{k=p}^{P_m} \frac{[\psi^a(S_p(n,t))][k]}{\frac{R_1}{S_k^a(t)} + R_2} \tag{9.3}$$

We can use this equality to deduce an expression of R_ω as a function of the model parameters (and independent of $S_k^a(t)$ that are variables of the model).
Let us consider a meta-organ of physiological age p and chronological age n, $\omega(S_p(n,t))$, having its blade area equal to the cumulated areas of all blades of the corresponding structure:

$$\forall p \in \{1..P_m\},\ \forall n \leq t,\ S_\omega(S_p(n,t)) = \sum_{k=p}^{P_m} [\psi^a(S_p(n,t))][k] \cdot S_k^a(t) \tag{9.4}$$

where $[\psi^a(S_p(n,t))][k]$ is the k^{th} component of the vector $\psi^a(S_p(n,t))$ as defined in section 2.5, i.e. the number of leaves of PA k in the structure $S_p(n,t)$. $S_k^a(t)$ denotes

the area of a blade of PA k in that structure. Let us define $[S^a(t)]$ as the vector of size P_m whose components are $S^a_k(t)$. Thus eq. 9.4 can be written under a more compact form as:

$$S_\omega(S_p(n,t)) = [\psi^a(S_p(n,t))] \cdot [S^a(t)] \tag{9.5}$$

where $\cdot$ denotes the scalar product. Note that the components $[\psi^a(S_p(n,t))][k]$ are null if $k < p$. As we assume that expansion duration lasts only one growth cycle, blade area only depends on its appearance cycle and not on its chronological age. Moreover, this assumption implies the following relation between blade areas of different physiological ages:

$$\forall (p,k) \in \{1..P_m\}^2,\ S^a_p(t) = \frac{P^a_p}{P^a_k} \cdot S^a_k(t) \tag{9.6}$$

Indeed, these leaves competed for the same biomass pool and got amounts of biomass in proportion to their respective sinks. To get the same biomass production from the meta-organ as from the complete structure, we need to set the equivalent resistance of the meta-organ equal to the resistance of the network composed of the structure leaves. It implies in particular that the blade area of each leaf of PA k can be expressed depending on the blade area of the meta-organ S_ω:

$$\begin{aligned} \forall k \in \{1..P_m\},\ S^a_k(t) &= \frac{S_\omega}{\sum_{k'=p}^{P_m} \frac{P^a_{k'}}{P^a_k} \cdot [\psi^a(S_p(n,t))][k']} \\ &= \frac{S_\omega \cdot P^a_k}{[\psi^a(S_p(n,t))] \cdot [P^a]} \end{aligned} \tag{9.7}$$

where $[P^a]$ is the vector of size P_m whose components are respectively the sinks of blades for each PA. Therefore from the equality of hydraulic resistances of the meta-organ and of the structure (eq. 9.3), an expression of R_ω can be deduced by replacing $S^a_k(t)$ by its expression:

$$R_\omega = \frac{1}{\sum_{k=p}^{P_m} \frac{[\psi^a(S_p(n,t))][k]}{\frac{R_1 \cdot [\psi^a(S_p(n,t))] \cdot [P^a]}{P^a_k \cdot S_\omega} + R_2}} \tag{9.8}$$

with the abbreviated notation $S_\omega = S\left(\omega(S_p(n,t))\right)$ representing the blade area of the meta-organ corresponding to the structure $S_p(n,t)$.
So if the resistance of the meta-organ $\omega(S_p(n,t))$ is calculated at every G.C. as in expression 9.8 as a function of its area S_ω and of the parameters of the complete model, then the biomass production of the meta-organ is identical to that of the structure $S_p(n,t)$.

Beer Law formalism

This part simplifies a lot if the Beer law formalism only is chosen, that is to say if we consider that water uptake is not a limiting factor and that hydraulic resistances have only a weak impact. In that case, the equation of biomass production is Eqn. 3.5:

$$Q(t) = E(t) \cdot \mu \cdot Sp \cdot \left(1 - e^{-k \cdot \frac{S(t)}{Sp}} \right)$$

Here biomass production is considered at whole-plant scale only. As soon as the total blade area is the same in the simplified model as in the complete model, the equation of biomass production with the same parameters gives the same result. Hence characterizing the parameters of the simplified model amounts to the problem of biomass allocation. If allocation to the biomass compartment is the same in the simplified model as in the complete model, the biomass production will be also identical.

9.3.2 Allocation

Sinks of equivalent structures

The demand of a meta-organ $\omega(S_p(n,t))$ is defined as the aggregated demand of the set of organs of the structure $S_p(n,t)$. Therefore the sinks of internode and blade compartments, respectively noted P^i_ω and P^a_ω, are expressed as follows:

$$\begin{cases} P^i_\omega = \sum_{k=p}^{P_m} [\psi^i(S_p(n,t)) - \psi^i(S_p(n-1,t-1))][k] \cdot P^i_k \\ P^a_\omega = \sum_{k=p}^{P_m} [\psi^a(S_p(n,t)) - \psi^a(S_p(n-1,t-1))][k] \cdot P^a_k \end{cases} \tag{9.9}$$

with the definitions of section 2.5: ψ^o gives the number of existing organs on the structures. To simplify the notations, let $[\psi'(S_p(n,t))]$ be the vector of the numbers of *new* metamers in the structure $S_p(n,t)$, *i.e.* the number of metamers appearing in the structure of PA p and CA n at GC t. Note that $[\psi'(S_p(n,t))][k] = 0$ if $k < p$ as there are no metamers of PA less than p on a structure of PA p. Therefore the meta-organ sinks are written as scalar products:

$$\begin{cases} P^i_\omega = [\psi'^i(S_p(n,t))] \cdot [P^i] \\ P^a_\omega = [\psi'^a(S_p(n,t))] \cdot [P^a] \end{cases} \tag{9.10}$$

At each growth cycle, a new internode appears in the internode compartment of the meta-organ. Its sinks depend on its rank. Figure 9.1 shows the shape of meta-organ

sinks for internodes in the rattle simplified model. The sink shape is similar for blades (but not identical, as blade sinks are set different from those of internodes in the complete model). As the $GL1$ version is considered, all branches of order 2 have the same development with only a temporal shift that depends on the time of their initiation. Therefore the graphs representing the demand of internode compartment of meta-organs as functions of time all have the same shape: there are only translated of one GC when switching from one meta-organ to the upper one. These values are calculated preliminary to the simulation and are given as inputs to simulate the simplified tree. The

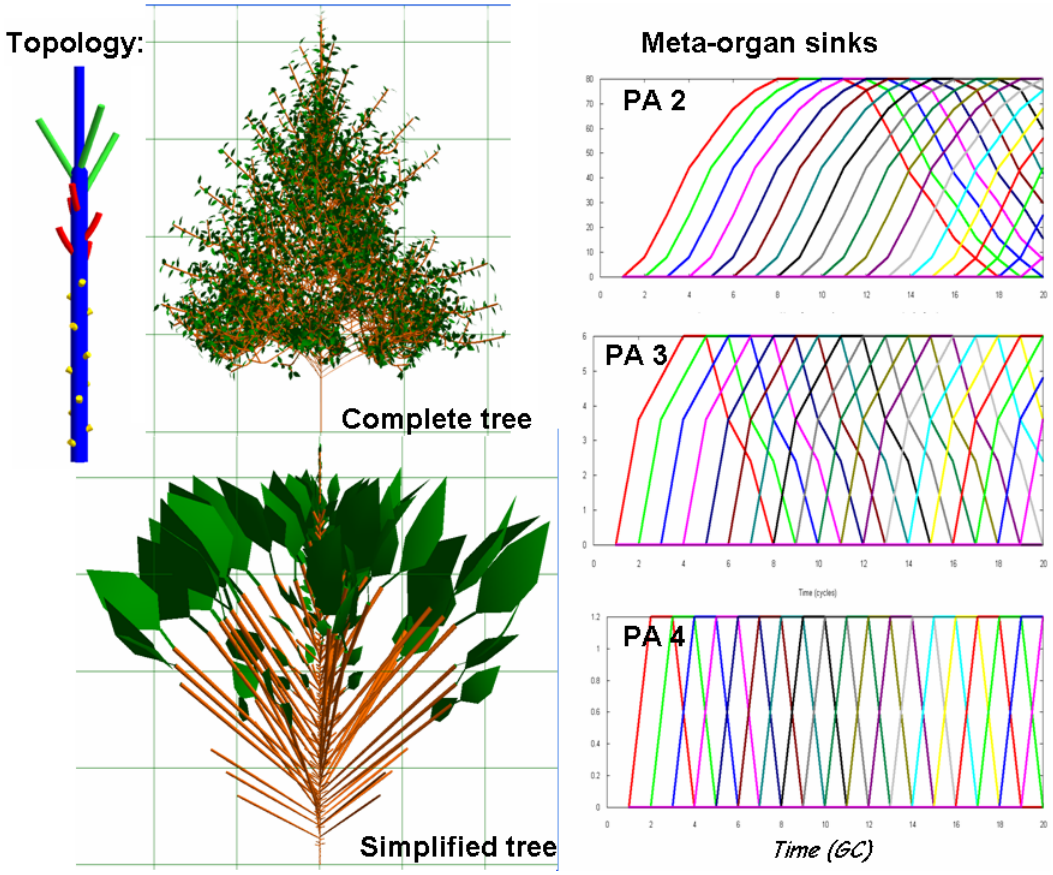


Figure 9.1: Simplified model with equivalent structures (meta-organs) at "rattle" level: shape of sink variation function for branches of order 2. Case of deterministic development ($GL1$) with no ring demand. The graphs represent the sink of meta-organ internode P^i_ω as a function of time.

complete and simplified trees presented in this figure thus have exactly the same profiles of biomass production and of allocation to compartments on branches of order 2. Their trunks are identical. The meta-organ sinks for blades and internodes only depend on the structure development and not on its topology nor on the structure morphology.

This is not the case for the ring compartment.

Allocation to ring compartments

Global allocation. Regarding secondary growth, let us first deal with the step of global allocation.

- If modes Q/D or Q (equations (3.24) and (3.26)) are chosen, the equality of biomass allocated to the ring compartments in the complete and simplified models is straightforward since we impose the conservation of these variables. Therefore the parameters of the global secondary sinks of the simplified model can be simply set identical to that of the complete model.

- In the case of the first allocation mode (*Leaf Number*), the parameters of the simplified model are of the form:

$$P_0^{rg}{}_{\omega} = P_0^{rg} \cdot \sum_{k=1}^{P_m} [\psi^a(S_1(t,t))][k] \tag{9.11}$$

  where P_0^{rg} is the sink of the ring compartment of the complete model and the sum represents the total number of leaves of the tree (structure of PA 1 and of CA equal to the current GC, t).

Thus we can have the same biomass allocated to the ring compartment $Q_{ring}(t)$ in the simplified model as in the complete model.

Local allocation. More difficulties are raised for the step of local allocation (section 3.2.3). Let us write the expression of the biomass allocated to the ring compartment of the meta-organ $\omega(S_p(n,t))$ at GC t. It imposes classifying the different metamers not only according to their physiological and chronological ages but also to their ontogenetic age. Let $[\psi'_o(S_p(n,t))]$ be the vector of the numbers of new metamers for each PA in the structure $S_p(n,t)$. Thus we have:

$$q_{\omega}^{rg} = \sum_{k=p}^{P_m} \sum_{m=1}^{n} \sum_{o=1}^{t} \underbrace{\left[\psi'_o\left(S_p(n-m+1, t-m+1)\right)\right][k]}_{\text{Nb phyto of CA m, PA k and OA o in Sp(n,t)}} \cdot q_k^{rg}(m,o,t) \tag{9.12}$$

Remark: Contrary to chronological age, ontogenetic age of a newly created organ can be higher than 1. This is why the index of the summation on o can reach the maximal value of t.

