Optimization and optimal control of plant growth : application of GreenLab model for decision aid in agriculture.

Rui Qi

ÉCOLE CENTRALE DES ARTS ET MANUFACTURES
«ÉCOLE CENTRALE PARIS »

INSTITUTE OF AUTOMATION CHINESE ACADEMY OF SCIENCES

Dissertation

presented by: Rui QI
for the degree of Doctor of Engineering

Specialty: Applied Mathematics, Computer Science

Laboratory: Mathématiques Appliquées aux Systèmes (MAS)
Laboratory: Sino-French Laboratory of Computer Science, Automation and Applied Mathematics (LIAMA)

Optimization and Optimal Control of Plant Growth: Application of GreenLab Model for Decision Aid in Agriculture.

Date of defence : 10/03/2010

Jury:

M. Paul-Henry COURNÈDE	Ecole Centrale Paris
M. Baogang HU	CASIA (Dissertation surpervisor)
M. Sanwen HUANG	Chinese Academy of Agriculture Sciences
M. René LECOUSTRE	CIRAD
M. Xiangdong LEI	Chinese Academy of Forestry (Reviewer)
M. Philippe de REFFYE	CIRAD-INRIA (Dissertation supervisor)
M. Christian SAGUEZ	Ecole Centrale Paris (Chair of the jury)
M. Baogui ZHANG	China Agriculture University (Reviewer)

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction (version française)

1.1.1 Composantes du rendement

Les facteurs explicatifs du rendement peuvent être classifiés en deux groupes : des composantes botaniques et des composantes écologiques. L'aspect botanique renvoie à la description des organes qui composent la plante, tandis que l'aspect écologique renvoie à l'influence des facteurs environnementaux sur la croissance de la plante et donc sur son rendement. Notre objet d'étude, tout au long de ce manuscrit, est l'optimisation du rendement et nous nous intéresserons à plusieurs de ses composantes. Nous allons donc tout d'abord introduire ces facteurs et décrire leur influence sur la croissance de la plante, leurs valeurs et leurs relations.

Composantes botaniques du rendement

Les plantes sont composées de nombreux organes, parmi lesquels les feuilles, les entre-noeuds, les fleurs, les fruits, les graines, les racines (qui étant considérées comme des unités élémentaires dans notre approche), etc. Chaque type d'organe a ses propres fonctions et les interactions entre tous ces organes permettent à la plante de croitre et de se reproduire.
Les feuilles sont les organes les plus variables (Huston and Jeffree [2007]) en termes de forme et de taille entre différentes espèces, et même entre différentes plantes d'une même espèce. Leur principale fonction est la photosynthèse qui permet d'alimenter la croissance de la plante. Contrairement à d'autres organes (par exemple la tige, les racines) qui continuent à croitre aussi longtemps qu'elles reçoivent les ressources nécessaires pour cela (croissance secondaire), les feuilles ont une durée de croissance limitée et finissent par tomber.
Les tiges sont composées d'entre-noeuds. L'une des principales fonctions de la tige est le support mécanique des feuilles, fleurs ou fruits et de les positionner dans l'espace de

façon propice à la réalisation de leurs propres fonctions (interception lumineuse, reproduction, etc). L'architecture des tiges est un facteur crucial pour la quantité de carbone assimilé puisqu'elle détermine la distribution 3D des feuilles et la forme du houppier. De plus, les tiges servent au transport des fluides absorbés par les racines jusqu'aux extrémités du réseau d'axes formant la plante, à travers le xylème, et également au transport des sucres, à travers le phloème. Les tiges produisent également de nouveaux tissus (les cernes), qui jouent un rôle pour la stabilité mécanique de la plante et pour le stockage des nutriments. Le bois est produit par la formation de sylème secondaire, processus qui est observé pour les arbres. Cela permet leur croissance en diamètre. Les fonctions du bois sont le transport de ressources vers les organes en croissance et l'augmentation de diamètre des axes afin de leur permettre de se redresser.

Les fleurs sont des structures reproductives. Elles sont le siège de l'union du sperme mâle avec l'ovule femelle pour produire des graines. Pour cela, deux processus sont successivement mis en jeu : la pollenisation et la fertilisation. L'élément nécessaire à la pollenisation est le pollen, qui doit être transporté d'une anthère vers le stigmate. La fertilisation est définie par le contact du sperme et des ovules. Dans certains cas, les insectes jouent un rôle crucial pour le transport du pollen. Le fait que de nombreuses fleurs attirent seulement certaines espèces d'insectes soulève le problème de la coévolution. L'extinction de l'un des membres de la paire plante-insecte impliquerait presque certainement celle de l'autre.

Les racines sont la partie de la plante ne portant pas de feuilles et sont responsable de l'absorption d'eau et de matières inorganiques. Elles sont aussi responsables de l'ancrage de la plante dans le sol et du stockage de nutriments.

Le fruit est défini par les botanistes comme l'organe de la plante composé des graines et de leur enveloppe (Pennington and Fisher [2009]). Il succède à la fleur par transformation du pistil. Le terme de fruit est couramment employé dans un contexte culinaire, où il désigne des parties comestibles et généralement sucrées de la plante. Tous les fruits botaniques ne sont pas nécessairement des fruits au sens culinaire (par exemple l'aubergine ou le poivron sont considérés comme des légumes ; les grains céraliers sont des fruits particuliers au sens botanique, appelés caryopses) et réciproquement, tous les fruits au sens culinaire ne sont pas nécessairement des fruits au sens botanique (par exemple les pommes ou les fraises sont des faux-fruits : ils sont issus non pas du pistil d'une fleur mais d'autres parties de la fleur ou de la transformation de plusieurs fleurs d'une inflorescence).