We remind that the biomass $q_k^{rg}(m,o,t)$ allocated at growth cycle t to a metamer of physiological age k, chronological age m and ontogenetic age o is calculated as:

$$q_p^{rg}(m,o,t) = \left(\frac{1-\lambda}{D_{Pool}(t)} + \frac{\lambda \cdot N_p^{a,a}(m,o,t)}{D_{Pressler}(t)}\right) \cdot p_p^{rg} \cdot l_p(m,t) \cdot Q_{ring}(t) \tag{9.13}$$

where the notations have been defined in section 3.2.3. It depends in particular on the number of leaves located above the metamer in the architecture $N_k^{a,a}(m,o,t)$ (topological information) and on the length of the metamer $l_k(m,t)$ (morphological information). It is possible to deal with the topology problem since the number of leaves above a metamer can be defined from the parameters of the model in the $GL1$ case (see paragraph 2.5). But the metamer length varies with the dynamics of Q/D. Thus metamer length cannot be included as such in the expression of the parameters of the simplified model.

To express q_ω^{rg} as a function of the parameters of the complete model and of the state variables of the simplified model only, we can take advantage of the proportionality relation between weights of internodes that appeared at the same G.C.:

$$q_k^i(n,t) = \frac{P_k^i}{P_{k'}^i} \cdot q_{k'}^i(n,t) \tag{9.14}$$

where $q_k^i(n,t)$ is the weight of internode of PA k and CA n at GC t. As internode expansion lasts no longer than one growth cycle, internode sink is defined as a constant and noted P_k^i. Therefore a relation between internode weights in the complete model and the internode weight of the meta-organ q_ω^i can be derived:

$$\begin{aligned} q_\omega^i &= \sum_{k=p}^{P_m} [\psi'^i(S_p(n,t))][k] \cdot q_k^i(n,t) \\ \Longrightarrow q_k^i(n,t) &= q_\omega^i \cdot \frac{P_k^i}{[\psi'^i(S_p(n,t))] \cdot [P^i]} \\ &= q_\omega^i \cdot \frac{P_k^i}{P_\omega^i} \end{aligned} \tag{9.15}$$

where q_ω^i stands for $q_\omega^i(S_p(n,t))$. It represents the amount of biomass that appeared at cycle t in the internode compartment of structure $S_p(n,t))$.

The metamer length is obtained from the allometric rules defined in section 3.2.2:

$$\begin{aligned} l_k(n,t) &= \sqrt{Be}\left(q_k^i(n,t)\right)^{\frac{1+\beta}{2}} \\ &= l_\omega \cdot \left(\frac{P_k^i}{P_\omega^i}\right)^{\frac{1+\beta}{2}} \end{aligned} \tag{9.16}$$

By replacing $l_k(n,t)$ in equation 9.13 and then in equation 9.12, q_ω^{rg} eventually becomes:

$$q_\omega^{rg} = \left(\frac{1-\lambda}{D_{Pool}(t)} \cdot K_{\omega,Pool} + \frac{\lambda}{D_{Pressler}(t)} \cdot K_{\omega,Pressler}\right) \cdot Q_{ring}(t) \tag{9.17}$$

with the following notations:

$$\begin{cases} K_{\omega,Pool} = \sum_{m=1}^{n} \frac{l^i_{\omega(S_p(m,t))}}{P^{i\ \frac{1+\beta}{2}}_{\omega(S_p(m,t))}} [\psi'(S_p(n-m+1,t-m+1))] \cdot [p^{rg}(P^i)^{\frac{1+\beta}{2}}] \\ K_{\omega,Pres} = \sum_{m=1}^{n} \sum_{o=1}^{t} \frac{l^i_{\omega(S_p(m,t))}}{P^{i\ \frac{1+\beta}{2}}_{\omega(S_p(m,t))}} \cdot [\psi'_o(S_p(n-m+1,t-m+1))] \cdot [p^{rg}(P^i)^{\frac{1+\beta}{2}} N^{a,a}(m,o,t)] \end{cases} \tag{9.18}$$

where $l_\omega(S_p(m,t))$ is the length of the new metamer in meta-organ $S_p(m,t)$, assuming that the allometric rules are conserved (*i.e.* that $l^i_\omega = \sqrt{Be}(q_\omega)^{\frac{1+\beta}{2}}$). $D_{Pool}(t)$ and $D_{Pressler}(t)$ are calculated as the respective sums of the demands of growth units of the trunk and the sum of factors $K_{\omega,Pool}$ and $K_{\omega,Pressler}$ of all meta-organs of the plant at GC t. This formulation allows defining the equations and the parameters of a simplified model with conservation of the amount of biomass allocated to the ring compartment. Note that equations of the simplified model (eqn. 9.17) have not the same form as that of the complete model. It implies that in the simplified model, the calculation of the demand of growth units of the stem and that of meta-organs do not follow the same rules.

The shapes of parameter variations $K_{\omega,Pool}$ and $K_{\omega,Pressler}$ are given in figure 9.2 for a $GL1$ virtual tree. It represents the changes of K_ω with time for a meta-organ (in $GL1$, the shape is identical for all meta-organs). In the $K_{Pressler}$ calculation, the major components are demands of growth units of main axis: indeed, their demand is multiplied by the number of leaves of all their lateral structures. Hence mutations of the main axis induce the decline of this factor. By contrast in K_{Pool} parameter, all substructures have similar influences and therefore death of lateral branchlets can be observed in the variations of K_{Pool}.

To summarize, we have extracted the parameters that allows calculating the demand of a meta-organ as functions of the parameters defining the behaviour of its equivalent structure in the complete model. Under our assumptions (no fruits), meta-organ primary demand is characterized by two parameters: internode sink and blade sink. These parameters depend on meta-organ physiological age. They vary with time when meta-organs age (curves of figures 9.1). Meta-organ ring demand is characterized by two variables ($K_{\omega,Pool}$, $K_{\omega,Pressler}$). Note that these are variables and not parameters since they depend on pith length l_ω (figure 9.2).

If the formalism of Beer-Lambert law is chosen for biomass production (section 9.3.1), the equations of the simplified model are the same as that of the complete model (except for ring biomass repartition among meta-organs that conforms to equation 9.17).

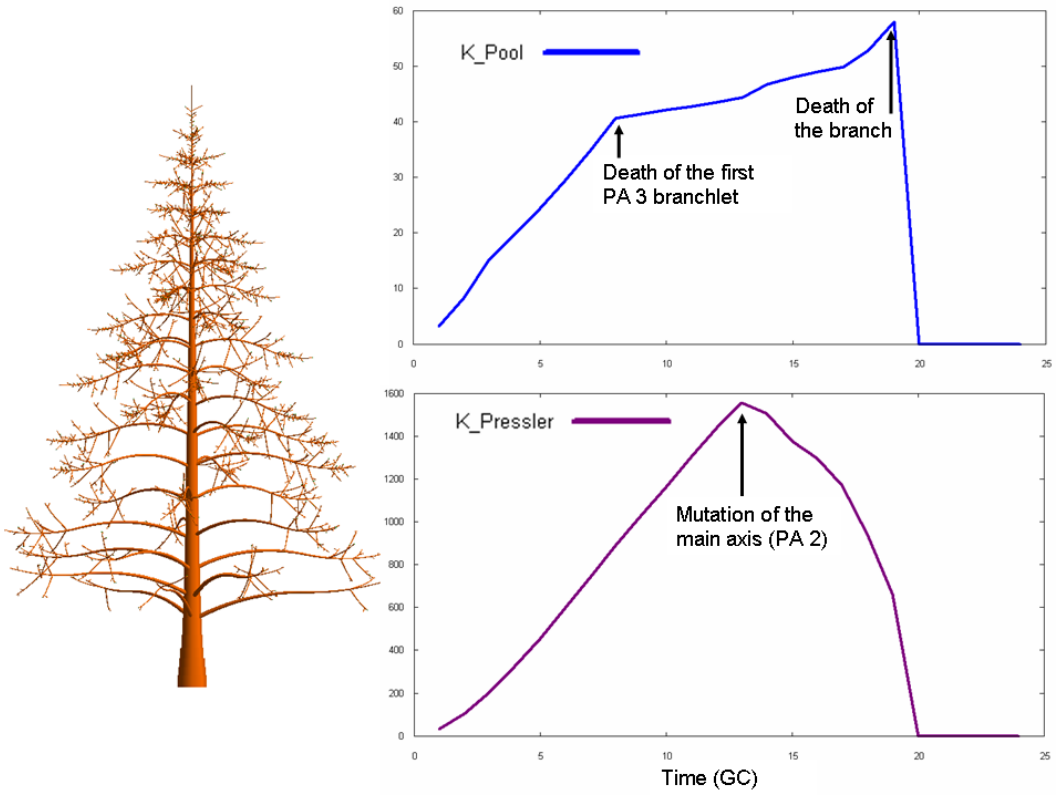


Figure 9.2: Changes of parameters $K_{\omega,Pool}$ and $K_{\omega,Pressler}$ with time for meta-organs of a virtual $GL1$ tree.

9.3.3 $GL3$ buds

Let us consider now the case where numbers of axillary axes per growth unit are not constant, as in the $GL3$ version [Mathieu, 2006]. We remind that hereafter new growth units are abusively named "buds".

First step: global allocation (number of potential buds). Biomass is first allocated to bud compartment (Q_{bud}) according to the demand of potential buds at whole-plant scale. We denote P_p^{pb} the sink strength of potential buds of physiological age p. The plant demand for potential bud compartment at growth cycle t, $D_{bud}^{pb}(t)$, is:

$$D_{bud}^{pb}(t) = \sum_{p=1}^{P_m} [\psi'^{pb}(S_1(t,t))][p] \cdot P_p^{pb} \tag{9.19}$$

with the notations of the previous section: $[\psi'^{pb}(S_1(t,t))][p]$ is the number of potential buds of physiological age p at growth cycle t. Therefore the demand of potential buds of a meta-organ can be defined as the sum of demands of potential buds of its equivalent

structure. This is similar to the method applied for defining sink strengths of internode and blade compartments (section 9.3.2). Sink strengths of meta-organ potential buds (P_ω^{pb}) are thus defined as follows:

$$P_\omega^{pb} = \sum_{k=p}^{P_m} [\psi'^{pb}(S_p(n,t))][k] \cdot P_k^{pb} \tag{9.20}$$

The plant demand for bud compartments is calculated as the sum of meta-organs bud compartment demands.

Second step: Number of active buds. After the step of global allocation to bud compartments, the number of active buds per growth unit zone, *i.e.* buds that effectively gives rise to new shoots, are determined as affine function of the ratio $\frac{Q_{bud}}{D_{bud}^{pb}}(t)$ in the complete model [Mathieu, 2006]. It can be bounded by the number of potential buds per zone as:

$$b_{pq}(t) = \min \left\{ b_{pq}^{pb}(t),\ \lfloor B_{pq}^1 + B_{pq}^2 \cdot \frac{Q_{bud}}{D_{bud}^{pb}}(t) \rfloor \right\} \tag{9.21}$$

where $b_{pq}^{pb}(t)$ is the number of potential buds in a zone Z_{pq} at growth cycle t. It varies with time if there are bud dormancy or if the number of metamer per zone is variable.

To aggregate the variables of this complete model equation, we introduce the notion of *active bud sink* of a meta-organ. Indeed, determining the *number* of active buds in the complete structure amounts to determining which *proportion* of potential bud sink becomes active in the simplified model. There is a switch from discrete variables in the complete model to continuous variables in the simplified model. Therefore equivalences will be of course only approximate for that part. Active bud sink strength of a meta-organ $\omega(S_p(n,t))$ can be bounded by its potential bud sink strength P_ω^{pb}. It is written as:

$$P_\omega^{ab} = \min \left\{ P_\omega^{pb},\ B_{p,\omega}^1 + B_{p,\omega}^2 \cdot \frac{Q_{bud}}{D_{bud}^{pb}}(t) \right\} \tag{9.22}$$

The number of active buds in a structure is a piecewise constant function of Q/D in the complete model; we choose an affine function for the active bud sink strength in the simplified model (see figure 9.3).