En plus de leur importance pour la croissance et la survie de la plante, chacun de ces types d'organes peuvent avoir une valeur commerciale, du moins pour certaines espèces. Les feuilles peuvent être consommées (par exemple la salade, le thé ou le tabac). Les tiges peuvent fournir du sucre (canne à sucre, érable) ou des légumes (pousses de bambou), ou encore du bois de construction. Ces organes représentent également de précieuses ressources pour les médicaments. Les fleurs ont longtemps été utilisées par les hommes pour servir de décoration ou comme source de nourriture

(par exemple, le thé de chrysanthème). Les racines peuvent servir non seulement à se nourrir (par exemple la patate douce ou la betterave) mais également à protéger leur environnement en empêchant l'érosion du sol. Concernant les fruits, plusieurs centaines d'entre eux peuvent être consommés par les hommes. Enfin, les plantes peuvent servir à la production de carburant.

Composantes écologiques du rendement

La croissance d'une plante est le résultat des interactions entre son fond génétique et son environnement (Walter and Schurr [2005]). Les conditions environnementales ont une influence sur la croissance et le développement de la plante soit de manière directe, par l'intermédiaire des conditions physiques comme la lumière, la température, l'humidité, soit de manière indirecte par l'adaptation développementale (Chelle [2005], Walter and Schurr [2005]). La dépendance de la croissance de la plante à son environnement repose sur des processus biologiques qui contrôlent la réponse de la plante à des variables physiques : par exemple la photosynthèse, l'absorption racinaire, l'ouverture des stomates ou le développment de maladies.
Parmi les nombreux facteurs influençant la croissance de la plante, on trouve l'eau, la lumière, la température ou l'humidité, qui sont des facteurs abiotiques, et les ravageurs, les pathogènes, etc, qui sont des facteurs biotiques.
L'énergie lumineuse interceptée est l'un des facteurs les plus importants du bilan énergétique de la plante et de son accumulation de biomasse. Des conditions lumineuses différentes mènent des architectures de plantes très différentes, ce qui influence les fonctions principales du houppier telles que la photosynthèse, la transpiration et les flux d'eau ou de nutriments (Pearcy et al. [2005]).
L'eau est également un important élément de la réaction de photosynthèse. Les changements de la disponibilité en eau du sol entrainent des variations dans le fonctionnement des feuilles et dans le paramètre d'efficience de l'eau (Blum [2005]). Ils affectent aussi les parties souterraines de la plante, c'est-à-dire les racines, des points de vue architectural (longueur et profondeur) et physiologique (masse) (Blum [2005],Walter and Schurr [2005]). Sous l'effet d'un stress hydrique sévère, le rapport entre partie racinaire et aérienne a tendance à augmenter (Trewavas [2003]).
La température est impliquée dans la plupart des processus de croissance de la plante. Les changements de température entrainent des différences marquées des caractéristiques de la plante, comme par exemple la photosynthèse (Bertamini et al. [2005]), la surface foliaire, la vitesse de croissance (Loveys et al. [2002]), la date de floraison (Reeves and Coupland [2000]) ou la germination (Kamizi et al. [2006]). Les conditions de température jouent aussi un rôle pour la localisation géographique des plantes (Bertamini et al. [2005]).
D'une manière générale, des changements dans leurs conditions environnementales impliquent pour les plantes une série de mécanismes en réponse : changements d'architect-

ures (par exemple architecture du houppier (Pearcy et al. [2005], élagage de racines, de feuilles ou de pousses (Trewavas [2003])) et changements physiologiques. L'étude de ces phénomènes relève de nombreux domaines scientifiques (Zhou and Shao [2008]): anatomie, physiologie, biochimie, génétique, développement, évolution et biologie moléculaire, pour n'en citer que quelques uns. Les conditions environnementales influencent le rendement : par exemple, on observe une corrélation négative entre le rendement et la sécheresse (Blum [2005]), un déficit en eau dans le sol réduit la production de la plante (Zhou and Shao [2008]), différentes stratégies d'irrigation conduisent à différents rendements pour le cotton et le maïs simulés grâce au modèle EPIC (Ko et al. [2009]).

Les effets des facteurs environnementaux sur la croissance de la plante peuvent être positifs ou négatifs. Par exemple, si les facteurs environnementaux entrainent une compétition interne au sein de la plante (état de forte compétition trophique), leur effet est alors négatif et peut même conduire au sacrifice de certaines parties de la plante. Un exemple d'effet positif est celui de l'interaction avec certains champignons qui peuvent fournir des suppléments de phosphates aux plantes poussant sur des sols pauvres (Walter and Schurr [2005]). Par ailleurs, l'influence environnementale peut être forte ou insignifiante (Walter and Schurr [2005]): si les facteurs environnementaux changent abruptement, cela peut avoir des conséquences importantes sur la croissance de la plante ; alors qu'inversement, si ces changements se font de manière douce et continue, la plante a les capacités d'ajuster sa réponse aux facteurs correspondants.

On voit donc que la croissance de la plante est le résultat d'interactions entre son génotype et son environnement (Hammer et al. [2002], Yin et al. [2003],Tardieu [2003]). Cela peut se traduire par une compétition entre les différentes composantes botaniques du rendement en termes d'énergie : par exemple, en cas de stress hydrique sévère, le rapport entre les masses souterraines et aériennes augmente (Trewavas [2003]). Une augmentation de la biomasse allouée à un composant du rendement entraine une diminution de la biomasse allouée aux autres composants. Ces composants botaniques peuvent être groupés en deux catégories : les organes sources et les organes puits. Les organes sources sont ceux qui produisent de la biomasse alors que les organes puits sont ceux qui en reçoivent. Les feuilles sont des organes à la fois source et puits alors que les autres types d'organes sont généralement uniquement des puits, mis à part certains cas particuliers (Letort [2008]). Différentes trajectoires de croissance et différents phénotypes résultent ainsi de la compétition entre les puits. Ces puits dépendent a priori du génotype de la plante : sous des conditions environnementales identiques, différents cultivars de tomates ont des traits phénotypiques très différents, par exemple pour leur rendement ou leur hauteur (Eshed and Zamir [1995]). Hallé [2005] mentionne que le rendement de l'Hévéa ou du Manioc (pour le caoutchouc) peut être amélioré par certains types de greffes. Cela aussi est le résultat de la compétition des organes de la plante pour la biomasse. Cette compétition est rarement prise en compte, malgré son importance pour la détermination du rendement final. La Figure 1.2 illustre les effets de la compétition entre les différentes composantes botaniques du rendement sur sa valeur.

Figure 1.1: Illustration des effets de la compétition entre les différentes composantes botaniques du rendement sur sa valeur. Le puits du fruit augmente de 0 jusqu'à 6.

1.1.2 Problématique, objectifs et démarche de la thèse

L'objectif principal de cette thèse est de se baser sur les composantes botaniques du rendement identifiées dans la section ci-dessus pour étudier les possibilités d'amélioration du rendement, en utilisant des méthodes d'optimisation et de contrôle optimal, et en les adaptant pour les appliquer dans le cadre d'un modèle structure-fonction de croissance de plante (FSPM). Les résultats d'optimisation sont utiles pour analyser les mécanismes de croissance et de développement de la plante en termes de dynamiques source-puits. Ces résultats peuvent ainsi fournir des références susceptibles de guider la définition d'idéotypes et des stratégies de sélection génétique ainsi que l'amélioration de modalités de culture.
Comme nous l'avons introduit dans le paragraphe précédent, le rendement est le résultat de processus complexes impliquant des interactions entre de nombreuses composantes biologiques et écologiques. Cependant, dans le cadre de cette thèse, nous allons aborder le problème sous deux angles distincts : tout d'abord, l'optimisation de facteurs endogènes en supposant les conditions environnementales constantes, et deuxièmement le contrôle optimal de facteurs exogènes pour un génotype fixé.
Parmi les facteurs biotiques, nous avons étudié l'influence des interactions entre des insectes ravageurs et la croissance d'une plante-hôte. Comme les ravageurs ont leur propre dynamique de développement, cela nécessite le développement d'un modèle approprié. La difficulté réside en ce que ce modèle doit être compatible avec le modèle de croissance de la plante, en terme d'échelles temporelles et spatiales.
L'étude de l'amélioration du rendement peut se faire selon deux approches principales : par la voie de l'expérimentation ou bien par celle de la modélisation. L'un des principaux inconvénients de l'approche expérimentale est la longue durée de temps nécessaire

pour obtenir des résultats (10 ans pour Dencic [1994] et pour Lauri and Costes [2004], 20 ans pour Peng et al. [2008]). De plus, cela nécessite des ressources en surface, irrigation, temps de travail, etc., d'autant plus importantes que le nombre de plantes cultivées est grand (Dencic [1994]). De plus, les performances obtenues par exemple pour l'amélioration du rendement d'une plante céréalière peuvent être décevantes si l'expérimentation est faite sous certaines conditions environnementales et que celles-ci varient par la suite (Peng et al. [2008]). Du fait de toutes ces raisons, l'approche par modélisation est donc de plus en plus répandue. Les modèles de croissance de plantes émergent en tant qu'outils efficaces pour analyser les trajectoires de croissance des plantes (Tardieu [2003], Herndl et al. [2007], Letort et al. [2008b]). Les modèles peuvent également servir en tant que compléments d'une approche expérimentale, pour optimiser le plan expérimental, et dans tous les cas, permettent d'économiser du temps et de l'argent. Les modèles ne sont cependant évidemment pas une solution universelle à tous les problèmes, ils ont également leurs propres limitations. Tout modèle est basé sur des hypothèses qui conditionnent son application. D'autre part, il est souvent impossible de prendre en compte tous les processus biologiques, ce qui impose de faire des simplifications ou de négliger des interactions. Enfin, les modèles dépendent de l'expérimentation lorsqu'il s'agit de leur paramétrisation ou de leur validation. L'un des domaines pour lequel la modélisation peut apporter un gain important est celui de la sélection variétale (Yin et al. [2003], Cilas et al. [2006]). Parmi les travaux réalisés dans ce domaine, on trouve celui de Cilas et al. [2006] qui se base sur des critères architecturaux pour définir un idéotype et celui de Yin et al. [2003] qui considère des critères physiologiques en utilisant un modèle basé sur les processus. Ils s'accordent cependant à dire que les liens entre architecture et processus physiologiques sont importants (Rasmusson [1987], Kaitaniemi et al. [2000], Sievänen et al. [2000] , Luquet et al. [2006], et Fourcaud et al. [2008]). Dans cette thèse, le problème de l'amélioration du rendement est traité en utilisant un modèle considérant à la fois les aspects architecturaux et physiologiques de la croissance. Ce problème est formulé sous forme de plusieurs problèmes d'optimisation qui sont résolus en utilisant des techniques d'optimisation mathématique. Les difficultés liées à leur résolution sont : la non-convexité, la multimodalité, le fait que l'on ne peut pas différencier analytiquement les fonctions objectifs. Certains problèmes se présentent sous la forme de problèmes multi-objectifs : dans ce cas, la solution optimale est donc un ensemble de solutions possibles formant ce que l'on appelle un *front de Pareto*. Etant données ces contraintes, nous avons choisi d'utiliser des algorithmes populationnels, qui peuvent retourner plus d'une solution à chaque fois et qui ne requièrent pas de différenciation de la fonction objectif. Après comparaison de différents algorithmes de ce type, la solution retenue pour sa meilleure performance sur nos cas-tests est un algorithme populationnel heuristique appelé *Particle Swarm Optimization (PSO)*. Le principe de la méthode et de ses différentes variantes est présentée au chapitre 4, ainsi que ses performances en comparaison avec d'autres algorithmes d'optimisation.

1.1.3 Présentation générale des modèles structure-fonctions de croissance de plantes, des modèles de dynamique des populations et des méthodes d'optimisation appliquées aux modèles de croissance de plantes.