We get the parameters of this equation 9.22 by calculating a line equation approximating

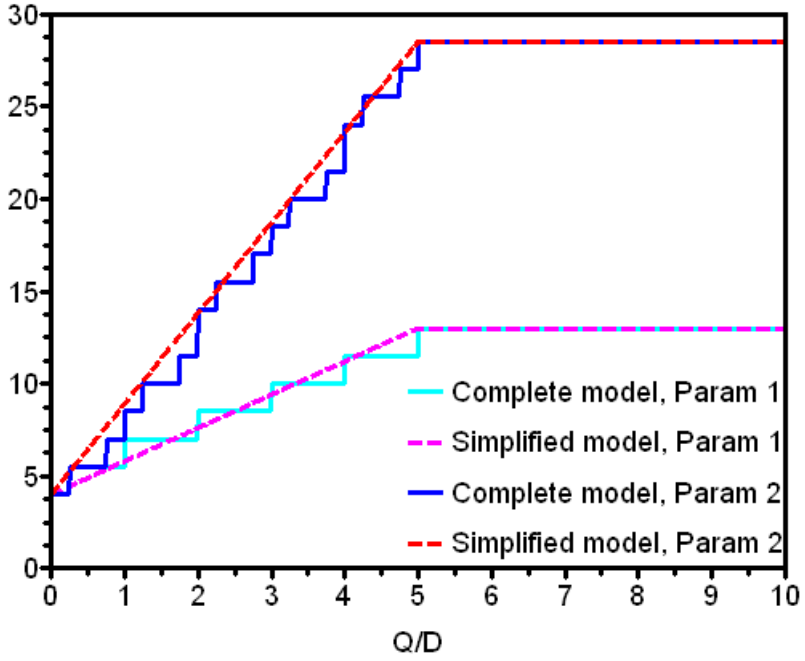


Figure 9.3: Variations with Q/D of structure bud sinks in the complete model (piecewise constant functions) and equivalent meta-organ bud sink (affine function) in the simplified model for two sets of parameter values[2]

the piecewise constant function:

$$\begin{cases} B^1_{p,\omega} &= \displaystyle\sum_{k=p}^{P_m}\sum_{q=k}^{P_m} \psi_\omega[p][q]\cdot P^{pb}_q \cdot \min\{b^{pb}_{pq}(t),\ \lfloor B^1_{pq}\rfloor\} \\ B^2_{p,\omega} &= \dfrac{\displaystyle\sum_{k=p}^{P_m}\sum_{q=k}^{P_m} \psi_\omega[p][q]\cdot P^{pb}_q \cdot \left(b^{pb}_{pq}(t) - \lfloor B^1_{pq}\rfloor\right)^+}{\max_{p,\ q}\left\{\dfrac{b^{pb}_{pq}(t) - B^1_{pq}}{B^2_{pq}}\right\}} \end{cases} \tag{9.23}$$

where $\psi_\omega[p][q]$ denotes the total number of zones Z_{pq} in the structure $S_p(n,t)$.
These equations are sufficient if the number of potential buds per zone does not depend on time. But parameters $b^{pb}_{pq}(t)$ vary in particular with the number of metamers per

[2]Parameter values for a virtual structure containing three zones:

- Parameter values, set 1. Bud sinks: $P^{pb}_2 = 1$, $P^{pb}_3 = 1.5$; Zone Z_{12}: $b^{pb}_{12} = 1$, $B^1_{12} = 1$, $B^2_{12} = 0.5$; Zone Z_{23}: $b^{pb}_{23} = 1$, $B^1_{23} = 0.5$, $B^2_{23} = 2$; Zone Z_{13}: $b^{pb}_{13} = 7$, $B^1_{13} = 2$, $B^2_{13} = 1$
- Parameter values, set 2. Same values as set 1 except: $b^{pb}_{12} = 3$; $b^{pb}_{23} = 10$

zone which is also an affine function of Q/D ratio in the $GL3$ version [Mathieu, 2006]:

$$b_{pq}^{pb}(t) = M_{pq}(t-1) \cdot m_{pq} \text{ with } M_{pq}(t-1) = \lfloor M_{pq}^1 + M_{pq}^2 \cdot q_p^b(t) \rfloor \tag{9.24}$$

where $M_{pq}(t)$ is the number of metamers in zone Z_{pq} at growth cycle t, m_{pq} is the number of buds per metamer,$q_p^b(t)$ is the biomass amount allocated to a growth unit of physiological p at growth cycle t, M_{pq}^1 and M_{pq}^2 are parameters of the affine function. No simple expression was found: the parameters have to be defined after successive iterations of the model to get the values of Q/D. What is important at this point is that active bud sinks of meta-organ depend on potential bud sink P_ω^{pb}, on the ratio $\frac{Q_{bud}}{D_{bud}^{pb}}(t)$ but also on the biomass allocated to each growth unit of the meta-organ at the previous growth cycle.

9.4 Perpsectives for simplified models

In the previous section, we have extracted the values of the parameters of the simplified model as functions of the parameters of the complete model. These equations allow keeping strict equivalences between complete and simplified models, at least in the $GL1$ version. It raises now the problem: how to deal with real trees ? Answering this question requires defining the equations of simplified models for applications to fitting of real trees. The equations presented above can help for this choice.

9.4.1 First approach: beta functions to define parameter variations.

The variations of aggregated parameters, *i.e.* meta-organ sinks, mainly take bell shapes (see figures 9.1 and 9.2), at least in $GL1$. Therefore they can be approximated by beta density functions. So a reasonable strategy would be to define the meta-organ sinks as beta functions whose parameters would be fitted against the data. This is of course an option but it does not take advantage of the knowledge acquired from the complete model. Moreover, it does not account for the tree species: beech trees or pine trees would differ only from their final fitted values for beta law parameters. It means no *a priori* information would be integrated before fitting, such as general branching rules or the hierarchical organization of physiological ages.

9.4.2 Second approach

Discrete coefficient values for internode and blade sinks.

Case $GL1$

We can benefit from the $GL1$ property that the sequence of structures of order 2 at a given time from the top branch to the bottom of the tree represent the topological development of a structure that would be one GC older from one level to the following one. Their sink shapes are the same for all meta-organs (see figure 9.1). Therefore it is possible to extract these data from biomass measurements at rattle level.
Let $W_k^i(n,d)$ and $W_k^a(n,d)$ denote respectively the internode and blade biomass increments added at cycle d on the structure of CA n and PA k. Thanks to the common pool assumption, we get:

$$\begin{aligned}\frac{Q}{D}(d) &= \frac{W_k^i(0,d)}{P_\omega^i(S_k(0,d))} = \cdots = \frac{W_k^i(d,d)}{P_\omega^i(S_k(d,d))} \\ &= \frac{W_k^a(0,d)}{P_\omega^a(S_k(0,d))} = \cdots = \frac{W_k^a(d,d)}{P_\omega^a(S_k(d,d))}\end{aligned} \tag{9.25}$$

where $W_k^a(0,d)$ is the biomass of the trunk blade whose axillary bud gives rise to the structure S_k. By definition, the blade sink of PA 1 is set to one. Therefore we can extract the expression of the meta-organ sinks:

$$\forall n \in \{1..d\},\ \forall k \in \{1..P_m\},\ \begin{cases} P_\omega^i(S_k(n,d)) &= \dfrac{W_k^i(n,d)}{W_k^a(0,d)} \\ P_\omega^a(S_k(n,d)) &= \dfrac{W_k^a(n,d)}{W_k^a(0,d)} \end{cases} \tag{9.26}$$

Figure 9.4 shows the curves obtained from this calculation on data from the simulated tree of figure 9.1. Of course, the sink shape is the same as the theoretical one (calculated from numbers and sinks of organs on the structure). It allows simulating the simplified tree presented in figure 9.1 with conservation of the variables listed in section 9.2.3. As the $GL1$ version is considered, these equivalences are verified for the whole growth duration although data are virtually collected only at the last growth cycle ($d = 20$ on figure 9.4).
Thus it is possible to define the variations of internode and blade compartment sinks as vectors of coefficients giving their value in function of meta-organ chronological age. Unfortunately, the variations of $K_{\omega,Pool}$ and $K_{\omega,Pressler}$ cannot be directly assessed from the data in the same way. They remain to be represented by beta functions and fitted. Another option is to give up equation 9.17, as discussed in section 9.4.2.

For trees, the $GL1$ case is hardly ever encountered: the topological structure is not as regular and depends strongly on ontogenetic or environmental changes. However, we

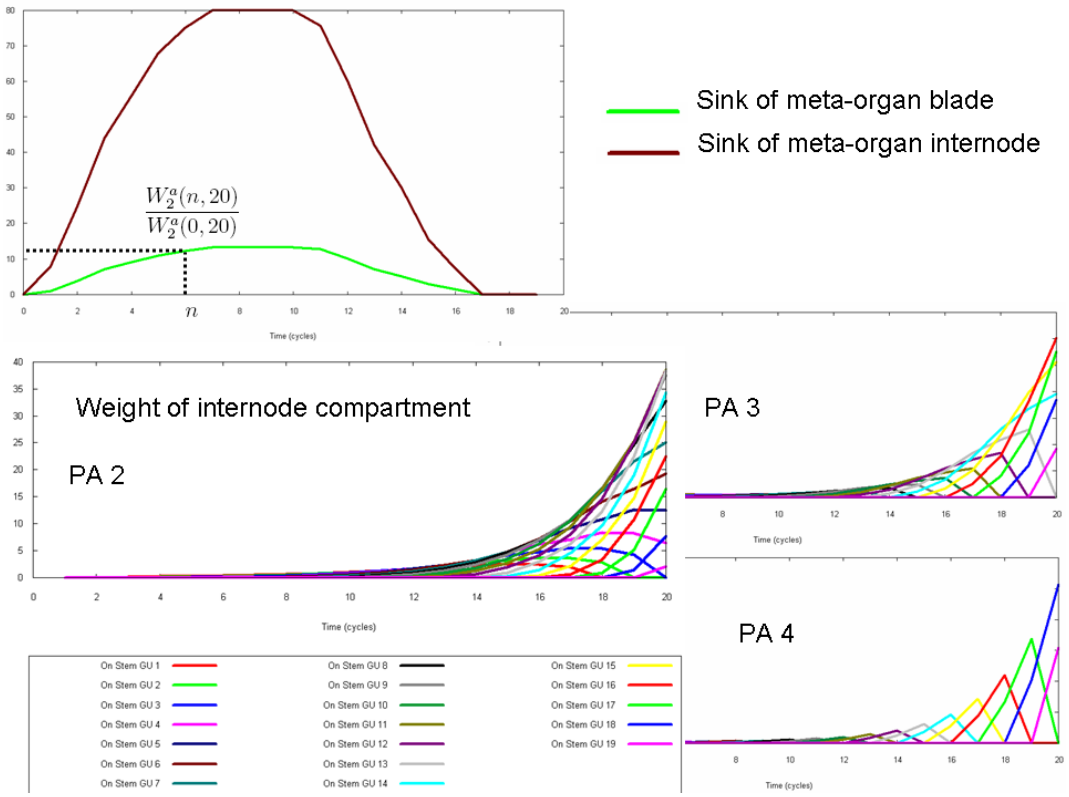


Figure 9.4: Fitting at rattle level for a virtual tree: meta-organ sinks calculated from "rattle" data and corresponding variations of meta-organ compartment weights for each PA. The sink shape is the same as the theoretical one (figure 9.1).

can use the corresponding tree topology as a reference framework to propose solutions for the $GL3$ version.

Case $GL3$

In the $GL3$ case, the method of parameter aggregation did not provide simple expressions of parameters for a simplified model. Moreover, when considering real trees, the $GL3$ parameters driving numbers of new axes and new metamers would need to assessed. Therefore we propose the following approach:

1. From botanical analysis, from bibliography or from direct observations, a reference (or potential) topology can be chosen. This was done for example for beech trees in section 8.1.3: it corresponds to the level of topology description referenced as T2 level. It means defining the topological rules of a $GL1$ automaton that the plant would follow in its steady state.

2. From sample data on branches tips, the values of sink strengths for organs of each

physiological age can be obtained. Indeed, assuming that expansion duration is one growth cycle only implies that sink ratios can be directly assessed from measurements. This was done for pine trees (e.g. see figure 7.3 for young pine trees sinks).

3. These data are sufficient to calculate the meta-organ sinks of the reference $GL1$ tree, using equations 9.10. Thus we have the shapes of internode and blade sink variations of meta-organs. They are respectively denoted $P^i_{\omega,GL1}$, $P^a_{\omega,GL1}$. Note that in the $GL3$ case, we can also include bud sink variation, $P^b_{\omega,GL1}$ (bud sinks can be estimated from sample data the same way as internode or blade sinks: it was done for beech trees, section 8.4.1).

4. These sink variation shapes are input as default values for these parameters. They are then updated at each simulation growth cycle depending on Q/D ratio. To be consistent with $GL3$ modelling choices and equation 9.22, we propose the following expression:

$$P_\omega = P_{\omega,GL1} \cdot (B^1_p \cdot q^b_\omega(t-1) + B^2_p \cdot \frac{Q}{D}(t)) \tag{9.27}$$

where P_ω stands for respective meta-organ sinks and $q^b_\omega(t-1)$ is the biomass received by meta-organ bud compartment (*i.e.* new shoots) at cycle $t-1$. This term accounts for the influence of number of metamers initiated at growth cycle $t-1$ on the number of potential buds at cycle t, as presented in equation 9.24. B^1_p and B^1_p are parameters of the simplified model that need to be fitted. They take different values according to the type of parameter (internode, blade or bud sinks) and according to meta-organ physiological age (p).

An advantage of this method is that we keep information on the default branching rules. Therefore, species-specific topological characteristics are taken into account in the fitting procedure and in the results. In the $GL1$ case, the equivalences between simplified and complete models are verified since we use equations with aggregated variables derived directly from the equations of the complete model. In $GL3$, there is no strict equivalence with a complete model that would be fitted on the same data but it is consistent with the $GL3$ philosophy.

Simplified model for ring biomass partitioning

Ring sinks. Firstly we can notice that it is possible to estimate the secondary sinks for ring biomass allocation to metamers (p^{rg}_k) from sample data as it was done for primary growth. Indeed, equation 9.13 implies a relationship between ring sinks of several physiological ages that appeared at the same growth cycle. Let us introduce the

following variable:

$$F_k = \frac{\dfrac{q_k^{rg}(m,t)}{l_k(m,t)} - \dfrac{q_k^{rg}(m',t)}{l_k(m',t)}}{N^{a,a}(m,t) - N^{a,a}(m',t)} \text{ with } m,\ m' \in \{1..t\} \tag{9.28}$$

We have:

$$\forall k,\ k' \in \{1..P_m\},\ \frac{p_k^{rg}}{p_{k'}^{rg}} = \frac{F_k}{F_{k'}} \tag{9.29}$$

As we set for reference value that $p_1^{rg} = 1$, this relation provides estimations of ring sinks for every higher physiological age.
Therefore it would be possible to define simplified equations for ring partitioning using the same method as presented above for primary growth. It implies adding a new assumption on internode lengths in the complete model: they have to be arbitrarily fixed. It amounts to considering that structure demand for ring increment is calculated according to numbers of metamers rather than metamer lengths. Under this strong assumption, it is possible to define $K_{\omega,Pool,GL1}$ and $K_{\omega,Pressler,GL1}$ as parameters that can be calculated from parameters of the complete model.
An immediate consequence is that if pruning is taken into account in the calculation of default parameters $K_{\omega,Pool,GL1}$ and $K_{\omega,Pressler,GL1}$, then ring demand is calculated considering only living branches.

More relevant choice for ring biomass partitioning A major inconvenient of the simplified model issued from variable aggregation is that simulation of ring biomass partitioning among meta-organs is not mechanistic enough. Indeed, it relies on equation 9.17 where the effective meta-organ blade biomass does not intervene. The influence of blade biomass on ring partitioning is only taken into account through the definition of variables $K_{\omega,Pressler,GL1}$ *i.e.* through the default $GL1$ topology. This modelling choice allowed conserving equivalences with the complete model. But it could be modified if more degrees of freedom were left to the simplified model. We propose a solution based on studies of possible slight modifications of the complete model in some particular cases. It does not strictly respect the equivalences in the general case but it provides a mechanistic model consistent with the GreenLab modelling approach.
A first remark is that defining simplified models would be easier if continuous additive variables were used in the complete model. For instance in the process of ring biomass partitioning, blade areas could be considered instead of numbers of leaves as driving variables. Under our assumptions, cumulated blade areas are conserved when switching from the simplified to the complete model whereas numbers of leaves change (several leaves on a structure but only one leaf on its associated meta-organ). Similarly, we switch from internode pith length to pith biomass which is an additive variable in the equation of ring biomass partitioning.

Let us consider that all internodes of a given structure have the same primary and secondary sinks (*i.e.* P^i and P^{rg} do not depend on physiological age). In that case, leaf areas above every internode appeared at a given growth cycle forms a partition of total leaf area at the current cycle (see figure 9.5).

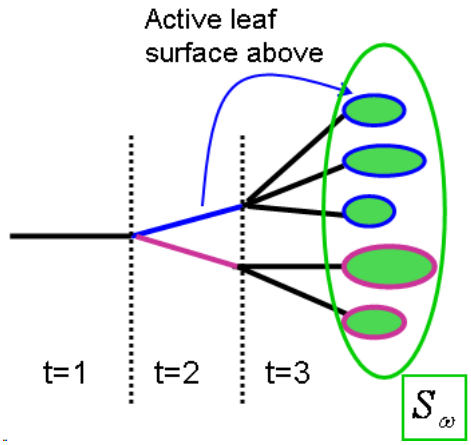


Figure 9.5: Blade area of a meta-organ is partitioned between leaf areas seen by sets of metamers appeared at every growth cycles.

Meta-organ blade area can be factorized in the second term of equation 9.17 which implies the following simplified expression for ring biomass increment of this structure at cycle t:

$$q_\omega^{rg} = \left(\frac{1-\lambda}{D_{Pool}(t)} \sum_{m=1}^{n} q_\omega^i(S_p(n-m+1, t-m+1)) + \frac{\lambda}{D_{Pressler}(t)} \cdot S_\omega^a \right.$$
$$\left. \cdot \sum_{m=1}^{n} \frac{P^i}{P_\omega^i(S_p(n-m+1, t-m+1))} q_\omega^i(S_p(n-m+1, t-m+1))\right) \cdot P^{rg} \cdot Q_{ring}(t) \tag{9.30}$$

where $q_\omega^i(S_p(n-m+1, t-m+1))$ is the amount of biomass allocated to meta-organ internode compartment at cycle $t-m+1$. It corresponds to biomass of piths of chronological age m.
Therefore, a coherent proposal for calculation of biomass allocated to meta-organ ring compartment is the following:

$$q_\omega^{rg} = \left(\frac{\lambda S_\omega}{D_{Pressler}(t)} \cdot \sum_{m=1}^{n} \frac{q_\omega^i(S_p(n-m+1, t-m+1))}{P_\omega^i(S_p(n-m+1, t-m+1))} \right.$$
$$\left. + \frac{1-\lambda}{D_{Pool}(t)} \cdot \sum_{m=1}^{n} q_\omega^i(S_p(n-m+1, t-m+1))\right) \cdot Q_{ring}(t) \tag{9.31}$$

We remind that D_{Pool} and $D_{Pressler}$ are calculated by summing the components of plant ring demand (demand of growth units of the main stem and demands of existing meta-organs).

To summarize, we have studied in this chapter the consequences of variable aggregation of the complete model in order to extract the equations of a simplified model. We focused on "Rattle" level which is the most detailed so that the results are easily generalized. Given a set of equivalences to conserve at every scale change, we first defined the equations of the simplified model. These equations are similar to those of the complete model except for biomass allocation to rings and $GL3$ parameters that drive the topological development. These last processes are strongly dependent on plant topology and thus only approximative equivalences can be obtained.
Then we proposed solutions to deal with real plants. The equations derived from variable aggregation of the complete model do not necessary provide the more relevant modelling choice. Therefore other options can be considered: equivalences with complete model are not conserved but the approach is consistent with the GreenLab philosophy.
Concerning biomass allocation to primary growth (meta-organ internode and blade compartments), two approaches are distinguished according to the model version. In $GL1$, meta-organ sinks can be assessed directly from data (figure 9.1). In $GL3$, we propose to take advantage of a $GL1$ reference automaton to define the default shapes of meta-organ sinks. The effective values are then updated at each growth cycle depending on Q/D ratio.
Concerning biomass allocation to rings, a similar method can be applied under the assumption that structure demand for ring increment is calculated according to numbers of metamers rather than metamer lengths. But a more mechanistic modelling is that of equation 9.31 where meta-organ leaf area is taken into account.

Conclusion

This work took place in a particular development phase of the GreenLab model. GreenLab is a mathematical model of plant growth that can be expressed under the form of a discrete dynamic system. It is a generic structure-function model that accounts for processes of plant topological development, biomass production and allocation at mesoscale (organ level). After the Phd of Amélie Mathieu [Mathieu, 2006] who introduced feedback effects of plant trophic balance on its development, an important step of confrontation of the model to real plants was needed (which is of course an on-going work). To deal with real plants, we needed to consider the specificities of each of them: it implied introducing some new features in the model such as biomass remobilization at organ senescence or variations of several functional and topological variables (sinks, allometries, RUE, number of axes ...) with time or with phytomer rank. We studied the changes in biomass production and allocation due to changes of plant topology. We found that similar production sequences could be generated by plants with different development sequences under some conditions: either if organ expansion duration is equal to one growth cycle or if blade functioning time is infinite.

We presented in this work several methods to deal with parameter identification of branched plants with complex topology. The most detailed target data include a complete topological description of plants and weights of their individual organs. This kind of data was for instance collected on young Chinese pines (*Pinus tabulaeformis*, data from Guo Hong, CAF) and on *Cecropia sciadophylla* (Patrick Heuret and Camilo Zalamea, CIRAD). We defined three levels of simplification of target data both for topological and dimensional plant description. Regarding dimensional data, these levels were characterized as cumulated ($B0$, compartment weights only), "lollipop" ($B1$, complete description of main stem, compartment weights of all lateral branches) and "rattle"($B2$, complete description of main stem, compartment weights of each lateral branch). Regarding topology, three levels of simplification were also distinguished: sample data ($T0$), mean and variance of numbers of organs per physiological age ($T1$), description of a default automaton ($T2$).

For instance, *Arabidopsis thaliana* was measured following a rattle pattern for biomass ($B0$ level) and complete topological description was recorded. It allowed defining its development rules for inflorescence axes ($T2$ level). Pine trees (*Pinus tabulaeformis*, data from Guo Hong, CAF) were simulated with a complete average topology defined

from the measurement ($T2$). The fitting results were compared in the case of a complete average target and a simplified target (section 8.2.2).

Measurements on wheat with tillers included numbers of organs at several growth stages with repetitions (Kang MengZhen, Univ. of Wageningen-LIAMA). Although wheat topology is not really stochastic (the numbers of tillers can be related to plant density), the data were adequate to test our method. It was thus possible to fit the stochastic version of the model using the theoretical mean and variance of the numbers of organs ($T1$). Only a maximal topology was set; the effective number of organs that appeared at each growth cycle was determined according to the probability values. These numbers were non-integer values since theoretical expressions of average demand were considered. Then functional parameters were estimated using the mean number of organs at each growth cycle. Therefore parameter identification is done for the average plant and not on particular realizations of stochastic plants that would each have a different topology (section 8.3.2). Similar procedure is on progress for cotton tree data with Li Dong (CAU).