Panorama des modèles structure-fonction de croissance de plantes.

Les modèles structures-fonction de croissance des plantes (FSPM) sont des modèles qui intègrent à la fois des processus physiologiques et le développement architectural d'une plante. Ils représentent la plante comme un ensemble de composants interconnectés (entre-noeuds, feuilles, etc.) qui forment le support pour la modélisation des processus physiologiques (assimilation du carbone, photosynthèse, etc.).

Les FSPMs ont vu leur émergence pour la simulation de l'interaction entre architecture et physiologie de la plante. D'une part, l'architecture de la plante est le résultat à la fois du fonctionnement des méristèmes et de la photosynthèse, ce qui implique des processus complexes avec des interactions environnementales. L'observation de l'architecture permet de retracer la trajectoire de croissance de la plante: par exemple des marqueurs morphologiques identifiés par les botanistes permettent de retrouver le nombre d'entre-noeuds produits à chaque cycle de croissance ou bien des variables comme la durée d'expansion ou de fonctionnement des organes. D'autre part, l'architecture d'une plante a une influence sur de nombreuses fonctions physiologiques et physiques associées à sa croissance, comme par exemple l'interception lumineuse, l'acquisition de carbone, les capacités de compétition, la reproduction, la réponse aux stress, les contraintes hydrauliques pour la conduction d'eau ou de sève, les contraintes biomécaniques, etc. Les FSPMs ont plusieurs domaines d'application. Le plus ambitieux est la prédiction de la croissance de plantes d'une espèce donnée sous des conditions environnementales non testées expérimentalement. Un autre objectif est celui de l'intégration des connaissances biologiques à l'échelle de la plante. Un autre objectif, enfin, est celui de l'optimisation d'itinéraires culturaux sous diverses contraintes, qui nécessitent l'intégration des effets de l'architecture et de la physiologie de la plante, comme c'est le cas pour les exemples suivants. Les attaques par des ravageurs sur des cultures agronomiques ou des arbres entrainent des variations de l'architecture qui à leur tour influencent les possibilités d'attaques futures. En parallèle, pour étudier le comportement des insectes, le couplage de l'architecture et du fonctionnement est nécessaire puisque les insectes se nourrissent de la plante et dépendent donc d'une certaine façon de sa croissance. Une autre application des FSPMs est la mise au point de stratégies d'élagage ou de récolte pour l'horticulture et la sylviculture. Ces traitements modifient par nature l'architecture de la plante et ainsi influencent ses capacités photosynthétiques. De plus, il parait important de considérer l'âge et la position des organes pour décrire leur comportement à une date donnée. Les FSPMs représentent donc une direction de recherche intéressante à développer et à appliquer pour analyser des phénomènes complexes, pour

optimiser les trajectoires culturales ou pour construire des systèmes d'aide à la décision pour optimiser le rendement d'une plantation. Les FSPMs sont aujourd'hui de plus en plus considérés comme des outils prometteurs pour optimiser la sélection génétique avec différents objectifs comme le rendement, la résistance à des ravageurs ou des maladies, la tolérance au stress hydrique, etc. (Hammer et al. [2002], Tardieu [2003], Yin et al. [2004], Hammer et al. [2006]).

La caractéristique distinctve des FSPMs est le fait de lier l'architecture et le fonctionnement. Parmi les FSPMs existants, on trouve différentes méthodes pour effectuer ce couplage. Les FSPMs ont hérité des avancées faites dans le cadre du développement des modèles géométriques, dont le but est une représentation réaliste de la plante, sans généralement inclure de connaissance physiologiques. Ces modèles géométriques ont permis d'introduire des techniques à présent matures comme la simulation de l'architecture des plantes à partir d'approches basées sur des règles. De nombreux FSPMs utilisent donc ces techniques : L-systèmes (Lindenmayer [1984]), Automate double-échelle (Zhao et al. [2003]). La dimension des organes (par exemple la longueur des entre-noeuds, la longueur et la largeur des feuilles) dépend de facteurs environnementaux (température, radiations lumineuses, ...) et des dynamiques source-puits. Le nombre d'organes initiés à chaque intervalle de temps doit être déterminé. Dans de nombreux FSPMs, ce nombre est déterminé indépendamment des conditions physiologiques de la plante. Ce type de modélisation peut être considéré valable pour des plantes agronomiques à croissance déterministe et relativement stable, comme le maïs ou la betterave, mais est une simplification trop forte pour des arbres, pour lesquels il est nécessaire de prendre en compte les phénomènes de rétro-action entre l'organogénèse et le fonctionnement. Une méthode alternative à l'approche déterministe est d'utiliser un modèle stochastique. L'initiation des organes ou la levée de dormance des bourgeons dépend dans ce cas de lois de probabilités définies par les utilisateurs. Ces probabilités peuvent éventuellement être reliées à des facteurs environnementaux.

Pour la partie concernant le fonctionnement, l'un des processus les plus importants à prendre en compte lorsque l'on développe un FSPM est l'allocation de carbone. Selon la classification réalisée par Lacointe [2000], on trouve quatre catégories de méthodes pour représenter ce processus. Historiquement, l'approche développée tout d'abord est la méthode empirique : selon cette méthode, la relation entre les sorties et les entrées du modèle sont basés sur des fonctions paramétriques. Les paramètres de cette fonction n'ont pas de signification biologique et la fonction elle-même est choisie sans prendre en compte les propriétés des mécanismes biologiques sous-jacents de l'objet modélisé. Cette méthode permet de produire des résultats en bon accord avec les données observées mais a un pouvoir prédictif limité, notamment dans des conditions non testées, ce qui limite son intérêt pour des applications. Une approche plus flexible est la méthode basée sur des règles de croissance : pour les allométries des organes, une dimension est exprimée en fonction des autres dimensions de l'organe ; pour l'architecture, le nombre d'organes est déterminé en fonction de facteurs environnementaux. La troisième classe