Regarding beech trees (Thiéry Constant, LERFoB), a potential topology was chosen according to the botanical knowledge on its architecture. Then the topological and functional parameters were fitted on data from biomass compartments for branches and individual organs for the trunk ($B2$). Measurement of ring widths at several positions along the trunk allowed testing the model for biomass allocation and partitioning to ring growth. Including ring profile in the set of target data provides precious clues about the past trophic dynamics of trees. In particular in the model, it is possible to link ring width variations to variations of blade area ratios of meta-organs at each level and at different growth cycles (section 8.4.1). We wish to point out that both for pine trees and beech trees, a major difficulty in the experimental work associated to the fitting procedure was to determine the age of the target tree. As several morphological markers tend to disappear or to be deformed with time and as ring separations are not always clear, there might be errors on the determination of tree chronological age. This can be an important source of errors since morphogenesis processes are strongly linked with time.

Plants can be classified according to the level of details of their description (figure 8.19) and to the model version chosen for their simulation (deterministic $GL1$, stochastic $GL2$, mechanistic $GL3\ldots$). These choices also depend on the objectives of the fitting procedure and on the potential applications. For real trees, the $GL1$ case is hardly ever encountered: the topological structure is not as regular and depends strongly on ontogenetic or environmental changes. However, it is possible to consider a regular $GL1$ topology as a working approximation. But it is also important to propose solutions for the $GL3$ version as done for beech trees and in the perspectives presented in chapter 9. Note that the $GL2$ version could be considered also for trees: the same procedure as presented in section 8.3.1 for wheat could be applied at plant level. Means and variances would be calculated from branch and metamer numbers of a single tree instead

of being calculated from measurements of several plants of a population.

The applications presented in that work consist of parameter identification for equations of the complete model but with simplified targets (section 8). Different procedures have been proposed according to the model version ($GL1$, $GL2$, $GL3$). From this work, a second approach has emerged. It consists in defining new equations and variables that are consistent with the level of target description (chapter 9).

For some applications, the first solution should be privileged. As discussed in chapter 6, GreenLab trees have a simplified architecture that allows computational efficiency and realistic simulations of topological changes due to ontogenetic or environmental constraints. Moreover, tree architecture is important for several kinds of applications, even under its simplified form: for instance to simulate biomechanical stresses in the trunk, to account for - and to be influenced by - competition effects (especially competition for light), for visualization at tree scale. As branching rules are defined according to botanical knowledge, it allows characterizing tree species and their architectural models.

However, a more simplified model can be useful for stand- or landscape-scale simulations where only the dynamics of biomass allocation are required or where large numbers of individuals are involved (even with structural factorization, reducing computational can be necessary if trees are old and with potentially high physiological ages). Visually, all trees look the same regardless of their species (they take "lollipop" shape or "rattle" shape). For visualization purpose, representations of simplified trees would need to be improved, for instance using foliar density instead of a single blade in meta-organs. Different meta-organ shapes could be defined according to tree species.

As described in paragraph 9.4, we can benefit from topological information derived from the complete model to define default shapes for meta-organ sinks. In $GL1$, these default sinks gives the same biomass partitioning in meta-organs as in the corresponding structure of the complete model. In $GL3$, it provides reference values for meta-organ sinks that are then updated at each growth cycle according to values of Q/D ratio. It implies that meta-organ growth conforms to the same growth phases as the corresponding structures. Their default sink curves account for basal effect in the first stages of structure development and for aging effects until branch death. Thus defining simplified rules for meta-organ growth is analogous to defining rules of the automaton in the complete model. Another argument in favour to simplified models is the reduction of the number of parameters. With simplification of target description comes a reduction of the number of data available. Therefore it is important to reduce also the number of parameters. This is precisely the objective of simplified models: simulating the dynamics of aggregated variables with a limited set of parameters.

Equations of simplified models were extracted from these of the complete model using a method of variable aggregation. These equations gives exact equivalences between the complete and simplified models for a set of key state variables (determined according to the level of simplification chosen). However they are not suitable as such to define

the simplified model because parameters and variables of the complete model cannot be separated. Moreover, these formulations would have to be updated when the complete model changes. Therefore we proposed other solutions in adequacy with GreenLab modelling approach.
To conclude, we proposed in this work several procedures for parameter identification of the GreenLab model on plants with complex architectures. It opens access to applications concerning prediction and optimization of tree growth for decision support systems in forest management. It can also be used as a framework for other research fields such as analysis of biomechanical stresses in tree trunks.

Part III

Appendices

Appendix A

Notations and Parameters

Table A.1: Table of main abbreviations and parameter notations. To facilitate the reading, the same index meanings were chosen in most formulae.

Abbreviations	
G.C.	Growth Cycle
G.U.	Growth Unit (set of metamers produced during one growth cycle)
P.A.	Physiological Age
C.A.	Chronological Age
O.A.	Ontogenetic Age (number of GC since the initiation of the axis)
GL	GreenLab
$GL1$	Deterministic version with fixed development [Yan et al., 2004]
$GL2$	Version with stochastic development [Kang et al., 2003]
$GL3$	Deterministic version with Q/D feedback control [Mathieu, 2006]
Index conventions	
t	Growth cycle (chronological age of the plant)
n	Organ chronological age
p	Physiological age of bearing metamer
q	Physiological age of axillary metamer
o	Organ
a	Leaf
i	Internode
f	Fruit, flower
r	Root
rg	Ring
Parameters	
P_m	Maximal Physiological age
N_m	Maximal growth cycle of simulation
$S_p(n,t)$	Structure of basis of PA p and CA n at G.C. t
$[\psi^o(S_p)][k]$	Number of organs o of PA k in the structure S_p
Z_{pq}	Set of metamers of PA p and axillary buds of PA q
$b_{pq}(t,t)$	Number of active buds in a zone Z_{pq} at GC t
$a_{pq}(t,t)$	Total number of active buds in plant zones Z_{pq} at GC t
Q_t	Biomass production at GC t
t_a	Duration of leaf photosynthetic activity (in G.C.)
E_t	Value of environmental variable at GC t
μ	Parameter of biomass conversion efficiency
Sp	Reference ground surface of the plant
D_t	Plant demand at GC t
t_{exp}	Duration of organ expansion (in G.C.)
P_k^o	Sink strength of organ o and PA k
ϕ^o	Sink variation function of organ o

Appendix B

Normalization factor for sink variation function

We remind that the final weight of an organ of physiological age j appeared at cycle k can be written:

$$Q_o(j,k) = P(j) \cdot \sum_{i=1}^{T_{exp}} \phi(i, T_{exp}) \cdot \frac{Q(k+i-1)}{D(k+i-1)} \tag{B.1}$$

We adopt the following notation:

$$\beta[a,b,T_{exp}](i) = \left(\frac{i+0.5}{T_{exp}}\right)^{a-1} \cdot \left(1 - \frac{i+0.5}{T_{exp}}\right)^{b-1} \tag{B.2}$$

Expansion duration T_{exp} can be variable according to the appearance cycle of organs, k. The sink variation coefficient at chronological age j for an organ appeared at cycle k can be written:

$$\phi(j,k) = \frac{1}{N(k)} \cdot \beta[a,b,T_{exp}(k)](j) \tag{B.3}$$

The role of this normalization factor is important when T_{exp} is variable. In the following paragraph, shapes of sink variation functions are defined for a reference value of T_{exp} (usually the maximal one), denoted T_{exp}^{ref}, and we analyze the changes in this shape for a different value of T_{exp}.

B.1 First method: normalization of integral value

In this mode, we normalize the area covered by the curve associated to sink variation functions. As only discrete values are considered for coefficients of sink variation, it boils down to normalizing the sum of the beta coefficients:

$$N = \sum_{i=1}^{T_{exp}} \beta[a,b,T_{exp}(k)](i) \tag{B.4}$$

The calculation of N depends on the expansion duration T_{exp}. It has to be calculated for each organ independently if the expansion durations are different (see graph B.1A). Thus the sum of the coefficients $\phi(j)$ is always equal to 1. In that case, when the expansion duration T_{exp} decreases, the values of the sink variation function increase. **The notion of sink strength corresponds to the cumulated demand of the organ through its whole life-time**. It implies that the sink strength of an organ can be defined as the final weight that the organ would take if the ratio of biomass supply over demand (Q/D) were constant and equal to 1:

$$Q_o = P \cdot \underbrace{\sum_{i=1}^{T_{exp}} \cdot \phi(i)}_{=1} = P \tag{B.5}$$

The final weight does not depend on the sink variation function. The shape of the beta law only influences the dynamics of the organ growth but not the final biomass.

B.2 Second method: homothetic transformation

When expansion duration changes, all points of the curve are multiplied by the ratio of initial to final expansion duration.

$$N(k) = \frac{T_{exp}^{ref}}{T_{exp}(k)} \cdot \sum_{i=1}^{T_{exp}^{ref}} \beta[a, b, T_{exp}^{ref}](i) \tag{B.6}$$

Therefore sink variation coefficients of an organ appeared at cycle k are written:

$$\phi(j, k) = \frac{T_{exp}(k)}{T_{exp}^{ref} \cdot \sum_{i=1}^{T_{exp}^{ref}}} \cdot \beta[a, b, T_{exp}(k)](j) \tag{B.7}$$

The sum of coefficients for all chronological ages j is equal to 1 only for organs having an expansion duration equal to T_{exp}^{ref} (see graph B of figure B.1). If expansion duration decreases, organ sink strength decreases also.

B.3 Third method: contraction transformation

The same normalization factor is kept for all organs regardless of their expansion duration. It is calculated for the reference expansion duration and kept as a constant:

$$N(k) = \sum_{i=1}^{T_{exp}^{ref}} \beta[a, b, T_{exp}^{ref}](i) \tag{B.8}$$

Therefore sink variation coefficients of an organ appeared at cycle k are written:

$$\phi(j,k) = \frac{1}{N} \cdot \beta[a,b,T_{exp}(k)](j) \tag{B.9}$$

As the normalization factor is independent of k, the sum of coefficients for all chronological ages j is also equal to 1 only for organs having an expansion duration equal to T_{exp}^{ref} (see graph C of figure B.1).

B.4 Fourth method : maximum set to 1

We choose to normalize the sink variation function by setting its maximum to one. If $a > 1$ and $b > 1$, we get:

$$N(k) = \left(\frac{a-1}{a+b-2}\right)^{a-1} \cdot \left(\frac{b-1}{a+b-2}\right)^{b-1} \tag{B.10}$$

It is independent of k and more importantly, it is independent of expansion durations. In cases where $a \leq 1$ or $b \leq 1$, the normalization factor is equal to the maximal value taken by coefficients $\beta[a,b,T_{exp}^{ref}](j)$. In that case, **sink strength is defined as the maximal value taken by organ demand**.

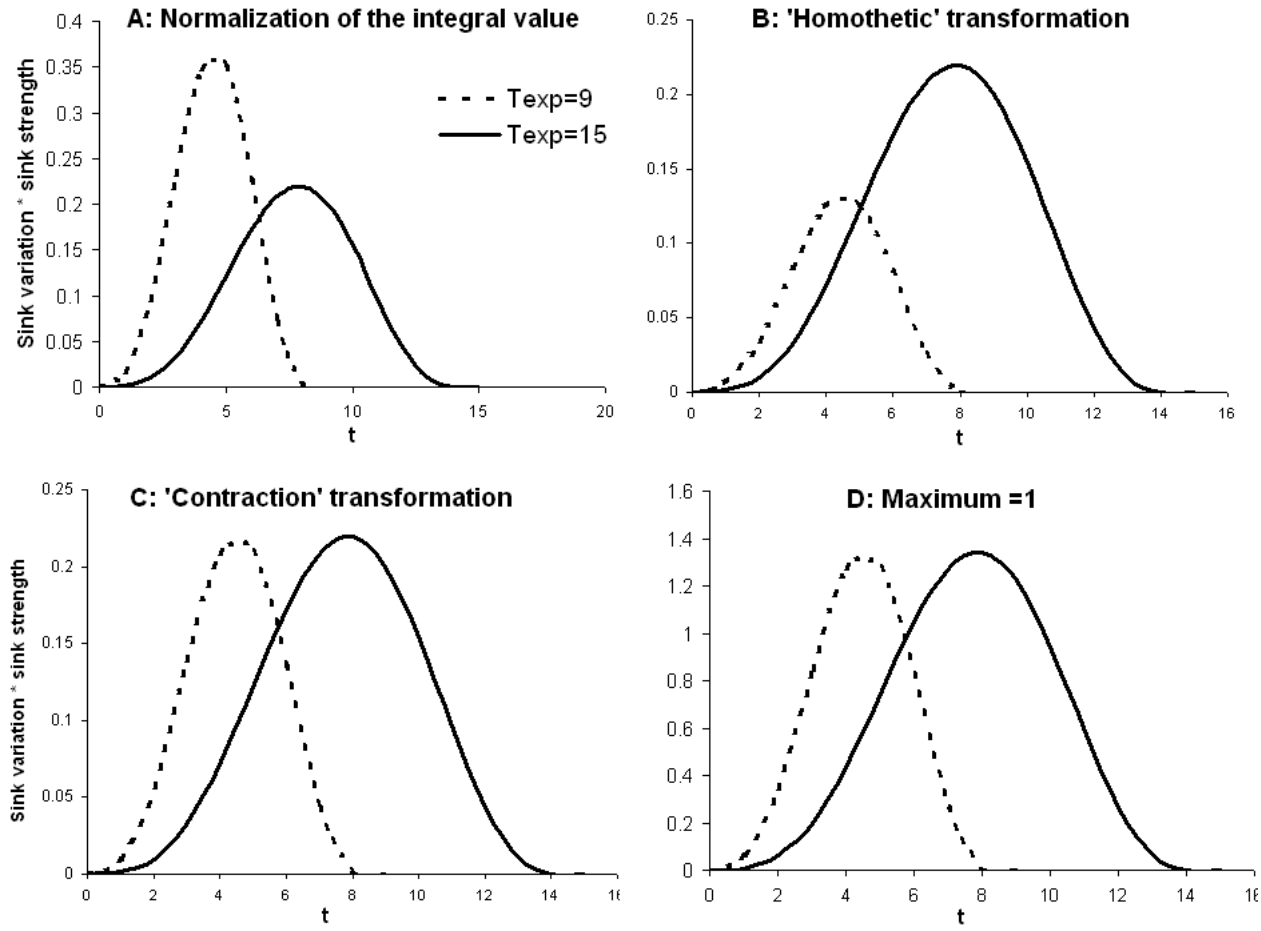


Figure B.1: Two curves representing the function of sink variations multiplied by sink strength (P=1.35), with different methods for the calculation of the normalization factor(N): integral value set to one (A), homothetic transformation (B), contraction (C) and maximum set to one (D).

Appendix C

Computation of the plant total demand

In the GreenLab model, the demand of some of the plant compartments (bud, root, ring) is a function of the ratio of Q/D (see section 3.2). Hence the total plant demand is calculated at each time step as the solution of the following equation:

$$D(t) = d^0_{bud}(t) + d^0_{exp}(t) + d^0_{ring}(t) + d^0_{root}(t) + \left(d^1_{bud}(t) + d^1_{ring}(t) + d^1_{root}(t)\right) \cdot \frac{Q(t)}{D(t)} + d^2_{ring}(t) \cdot \left(\frac{Q(t)}{D(t)}\right)^{\gamma} \quad \text{(C.1)}$$

with $0 < \gamma < 1$. This is the most general expression but in many cases, several factors can be equal to zero.
In this section, we prove that this equation always has a unique strictly positive solution and we describe how to choose a good initial point so that the Newton method converges to this solution. We recall that Newton's method (also called Newton-Raphson's method) for a function $f \in C^2(\mathbb{R})$ consists in building the sequence $(x_n)_{n\in\mathbb{N}}$ defined by the recurrence relation: $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$. For any zero z of f, it exists a neighborhood B_z so that if $x_0 \in B_z$, then $(x_n)_{n\in\mathbb{N}}$ converges to z. For a zero of multiplicity one verifying $f'(z) \neq 0$, the convergence is quadratic.

Case 1: $\gamma = 1$. In that case, we also have $d^2_{ring} = 0$ (see section 3.2). The equation to solve is simply polynomial of degree 2. The only positive solution is:

$$D(t) = \frac{1}{2}\left(d^0_{bud}(t) + d^0_{exp}(t) + d^0_{ring}(t) + d^0_{root}(t)\right) + \frac{1}{2}\sqrt{\left(d^0_{bud}(t) + d^0_{exp}(t) + d^0_{ring}(t) + d^0_{root}(t)\right)^2 + 4\left(d^1_{bud}(t) + d^1_{ring}(t) + d^1_{root}(t)\right)Q(t)} \quad \text{(C.2)}$$

If at least the demand of one compartment is non zero (which is necessary the case if the plant is not dead), then the demand value is strictly positive.

Case 2: $\gamma > 1$**.** Let us define the function $f : \mathbb{R}^+ \to \mathbb{R}$ as:

$$f(D) = D^{\gamma+1} - d_1 D^{\gamma} - d_2 D^{\gamma-1} - d_3 \tag{C.3}$$

where the positive time-dependent coefficients d_1, d_2 and d_3 are deduced from equation C.1 (multiplied by D^{γ}). f is continuous and its values at its domain boundaries are of opposite signs. Therefore the equation $f(D) = 0$ has at least one solution. The first derivative of f can be written:

$$f'(D) = D^{\gamma-2} \cdot \big((\gamma+1)D^2 - d_1\gamma D - d_2(\gamma-1)\big) \tag{C.4}$$

The positive zero of f' is

$$\delta = \frac{\gamma d_1 + \sqrt{\gamma^2 d_1^2 + 4d_2(\gamma^2-1)}}{2(\gamma+1)}$$

Therefore f is decreasing on $[0;\delta]$ and increasing on $[\delta;0]$. Thus we know that the unique positive zero of f is higher than δ and that any initial point higher than δ allows for the Newton method to converge to the solution.

Case 3: $\gamma < 1$**.** Similarly, let us define the function $g : \mathbb{R}^+ \to \mathbb{R}$ as:

$$g(D) = D^2 - e_1 D - e_2 - e_3 D^{1-\gamma} \tag{C.5}$$

where e_1, e_2 and e_3 are positive coefficients deduced from equation C.1 (multiplied by D). With the same arguments as above, $g(D) = 0$ has at least one solution. We have:

$$g'(D) = 2D - e_1 - e_3(1-\gamma)D^{-\gamma} \tag{C.6}$$

The second derivative of g, g'', takes strictly positive values on $\mathbb{R}^+$. Therefore g' is continuously increasing. Let ϵ be the unique zero of g'. g is decreasing on $[0;\epsilon]$ and increasing on $[\epsilon;+\infty[$. Thus the unique positive zero of g is higher than ϵ. An adequate initial point is given by any point x_0 so that $g'(x_0) > 0$.
If $e_1 \geq 1$ and $e_1 \geq e_3$:

$$g'(e_1) \geq e_1 \left(1 - \frac{1-\gamma}{e_1^{\gamma}}\right) > 0$$

If $1 \geq e_1 \geq e_3$:

$$g'(1) = 2 - e_1 - e_3(1-\gamma) > 0$$

The cases where $e_3 \geq e_1$ are symmetrical.
Hence choosing $x_0 = max\{e_1, e_3, 1\}$ as initial point for the Newton algorithm ensures the convergence to the positive solution.

List of Tables

List of Figures

Bibliography

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.

J. Allirand, A. Jullien, A. Fortineau, A. Savin, and B. Ney. Parameterizing a simple model of photosynthetic active radiation absorption by complex aerial structures of oilseed rape resulting from genotype x nitrogen interactions. In *GCIRC: Proceedings of the 12th International Rapeseed Congress, Wuhan, Hubei, China*, pages 261–264. Science Press USA Inc., 2007.

F. Asch, A. Sow, and M. Dingkuhn. Reserve mobilization, dry matter partitioning and specific leaf area in seedlings of african rice cultivars differing in early vigor. *Field Crops Research*, 62(2-3):191–202, 1999.

P. Balandier, A. Lacointe, X. Le Roux, H. Sinoquet, P. Cruiziat, and S. Le Dizès. Simwal: A structural-functional model simulating single walnut tree growth in response to climate and pruning. *Annals of Forest Science*, 57:571–585, 2000.

P. Bancal and F. Soltani. Source-sink partitioning. do we need Münch? *Journal of Experimental Botany*, 53(376):1919–1928, 2002.

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*, 101:1125–1138, 2008.

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

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.

C. Basten, B. Weir, and Z. Zheng. *QTL Cartographer v1.17*. North Carolina State University. ftp://statgen.ncsu.edu/pub/qtlcart, 2005.

A. Beattie, J. Larsen, T. Mechaels, and K. Pauls. Mapping quantitative trait loci for a common bean (*Phaseolus vulgaris* L.) ideotype. *Genome*, 46:411–422, 2003.

M. Beaudet and C. Messier. Growth and morphological responses of yellow birch, sugar maple and beech seedlings growing under a natural light gradient. *Canadian Journal of Forest Resources*, 37:1007–1015, 1998.

F. Berninger, E. Nikinmaa, R. Sievänen, and P. Nygren. Modelling of reserve carbohydrate dynamics, regrowth and nodulation in a N_2-fixing tree managed by periodic prunings. *Plant, Cell and Environment*, 23:1025–1040, 2000.

N. Bertin, H. Gautier, and R. C. Number of cells in tomato fruits depending on fruit position and source-sink balance during plant development. *Plant Growth Regulation*, 0:1–8, 2001.

G. Bilbro and W. Snyder. Optimization with many minima. *Systems, Man and Cybernetics (IEEE Transaction)*, 21(4):840–849, 1991.

A. Bosc. EMILION, a tree functional-structural model: presentation and first application to the analysis of branch carbon balance. *Annals of Forest Science*, 57:555–569, 2000.

D. Boyes, A. Zayed, R. Ascenzi, A. McCaskill, N. Hoffman, K. Davis, and J. Görlach. Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *The Plant Cell*, 13:1499–1510, 2001.

J. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1):25–30, 1965.

R. Brouwer. Distribution of dry matter in the plant. *Neth. J. agric. Sci.*, 10(5):361–376, 1962.

G. Buck-Sorlin. The search for qtl in barley (*Hordeum vulgare* L) using a new mapping population. *Cellular and Molecular Biology Letters*, 7:523–535, 2002.

G. Buck-Sorlin and K. Bachmann. Simulating the morphology of barley spike phenotypes using genotype information. *Agronomie*, 20:691–702, 2000.

M. Canell and R. Dewar. Carbon allocation in trees: a review of concepts for modeling. In *Advances in ecological research*, pages 59–104. Begon, M. and Fitter A.H., academic press edition, 1994.

S. Chapman, M. Cooper, D. Podlich, and G. Hammer. Evaluating plant breeding strategies by simulating gene action and dryland environment effects. *Agronomy Journal*, 95:99–113, 2003.

A. Christophe, V. Letort, I. Hummel, P.-H. Cournède, P. De Reffye, and J. Lecoeur. Model-based analysis of the dynamics of trophic competition during *Arabidopsis thaliana* development. *Functional Plant Biology*, In Press., 2008.

M. Cline. Concepts and terminology of apical dominance. *American Journal of Botany*, 84(9):1064–1069, 1997.

A. Cofino, M. Ivanissevich, and J. Gutirrez. An hybrid evolutive genetic strategy for solving the inverse fractal problem of ifs models. *Lecture Notes in Artificial Intelligence*, 1952:467–476, 2000.