de méthodes d'allocation du carbone est celle des modèles transport-résistance, qui est dérivée de la théorie des circuits électriques. Chaque organe est considéré comme une résistance électrique, les organes sources représentent les entrées, et ainsi les règles de l'électricité permettent de calculer les flux dans le circuit. La dernière classe correspond à la méthode des échanges source-puits. Les organes qui requièrent de la biomasse sont appelés puits tandis que ceux qui en produisent sont les sources. L'allocation en carbone est alors supposée dépendre du potentiel compétitif d'un organe à importer des assimilats produits par la source, relativement aux autres puits présents. De nombreux FSPMs n'utilisent pas seulement l'une de ces quatre approches mais en intègrent plusieurs à la fois pour simuler différents processus physiologiques (par exemple, approches empiriques, basée sur règle de croissance et source-puits, ou bien empirique, basée sur règle de croissance et transport-résistance).

Les FSPMs développés actuellement ont différents objectifs. Par exemple, Hanan and Hearn [2003] et Renton et al. [2005] étudient plus particulièrement l'initiation des organes et leur distribution. Dans leurs modèles, le nombre d'organes initiés à chaque pas de temps est une fonction des facteurs environnementaux (par exemple le PAR (Radiation Photosyntétiquement Active) dans Renton et al. [2005] et la température dans Hanan and Hearn [2003]) mais l'allocation en carbone n'est pas prise en compte. Le critère d'évaluation de leur modèle est donc l'adéquation entre l'architecture simulée et celle observée ; par exemple, l'adéquation entre la simulation et les observations du nombre d'organes à une certaine position. Le modèle Y_plant (Pearcy et al. [2005]) concerne plus spécialement les effets de l'architecture sur les propriétés photosynthétiques (interception lumineuse, température des feuilles, transpiration) et sur les autres fonctions du houppier (support biomécanique, condutance hydraulique). D'autres modèles, enfin, tendent à représenter à la fois l'architecture, la production de biomasse et son allocation (Eschenbach [2005] and de Reffye et al. [2008]).

Ainsi, selon les contraintes et les objectifs du modélisateurs, différents types de FSPMs peuvent être utilisés. Nous résumons ici les principales caractéristiques de quelques FSPMs.

Quelques problèmes communs peuvent tout d'abord être relevés. Tout d'abord, certains paramètres de ces modèles ne peuvent pas être mesurés directement et sont donc estimés empiriquement ou bien ajustés suivant l'approche "essai-erreur". Certains paramètres peuvent même parfois ne pas être identifiables. Cela rend difficile la paramétrisation des FSPMs pour de nouvelles espèces et limite leur utilisation ou leur efficacité en termes de prédiction. Un autre problème à considérer est celui du pas de temps dans la simulation. S'il est trop petit, les modèles sont informatiquement lourds et la simulation d'une plante prend beaucoup de temps lorsqu'elle est âgée ou avec un grand nombre d'organes. Si au contraire le pas de temps est trop grand, les processus les plus fins ne peuvent être simulés en détails. La Table 1.2 liste les caractéristiques de quelques FSPMs. Au vu des objectifs de notre travail de thèse et des contraintes associées (simulation combinant architecture et physiologie, avec de potentielles interactions avec l'environnement,

simulation pas trop coûteuse de manière à pouvoir la lancer de nombreuses fois dans les procédures d'optimisation), ainsi que des compétences de l'équipe, le modèle GreenLab a été retenu. Par ailleurs, la formulation mathématique de GreenLab sous la forme d'une système dynamique facilite son utilisation au sein des algorithmes d'optimisation présentés dans cette thèse. Cela a aussi permis l'estimation de ses paramètres à partir de données expérimentales pour une large variété d'espèces et sous différentes conditions environnementales. La relative stabilité des paramètres du modèle au cours des saisons et selon les traitements (Ma et al. [2007], Ma et al. [2008]) nous conduit à considérer la possibilité d'établir un lien entre certains des paramètres et les génotypes des espèces considérées (Letort et al. [2008b]), même si la démonstration de l'existence de ces liens réclame encore une somme de travail considérable. Le modèle GreenLab a été appliqué déjà à de nombreuses espèces dans le but de représenter et d'analyser leur croissance : tournesol dans Guo et al. [2003], maïs dans Guo et al. [2006] et Ma et al. [2008], chrysanthème dans Kang et al. [2006], concombre dans Mathieu et al. [2007], tomate dans Dong et al. [2008], blé dans Kang et al. [2008b], pin dans Guo et al. [2008] et Wang et al. [2009], Arabidopsis thaliana dans Letort et al. [2008d], coton dans Zhan et al. [2008] et Li et al. [2009] et hêtre dans Letort et al. [2008a]. Ces divers travaux ont permis de construire les fondations sur lesquelles s'appuira notre travail de thèse.

Table 1.1: Caractéristiques de quelques modèles structure-fonction de croissance des plantes

Modèle	TS[1]	AM[2]	CAM[3]	SD[4]	PP[5]	ID[6]	OL[7]	Est[8]	Ref[9]
LIGNUM	annuel	RB[10]	RB	N	Y[15]	Y	Y	N[16]	Perttunen et al. [1998]
SIMWAL	horaire	PB[13]	TR[11]	Y	Y	Y	Y	N	Balandier et al. [2000]
L-OZlink	journalier	L-système	RB	Y	N	Y	Y	N	Hanan and Hearn [2003]
ALMIS	_[14]	PB	SS[12]	N	Y	Y	Y	N	Eschenbach [2005]
Y_plant	journalier	mesures digitales	RB	N	Y	Y	Y	N	Pearcy et al. [2005]
L-système									
+ Modèle canonique	1/2 année	L-système stochastique	RB	N	N	Y	Y	N	Renton et al. [2005]
ECOPHYS	horaire	Not clear	SS	Y	Y	Y	Y	N	Host et al. [2008]
L-PEACH	journalier	RB	SS	Y	Y	Y	Y	N	Lopez et al. [2008]
GRAAL	journalier	RB	SS	Y	Y	Y	Y	N	Drouet and Pagès [2003]
GreenLab	–	RB	SS	N	Y	Y	Y	Y	de Reffye et al. [2008]