C. Collet, O. Lanter, and M. Pardos. Effect of canopy opening on height and diameter growth in naturally regenerated beech seedlings. *Annals of Forest Science*, 58:127–134, 2001.

B. Comps, B. Thiébaut, G. Barrière, and J. Letouzey. Répartition de la biomasse entre organs végétatifs et reproducteurs chez le hêtre européen (*Fagus sylvatica* L), selon le secteur de la couronne et l'âge des arbres. *Annals of Forest Science*, 51:11–26, 1994.

R. M. Corless and D. E. Jeffrey, D. J.and Knuth. A sequence of series for the Lambert W function. In *International Symposium on Symbolic and Algebraic Computation, Maui, Hawaii.*, pages 197–204. New York: ACM Press, 1997.

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, C. Smith, R. Favreau, Y. Guédon, and T. DeJong. Linking carbon economy and architectural development of peach trees by integrating markovian models into L-PEACH. In P. Prusinkiewicz, J. Hanan, and B. Lane, editors, *Proceedings of the 5th International workshop on Functional-Structural Plant Models (FSPM07)*, page 39. HortResearch (Napier, New-Zealand), 2007.

P. Cournède, M. Kang, A. Mathieu, J. Barczi, H. Yan, B. Hu, and P. de Reffye. Structural factorization of plants to compute their functional and architectural growth. *Simulation*, 82(7):427–438, 2006.

P. 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):1207–1219, 2008.

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.*, pages 11–15. IEEE Computer Society (Los Alamitos, California), 2007.

P. Cruiziat. *Physiologie végétale*. Ecole Centrale Paris, 1991.

P. Cruiziat and M. Tyree. La montée de la sève dans les arbres. *La Recherche*, 21(220): 406–414, 1990.

D. Cvijovic and J. Klinowski. Taboo search, an approach to the multiple-minima problem. *Science*, 267:664–666, 1995.

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, 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, 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 J. Bouchon, P. de Reffye, and D. Barthélémy, editors, *Modélisation et Simulation de l'architecture des végétaux. Colloque INRA-CIRAD Modélisation et Simulation de l'architecture des arbres fruitiers et forestiers, Montpellier*, pages 255–423. INRA Science update, Paris, 1993.

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, 1997.

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. Unpublished, 2006.

P. De Reffye, E. Heuvelink, D. Barthélémy, and C. P.-H. *Biological and mathematical concepts for modelling plant growth and architecture*, page 3120 pp. Jorgensen, D.E. and Fath, B., elsevier edition, 2008.

D. de Vienne. *Les marqueurs moléculaires en génétique et biotechnologies végétales*. INRA Editions, 1998.

C. Deleuze. *Pour une dendrométrie fonctionnelle: essai sur l'intégration de connaissances écophysiologiques dans les modèles de production ligneuse*. PhD thesis, Univeristé Claude Bernard - Lyon I, 1996.

C. Deleuze and F. Houllier. A flexible radial increment taper equation derived from a process-based carbon partitioning model. *Annals of Forest Science*, 59:141–154, 2002.

C. Deleuze and F. Houllier. A transport model for tree ring width. *Silva Fennica*, 31 (3):239–250, 1997.

M. Dingkuhn. Modelling concepts for the phenotypic plasticity of dry matter and nitrogen partitioning in rice. *Agricultural Systems*, 52:383–397, 1996.

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.

D. Doerner. Plant meristems : a ménage à-trois to end it all. *Current Biology*, 11(19): 785–787, 2001.

C. Donald. The breeding of crop ideotypes. *Euphytica*, 17:385–403, 1968.

Q. Dong. *Study on functional-structural model of tomato growth based on accurate radiation transfer*. PhD thesis, China Agricultural University, Beijing, 2006.

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*, 101:1195–1206, 2008.

J. Durand, Y. Gédon, Y. Caraglio, and E. Costes. Analysis of the plant via tree-structured statistical models : the hidden Markov tree models. *New Phytologist*, 166: 813–825, 2005.

J. Farrar. Sink strength: What is it and how do we measure it? a summary. *Plant, Cell and Environment*, 16(9):1045–1046, 1993.

P. Ferraro and C. Godin. A distance measure between plant architectures. *Annals of Forest Sciences*, 57:445–461, 2000.

P. Ferraro, C. Godin, and P. Prusinkiewicz. Toward a quantification of self-similarity in plants. *Fractals*, 13:91–109, 2 2005.

T. Fourcaud, F. Blaise, P. Lac, P. Castéra, and P. de Reffye. Numerical modelling of shape regulation and growth stresses in trees I: An incremental static finite element formulation. *Trees*, 17:23–30, 2003.

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

V. Grange. GreenLab mathematical model behavior analysis on Roux architectural model. Rapport de stage d'option scientifique, LIAMA, 2006.

C. Granier, C. Massonet, O. Turc, B. Muller, K. Chenu, and F. Tardieu. Individual leaf development in *Arabidopsis thaliana* : a stable thermal-time-based programme. *Annals of Botany*, 89:595–604, 2002.

Y. Grossman and T. DeJong. PEACH: a model of reproductive and vegetative growth in peach trees. *Tree Physiology*, 14:329–345, 1994.

Y. Guédon, D. Barthélémy, C. Y., 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. Cournede, 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).*, pages 236–243. 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 *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 and R. Oldeman. *Essai sur l'architecture et la dynamique de croissance des arbres tropicaux*. Masson, Paris, 1970.

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, S. Chapman, E. Van Oosterom, and D. Podlich. Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems. *Australian Journal of Agricultural Research*, 56:947–960, 2005.

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.

A. Haywood. *Growth of advanced European beech trees in the transformation phase in the southern Black Forest*. PhD thesis, Forstwissenschaftlichen Fakultät der Albert-Ludwigs-Universität, Freiburg im Bresigau, 2002.

P. Heuret, D. Barthélémy, Y. Guédon, X. Coulmier, and J. Tancre. Synchronization of growth, branching and flowering processes in the south american tropical tree *Cecropia obtusa* (Cecropiaceae). *American Journal of Botany*, 89(7):1180–1187, 2000.

E. Heuvelink. Dry matter partitioning in a tomato plant: one common assimilate pool? *Journal of Experimental Botany*, 46(289):1025–1033, August 1995.

E. Heuvelink. Re-interpretation of an experiment on the role of assimilate transport-resistance in partitioning in tomato. *Annals of Botany*, 78:467–470, 1996.

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

E. Heuvelink and L. Marcelis. Concepts of modelling carbon allocation among plant 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 9, pages 107–115. Springer, 2007.

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

F. Houllier, J.-M. Leban, and F. Colin. Linking growth modelling to timber quality assessment for Norway spruce. *Forest Ecology and Management*, 74:91–102, 1995.

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

J. Isebrands, G. Host, K. Lenz, G. Wu, and H. Stech. *Hierarchical, parallel computing strategies using component onject model for process modelling responses of forest plantations to interacting multiple stresses*. The HAgue, The Netherlands, 2000.

M. Jeuffroy, B. Ney, and A. Ourry. Integrated physiological and agronomic modelling of N capture and use within the plant. *Journal of Experimental Botany*, 53:809–823, 2002.

C. Jirasek, P. Prusinkiewicz, and B. Moulia. Integrating biomechanics into a developmental plant models expressed using L-systems. In T. Speck, editor, *Proceedings of the 3rd Plant Biomechanics Conference, Freiburg-Badenweiler*, pages 615–624. Verlag, Stuttgart, 2000, 2000.

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

M. Kang, P. de Reffye, J. Barczi, B. 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. Kang, P.-H. Cournède, J.-P. Quadrat, and P. de Reffye. Introduction of a stochastic formal language for the GreenLab organogenesis model and computation of the probability distributions of plant structures. In T. Fourcaud and Z. X.P., editors, *Plant Growth Models and Applications (PMA06, Beijing, China)*. IEEE (Los Alamos, California), 2006a.

M. Kang, E. Heuvelink, and P. de Reffye. Building virtual chrysanthemum based on sink-source relationships: Preliminary results. *Acta Horticulturae*, 718:129–136, 2006b.

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*, In Press, 2007a.

M. Kang, J. Evers, J. Vos, and P. De Reffye. The derivation of sink functions of wheat organs using the GreenLab model. *Annals of Botany*, 101:1099–1108, 2007b.

R. Karwowski and P. Prusinkiewicz. Design and implementation of the L+C modeling language. *Electronic Notes in Theoretical Computer Science*, 86(2):1–19, 2003.

S. Kirkpatrick, C. Gelatt, and M. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.

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. Specifications of morphological models with L-systems and relational growth grammars. *Journal of Interdisciplinary Image Science*, 5, 2007.

W. Kurth. Morphological models of plant growth: Possibilities and ecological relevance. *Ecological Modelling*, 75/76:299–308, 1994.

W. Kurth. Some new formalisms for modelling the interactions between plant architecture, competition and carbon allocation. In *4th workshop on individual-based structural and functional models in ecology*. Wallenfels, Bayreuther forum ökologie, 9 1996.

A. Lacointe. Carbon allocation among tree organs : A review of basic processes and representation in functional-structural models. *Annals of Forest Sciences*, 57:521–533, 2000.

A. Lacointe and N. Donès. Functional-structural modelling using the generic tool PIAF-1: a simulation example on young walnut. In P. Prusinkiewicz, J. Hanan, and B. Lane, editors, *Proceedings of the 5th International workshop on Functional-Structural Plant Models (FSPM07)*, page 22. Hortresearch, Napier, New Zealand, 2007.

A. Lacointe, J. Isebrands, and H. G.E. A new way to account for the effect of source-sink spatial relationships in whole plant carbon allocation models. *Can. J. For. Res.*, 32:1838–1848, 2002.

Lanner. On the intensivity of height growth to spacing. *Forest Ecology and Management*, 13:143–148, 1985.

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. 2007.

X. Le Roux, A. Lacointe, A. Escobar-Gutiérrez, and S. Le Dizès. Carbon-based models of individual tree growth: A critical appraisal. *Annals of Forest Sciences*, 58:469–506, 2001.

J. Lecoeur and B. Ney. Change with time in potential radiation-use efficiency in field pea. *European Journal of Agronomy*, 19:91–105, 2003.

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.*, pages 65–69. IEEE Computer Society (Los Alamitos, California), 2007a.

V. Letort, P. Mahe, P.-H. Cournède, P. de Reffye, and B. Courtois. Optimizing plant growth model parameters for genetic selection based on `QTL` mapping. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, simulation, visualization and their Applications.*, pages 16–21. IEEE Computer Society (Los Alamitos, California), 2007b.

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*). *Functional Plant Biology*, In Press, 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):1243–1254, 2008b.

B. Li. Why is the holistic approach becoming so important in landscape ecology? *Landscape and Urban Planning*, 50:27–41, 2000.

G. Louarn, Q. Dong, Y. Wang, J. Barczi, and P. de Reffye. Parameter stability of the functional-structural model greenlab-tomato as affected by plant density and biomass data acquisition. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, simulation, visualization and their Applications.*, pages 142–148. IEEE Computer Society (Los Alamitos, California), 2007.

D. Luquet, Y. Song, S. Elbelt, D. This, A. Clément-Vidal, C. Périn, D. Fabre, and M. Dingkuhn. Model-assisted physiological analysis of Phyllo, a rice architectural mutant. *Functional Plant Biology*, 34:11–23, 2007.

Y. Ma, M. Wen, B. Li, P. 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)*, pages 188–193. IEEE (Los Alamitos, California), 2006.

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, 2007.

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):1185–1194, 2008.

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, B. Zhang, E. Heuvelink, S. Liu, P. 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 P. Prusinkiewicz, J. Hanan, and B. Lane, editors, *Proceedings of the 5^{th} international workshop on Functional-Structural Plant Models (FSPM07)*, page 5. HortResearch (Napier, New Zealand), 2007.

A. Mathieu, P. Cournède, 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):1233–1242, 2008.

R. McMurtrie, B. Medlyn, and R. Dewar. Increased understanding of nutrient immobilization in soil organic matter is critical for predicting the carbon sink strength of forest ecosystems over the next 100 years. *Tree Physiology*, 21(12-13):831–839, 2001.

P. Minchin. New understanding on phloem physiology and possible consequences to modelling long distance carbon transport. In C. Godin and et al., editors, *4th International Workshop on Functional-Structural Plant Models (FSPM04)*, pages 95–97. CIRAD, Montpellier (France), 2004.

P. Minchin and A. Lacointe. New understanding on phloem physiology and possible consequences for modelling long-distance carbon transport. *New Phytologist*, 166: 771–779, 2005.

P. Minchin, M. Thorpe, and J. Farrar. A simple mechanistic model of phloem transport which explains sink priority. *Journal of Experimental Botany*, 44(262):947–955, 1993.

K. Mitchell. Dynamics and simulated yield of Douglas-fir. *Forest Science (Monography 17)*, 21(4):39p, 1975.

A. Mäkelä. Process-based modelling of tree and stand growth: toward a hierarchical treatment of multiscale processes. *Canadian Journal of Forest Research*, 33:398–409, 2003.

A. Mäkelä. Implications of the pipe model theory on dry matter partitioning and height growth in trees. *J. theor. Biol.*, 123:103–120, 1986.

A. Mäkelä. Modeling structural-functional relationships in whole-tree growth: resource allocation. In R. Dixon, R. Meldahl, G. Ruark, and W. Warren, editors, *Process modelling of forest growth responses to environmental stress*, volume Chapter 7, pages 81–95. Timber Press, Portland, Oregon, 1990.

L. Mündermann, Y. Erasmus, B. Lane, E. Cohen, and P. Prusinkiewicz. Quantitative modeling of Arabidopsis development. *Plant Physiology*, 139:960–968, 2005.

E. Nicolini. Nouvelles observations sur la morphologie des unités de croissance du hêtre(*Fagus Sylvatica* L.). symétrie des pousses, reflet de la vigueur des arbres. *Canadian Journal of Botany*, 78:77–87, 2000.

E. Nicolini. Architecture et gradients morphogénétiques chez de jeunes hêtres (*Fagus Sylvatica* L.) en milieu forestier. *Can. J. Bot.*, 76:1232–1244, 1998.

E. Nicolini and B. Chanson. La pousse courte, un indicateur du degré de maturation chez le hêtre (*Fagus Sylvatica* L.). *Can. J. Bot.*, 77:1539–1550, 1999.

E. Nicolini, B. Chanson, and F. Bonne. Stem growth and epicormic branch formation in understorey beech trees (fagus sylvatica l.). *Annals of Botany*, 87:737–750, 2001.

E. Nikinmaa, H. Hakula, and R. Sievänen. modelling tree growth as a competition between sinks using reaction-transport approach in branched architecture. In C. Godin and et al., editors, *4th International Workshop on Functional-Structural Plant Models (FSPM04)*, pages 216–219. CIRAD, Montpellier (France), 2004.

J. Ottorini, N. Le Goff, and C. Cluzeau. Relationships between crown dimensions and stem development in *Fraxinus excelsior*. *Canadian Journal of Forest Research*, 26: 394–401, 1996.

B. Pallas, G. Louarn, A. Christophe, E. Lebon, and J. Lecoeur. Influence of intrashoot trophic competition on shoot development in two grapevine cultivars (*Vitis vinifera*). *Physiologia Plantarum*, In press, 2008.

J. Patrick. Sink strength: Whole plant considerations. *Plant, Cell and Environment*, 16(9):1019–1020, 1993.

D. Perry. Self-organizing systems across scales. *Tree*, 10(6):241–244, 1995.

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, and E. Nikinmaa. LIGNUM: a model combining the structure and the functioning of trees. *Ecological Modelling*, 108:189–198, 1998.

J. Perttunen, E. Nikinmaa, M. Lechowicz, R. Sievänen, and C. Messier. Application of the functional-structural tree model LIGNUM to sugar maple saplings (*Acer saccharum* Marsh) growing in forest gaps. *Annals of Botany*, 88:471–481, 2001.

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.

S. Pouderoux, C. Deleuze, and J. Dhôte. Analyse du rendement des houppiers dans un essai d'éclaircie de hêtre grâce à un modèle à base écophysiologique. *Ann. For. Sci.*, 58:261–275, 2001.

H. Pretzsch. The single tree-based stand simulator SILVA: construction, application and validation. *Forest Ecology and Management*, 162:3–21, 2002.

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

P. Prusinkiewicz. Modeling of spatial structure and development of plants : a review. *Scientia Horticulturae*, 74:113–149, 1998.

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

P. Prusinkiewicz, P. Allen, A. Escobar-Gutierrez, and T. DeJong. Numerical methods for transport-resistance source-sink allocation models. In J. Vos, L. Marcelis, P. de Visser, P. Struik, and J. Evers, editors, *Functional-structural plant modelling in crop production, Wageningen*, volume Chapter 11, pages 126–140. Springer, 2007.

D. Rasmusson. An evaluation of ideotype breeding. *Crop Science*, 27:1140–1146, 1987.

H. Rauscher, J. Isebrands, G. Host, R. Dickson, D. D.I., C. T.R., and M. D.A. ECOPHYS: An ecophysiological growth process model for juvenile poplar. *Tree Physiology*, 7:255–281, June 1990.

M. Renton, J. Hanan, and K. Burrage. Using the canonical modelling approach to simplify the simulation of function in functional-structural plant models. *New Phytologist*, 166:845–857, 2005.

M. Renton, D. Thornby, and J. Hanan. Canonical modelling: an approach for intermediate-level simulation of carbon allocation in functional-structural plant models. In J. Vos, L. Marcelis, P. de Visser, P. Struik, and J. Evers, editors, *Functional-structural plant modelling in crop production, Wageningen*, volume Chapter 13, pages 153–166. Springer, 2007.

H. Rey. *Utilisation de la modélisation 3D pour l'analyse et la simulation du développement et de la croissance végétative d'une plante de tournesol en conditions environnementales fluctuantes (température et rayonnement)*. PhD thesis, Ecole Nationale Supérieurs Agronomique de Montpellier, 2003.

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.

M. Reynolds, R. Trethowan, M. van Ginkel, and S. Rajaran. Application of physiology in wheat breeding (introduction). In M. Reynolds, J. Ortiz-Monasterio, and A. McNab, editors, *Application of Physiology in Wheat Breeding*, pages 2–10. Mexico, 2007.

J. Ribaut, H. William, M. Khairallah, A. Worland, and D. Hoisington. Genetic basis of physiological traits. In M. A. Reynolds MP, Ortiz-Monasterio JI, editor, *Application of Physiology in Wheat Breeding. Mexico*, pages 29–47, 2001.

J. Ricard. What do we mean by biological complexity? *C.R. Biologies*, 326:133–140, 2003.

S. Sabatier and D. Barthélémy. Growth dynamics and morphology of annual shoots according to their architectural position in young *Cedrus atlantica* (Endl.) Manetti ex Carrière (pinaceae). *Annals of Botany*, 84:387–392, 1999.

K. Sastry, G. Kendall, and D. Goldberg. *Genetic algorithms*, pages 97–125. Burke, E. K. and Kendall, G., springer edition, 2005.

Y. Shi and R. Eberhart. A modified particle swarm optimizer. In K. Belew and B. L.B., editors, *Evolutionary Computation Proceedings (IEEE World Congress on Computational Intelligence)*, pages 69–73. Morgan Kaufmann, 1998.

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

R. Shonkwiler, F. Mendivil, and A. Deliu. Genetic algorithms for the 1-D fractal inverse problem. In K. Belew and B. L.B., editors, *Proceedings of the 4th International Conference on Genetic Algorithms (ICGA), San Diego, CA, USA*, pages 495–501. Morgan Kaufmann, 1991.

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.

F. Sterck and F. Schieving. 3d growth patterns of trees: effects of carbon economy, meristem activity and selection. *Ecological Monographs*, 77:405–420, 2007.

F. Sterck, F. Schieving, A. Lemmens, and T. Pons. Performance of trees in forest canopies: explorations with a bottom-up functional-structural plant growth model. *New Phytologist*, 166:827–843, 2005.

L. Taiz and E. Zeiger. *Plant Physiology.* 4th edition. Sinauer associate, 2006.

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.

B. Thiebaut and S. Puech. Développement du hêtre commun. morphologie et architecture de l'arbre. *Revue Forestière Française*, 56(1):45–58, 1984.

M. Tomita, K. Hashimoto, K. Takahashi, T. Shimizu, y. Matsuzaki, F. Miyoshi, K. Saito, S. Tanida, K. Yugi, J. Venter, and C. Hutchison. E-CELL: software environment for whole cell simulation. *Bioinformatics*, 15:72–84, 1999.

J. Vos, L. Marcelis, and J. Evers. Functional-structural plant modelling in crop production. In J. Vos, L. Marcelis, P. de Visser, P. Struik, and J. Evers, editors, *Functional-structural plant modelling in crop production, Wageningen*, volume Chapter 1. Springer, 2007.

E. Walter and L. Pronzato. *Identification de modèles paramétriques.* Masson, 1994.

I. Wardlaw. The control of carbon partitioning in plants. *New Phytologist*, 116:341–381, 1990.

J. Warren-Wilson. *Ecological data on dry matter production by plants and plant communities*, pages 77–123. Bradley, E.F., and Denmead, O.T., interscience publishers, new york edition, 1967.

P. Wernecke, J. Müller, T. Dornbusch, A. Wernecke, and W. Diepenbrock. The virtual crop-modelling system 'VICA' specified for Barley. In J. Vos, L. Marcelis, P. de Visser, P. Struik, and J. Evers, editors, *Functional-structural plant modelling in crop production, Wageningen*, volume Chapter 5, pages 58–69. Springer, 2007.

J. White. The plant as a metapopulation. *Annu. Rev. Ecol. Syst*, 10:109–145, 1979.

W. White and G. Hoogenboom. Gene-based approaches to crop simulation: past experiences and future opportunities. *Agronomy Journal*, 95:52–64, 2003.

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 (CARI04)*, 3:194–207, 2005.

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.

D. Yang and X. Yang. Studies on biomass and productivity of pinus tabulaeformis plantation in the Wufengshang of Wudu, Gansu province. *Journal of Northwest Normal University Natural Science*, 1:70–73, 2004.

X. Yin, , M. Kropff, and P. Stam. The role of ecophysiological models in QTL analysis: the example of specific leaf area in barley. *Heredity*, 82:415–421, 1999.

X. Yin, S. Chasalow, P. Stam, M. Kropff, C. Dourleijn, I. Bos, and P. Bindraban. Use of component analysis in QTL mapping of complex crop traits: a case study on yield in barley. *Plant Breeding*, 121:314–319, 2002.

X. Yin, P. Stam, M. Kropff, and A. Schapendonk. Crop modeling, QTL mapping, and their complementary role in plant breeding. *Agronomy Journal*, 95:90–98, 2003.

X. Yin, P. Struik, and M. Kropff. Role of crop physiology in predicting gene-to-phenotype relationships. *Trends in Plant Science*, 9(9):426–432, 2004.

P. Zalamea, P. Stevenson, S. Madriñán, P.-M. Aubert, and P. Heuret. Growth pattern and age determination for *Cecropia sciadophylla* (Urticaceae). *American Journal of Botany*, 95:263–271, 2008.

Z. Zhan, Y. Wang, P. de Reffye, B. Wang, and Y. Xiong. Architectural modeling of wheat growth and validation study. In *ASAE Annual International Meeting, Milwaukee, Wisconsin*, July 9-12 2000.

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.

Y. Zhang, D. Reed, P. Cattelino, M. Gale, E. Jones, H. Liechty, and G. Mroz. A process-based growth model for young red pine. *Forest Ecology and Management*, 69: 21–40, 1994.

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.