[1] "TS" représente le pas de temps;
[2] "AM" les mécanismes architecturaux;
[3] "CAM" le mécanisme d'allocation du carbone;
[4] "SD" dépendant de l'espèce;
[5] "PP" paramètres physiologiques;
[6] "ID" individu-dépendant;
[7] "OL" échelle de l'organe;
[8] "Est" estimation;
[9] "Ref" référence;
[10] "RB" basé sur des règles;
[11] "TR" transport-résistance;
[12] "SS" dynamique source-puits ;
[13] "PB" probabiliste;
[14] "_" non spécifique;
[15] "Y" oui ;
[16] "N" non;

Présentation des modèles de dynamique des populations

De même que dans le domaine de la biologie végétale, les modèles sont des outils importants pour l'étude des écosystèmes pour anayser les risques et les incertitudes, pour améliorer notre compréhension des processus impliqués et pour aider à la décision dans leur management grâce à l'utilisation de techniques de contrôle optimal (Kropff et al. [1995], Legaspi et al. [1996], Sharov [1996], Buffoni and Gilioli [2003]). Par la simulation, les réponses à des questions ou bien l'évolution des écosystèmes sous des conditions particulières peuvent être facilement et rapidement obtenues : par exemple, la valeur critique des facteurs qui entrainent la prolifération de ravageurs, l'effet de pesticides sur un système biologique, etc. Il s'agit d'une approche complémentaire à l'approche conventionnelle reposant sur des expérimentations, avec les limitations qui lui sont inhérentes (Trichilo and Wilson [1993]). La modélisation de populations d'insectes date d'un siècle environ, avec initialement des modèles simples (Sharov lecture en ligne, chapitre 10.5). Puis le développement de modèles empiriques de régression a suivi (Pinnschmidt et al. [1995] page 195-196: review sur les modèles de régression; Marsh et al. [2000]). Depuis les années 60, ont été développés des modèles, appelés, modèles "life-systems" qui décrivent différents processus écologiques à différents stades de développement de la population (Sharov [1996]). Ces modèles sont plus réalistes que les modèles théoriques et peuvent être utilisés sous une large gamme de conditions environnementales, à la différence des modèles de régression.
Le parasitage des insectes ravageurs par d'autres insectes, appelés auxiliaires, est un phénomène couramment rencontré dans la nature et peut même être favorisé par des interventions humaines dans le cadre de la lutte biologique contre les ravageurs. La prise en compte de l'existence de ces auxiliaires est indispensable pour améliorer le réalisme de la simulation des interactions entre les ravageurs et les plantes (Price et al. [1980]). Dans ce cadre, la plante est l'objet central de l'écosystème et doit être protégée des nombreuses attaques de ravageurs (Lecoustre [1988]). La croissance de la plante est fortement liée, de manière directe ou indirecte, au développement des populations de ravageurs et d'auxiliaires puisque les ravageurs se nourissent des feuilles et que les auxiliaires se nourissent des ravageurs. De plus amples détails sur le rôle de la plante au sein de ces écosystèmes tri-trophiques peuvent être trouvés dans Price et al. [1980]. Cela souligne l'importance de modéliser de façon assez précise la croissance de la plante pour l'étude de ces systèmes. Cependant, la plupart des modèles ravageurs-auxiliaires existants considèrent seulement deux antagonistes et font l'hypothèse de ressources non limitées pour les ravageurs (Legaspi et al. [1996], Nguyen-Huu et al. [2006], Tonnang et al. [2009]) ou considèrent seulement un développement à court terme et local (Buffoni and Gilioli [2003]). Un nombre limité de modèles prennent en compte les trois antagonistes (plantes, ravageurs, auxiliaires) mais la plupart représentent la croissance de la plante sous la forme d'une unique variable (Plant et al. [1985], Brown et al. [2004]) ou comme la simple succession de quelques stades de développement bien identifiés (Gosselke et al. [2001]) ou encore à l'aide d'équations de régression (Marsh et al. [2000]).

Dans tous les cas, les mécanismes de croissance de la plante sont simplifiés à l'extrême et les processus physiologiques sous-jacents ne sont pas représentés (Trichilo and Wilson [1993]). Un autre fait important à prendre en compte est que l'impact des ravageurs et des auxiliaires sur la croissance de la plante dépend de leur stade de développement. De plus, les techniques chimiques de destruction des ravageurs ne tuent en fait que les ravageurs et les auxiliaires adultes, tandis que les oeufs et les individus juvéniles sont protégés par les feuilles. En conséquence, dans cette thèse nous développons un modèle de dynamique des populations qui prend en compte différents processus écologiques à différents stades de développement. Ce modèle de dynamique des populations est lié au modèle structure-fonction de croissance des plantes GreenLab, qui simule la croissance de la plante à l'échelle de l'organe et en intégrant des connaissances physiologiques, bien que sous forme simplifiée. Sur ces bases, nous construisons un modèle tri-trophique, c'est-à-dire qui permette de simuler l'évolution d'un système tri-trophique comprenant la plante, les ravageurs et les auxiliaires. Nous effectuons ensuite une analyse de sensibilité de ce système et nous nous servons de méthodes d'optimisation pour analyser le comportement du système et pour identifier les facteurs pouvant conduire à l'explosion de la population de ravageurs ainsi que les facteurs permettant d'augmenter l'efficacité des techniques habituelles d'élimination des ravageurs. A notre connaissance, le modèle que nous présentons est le seul qui considère la répartition des individus de la population sur les organes de la plante.

L'application des méthode d'optimisation aux modèles de croissance de plantes

Durant les dernières décennies, de nombreux travaux ont été réalisés sur l'optimisation basée sur des modèles de croissance de plantes agronomiques. Des revues sur l'optimisation des facteurs environnementaux peuvent être trouvées dans (van Straten et al. [2000], King and Sigrimis [2001], Ferentinos et al. [2006]). Les différents travaux présentés relèvent du domaine du contrôle optimal. Dans la plupart des cas, il s'agit d'optimiser des variables environnementales comme l'eau (Ho et al. [2004]), la température (Fink [1993], Cerasoli et al. [2009]). Dans d'autres cas, c'est le mode de culture qui est considéré, comme les quantités d'engrais (van Evert et al. [2006]), les stratégies de semis (Prasanna Kumar et al. [2009]), les stratégies d'approvisionnement en eau (Wu et al. [2005]). De nombreux travaux traitent de l'optimisation des conditions sous serre (Linker et al. [1998], Dieleman et al. [2006], van Henten et al. [2006]). En particulier, Morimoto et al. [1993] étudie le contrôle optimal de l'approvisionnement en eau (Morimoto et al. [1995]), de l'humidité (Morimoto et al. [1997]) et de la température (Morimoto et al. [2003]) pour une croissance de plante optimale. Mais le modèle de croissance de plante qu'il utilise est une sorte de "boîte noire" basée sur des réseaux de neurones, sans prise en compte de mécanismes physiologiques.
Concernant l'application de méthodes d'optimisation pour trouver des ideotypes, des modèles de plantes sont utilisées par Haverkort and Grashoff [2004] et Herndl et al.

[2007] pour trouver l'idéotype ayant une production optimale, en fonction de certains paramètres physiologiques. Cependant leurs résultats sont basés sur une approche de type essais - erreurs par la simulation. Habekotté [1997], eux, obtiennent leurs résultats grâce à une analyse de sensibilité.
Plusieurs auteurs étudient des problèmes d'optimisation multi-objectifs (Raju and Kumar [1999], Angelis and Stamatellos [2004], Francisco and Ali [2006], Buddadee et al. [2008]). La plupart de ces travaux s'appliquent aux systèmes agriculturaux et à la logistique. Les objectifs sont l'utilisation du territoire, la production agronomique, l'emploi de la force de travail, la gestion de l'eau, et cela pour différentes plantes agronomiques comme le maïs, le blé, la tomate. Les facteurs considérés sont tous liés aux modes de culture (par exemple, le plan d'irrigation). Les modèles de croissance de plante considérés dans ce cadre sont soit des modèles basés sur processus qui n'intègrent aucune information sur la structure des plantes, soit des modèles empiriques comme les réseaux de neurones. A notre connaissance, les seuls travaux mettant en oeuvre des méthodes d'optimisation appliquées à des modèles structure-fonction sont ceux de Wu et al. [2005] et Letort et al. [2008b], à partir du modèle GreenLab. Wu et al. [2005] recherche la stratégie optimale d'apport des ressources en eau, et Letort et al. [2008b] lie les paramètres de GreenLab à un modèle génétique simplifié et l'utilise pour déterminer les valeurs des paramètres génotypiques qui optimisent certains traits phénotypiques.
Les travaux présentés dans cette thèse se situent à la suite de ceux de Wu et al. [2005]. Il faut noter cependant que certains mécanismes physiologiques sont simulés différemment du fait de l'évolution du modèle GreenLab : par exemple, la production de biomasse et son transport sont représentés dans Wu et al. [2005] sur la base de la méthode des résistances hydrauliques ; dans notre thèse, en revanche, c'est la loi de Beer (McMurtrie [1985]), largement répandue parmi les modèles basés sur les processus, qui est utilisée. Nous avons utilisé la version déterministe du modèle GreenLab, GL1, afin d'éviter une trop grande complexité des facteurs. Ce modèle déterministe est cependant acceptable malgré ses simplifications, puisqu'il a pu être utilisé dans plusieurs études précédentes (par exemple pour le maïs dans Ma et al. [2008], la tomate dans Dong et al. [2008], le hêtre Letort et al. [2008a]).

1.1.4 Plan de la thèse

La thèse est organisée comme suit. Dans une première partie, nous nous intéressons à la modélisation : le modèle GreenLab est présenté dans le chapitre 2 et le modèle de dynamique des populations d'insectes que nous avons développé se trouve au chapitre 3. Dans une seconde partie, nous formulons les problèmes d'optimisation basés sur GreenLab au chapitre 4, et nous présentons les algorithmes d'optimisation et de contrôle optimal utilisés ainsi qu'une étude comparative. Dans une troisième partie, nous présentons les applications réalisées. Tout d'abord, l'analyse des paramètres de GreenLab à travers l'optimisation d'un modèle de Corner virtuel est présentée au chapitre 5. Nous décrivons

ensuite l'analyse de 44 génotypes de tomates sur la base des paramètres estimés pour le modèle GreenLab, au chapitre 6. Puis nous étudions au chapitre 7 des problèmes d'optimisation simples et multi-objectifs sur le maïs. Au chapitre 8, nous traitons le problème de la maximisation de la production de bois, en tenant compte de la qualité du bois récolté par l'introduction de contraintes biomécaniques dans le modèle. Le contrôle optimal de la stratégie de récolte de feuilles est introduit en chapitre 9. Enfin, nous résolvons un problème d'optimisation sur un écosystème consistant en la plante, les insectes ravageurs et les auxilaires au chapitre 10. Nous finissons en donnant nos conclusions et des perspectives de ce travail.

1.2 Introduction (English version)

1.2.1 Yield components

As the optimization object of this thesis is plant yield components, we first introduce their functions during the plant growth, their values and their relationships in this section. Yield components can be classified as botanical yield components and ecological yield components. Botanical yield components are the organs that compose a plant, and ecological yield components are the ones that are related to environmental factors, which affect plant growth behavior.

Botanical yield components

Plants are composed of many parts, e.g. leaves, internodes, flowers, fruits, seeds, roots, which are also considered as elementary organs. Each kind of organs has its own functions that enable a plant to grow and to reproduce (the contents of the organ functions in the following paragraphs are referred from wikipedia).
Leaves are the most variable plant organs (Huston and Jeffree [2007]), which differ in shape, size between different species, and even within individual plants. Through photosynthesis on the basis of specified photosynthetic function of leaves, a plant can have resource to support its growth. Unlike other organs (e.g. stems, roots) which continue to grow as long as they have resource to do so, leaves are limited for growth.
Stems are collection of internodes. One of stem functions is to support leaves, flowers and fruits, and to elevate them to appropriate positions in order to let them be able to function (e.g. light interception, reproduction). The architecture of stems is crucial for the amount of carbon assimilation, as it determines 3D distribution of leaves or crown architecture. In addition, it transports fluids (e.g. water) absorbed by roots to the tips of the axes network, or sugar downwards from leaves in the xylem and phloem. It also produces new living tissues (e.g. rings), which play a mechanical role for stability and provides a place to store nutrients. Wood is produced by the formation of secondary xylem, which is defined specifically for trees. Wood increases

trees in diameter. The functions of wood includes the conduction of nutrient resources to leaves and other growth tissues and enabling woody parts to reach large size or to stand up for themselves.
Flowers are reproductive structures. The primary purpose of a flower, reproduction, determines its biological function of mediating the union of male sperm with female ovum in order to produce seeds. To produce seeds, it requires two successive processes: pollination and fertilization. Fertilization is defined as the joining of the sperm to the ovules. The element and necessary material for the pollination is pollen, moving from the anthers to the stigma. In some cases, one way to propagate pollen is by insects. The fact that many flowers attract only one specific species of insects raises the problem of coevolution. The extinction of either member would mean almost certain extinction of the other member as well.
Roots are part of a plant that bear no leaves and lack of nodes (referred from wikipedia). It is responsible of absorption of water and inorganic nutrients, and of anchoring of the plant body to the ground and storage of food and nutrients.
The term fruit is defined botanically as "the seeds and surrounding tissues of a plant" (Pennington and Fisher [2009]). It results from the transformation of the flower's pistil. The term fruit is commonly used in the context of food preparation, where it defines the parts of certain plants that are edible and sweet in their raw state. Some fruits, in the botanical sense of the term, are not commonly considered as fruits (e.g. eggplants or sweetpepper are considered as vegetables ; cereals are particular fruits in the botanical sense, called caryopses). Reciprocally, some fruits in the common sense are not fruits in the strict botanical sense (e.g. apples or strawberries are called false-fruits: they are not issued from the transformation of a single ovary but from other parts of the flowers or from several flowers of an inflorescence).
Besides the functions of each part for plant survival, each type of organs have their own commercial values for certain species. Leaves are used for food as salad, tea or tobacco. Stems could provide sugar from sugarcane, maple sugar, and vegetables from bamboo shoots. It is also a staple for medicines and wood. Flowers have long been used for humans, mainly to beautify environment but also as a source of food, for example, chrysanthemum tea. Besides as edible food from sweet potato and sugar beet, roots can also protect the environment by holding the soil to prevent soil erosion. Many hundreds of fruits are consummated by human as food, including vegetables (e.g. tomato, cucumber), as well as seeds (e.g. maize, wheat). Besides as fuel, wood has extensively applications in the domain of construction.

Ecological yield components

The general growth pattern of plant organs is fixed due to genetic factors, however, it can be strongly modified by environmental conditions (Walter and Schurr [2005]). Environmental conditions impact on plant growth and development either directly in terms of physical conditions on primary growth processes, such as light, temperature

and humidity, or indirectly, due to developmental adaptation (Chelle [2005], Walter and Schurr [2005]). On the other hand, plant growth and development rely on environmental factors. The dependency relationship between them result from biological processes, responding to physical variables, e.g. photosynthesis, mineral root absorption, stomatal opening or disease development.

Numerous environmental factors affecting plant growth includes water, light, temperature, humidity, etc, which belong to abiotic factors, and pests, pathogen, etc, which belong to biotic factors.

Intercepted light is an important component for plant energy balance and plant biomass accumulation. Different light environmental conditions (e.g. sun or shade) lead to different plant crown architectures, and influence integrated crown functions such as photosynthesis, transpiration and water flow, which are central to the plant growth and development (Pearcy et al. [2005]).

Besides light, water is an important material for photosynthetic reactions. The change of the amount of water in soil results in variation of leaf functioning and water use efficiency (Blum [2005]). Besides physiological functions of upper parts of plants are affected by water condition in soil, underground part of plants, i.e. root, is strongly affected architecturally (length and depth) and physiologically (mass) (Blum [2005],Walter and Schurr [2005]). With more severe water stress, the ratio of root to shoot increases (Trewavas [2003]).

Temperature is involved in most of plant growth processes. Changes in temperature result in markedly different physiological outcomes of plant growth, e.g. photosynthesis, leaf area, growth rate (Loveys et al. [2002]), flowering time (Reeves and Coupland [2000]), and seed germination (Kamizi et al. [2006]).

Therefore, to respond to the changes of environmental conditions, plants have a series of fine mechanisms through architectural change (e.g. crown architectural (Pearcy et al. [2005], removal of root or shoot or leaves (Trewavas [2003])) and the changes of physiological variables. These mechanisms are involved in many aspects (Zhou and Shao [2008]): anatomy, physiology, biochemistry (e.g. proteins), genetics, development, evolution and molecular biology. Hence, environmental conditions will influence plant yield, for example, yield and drought resistance have negative relationship (Blum [2005]), water deficit in soil reduces plant production (Zhou and Shao [2008]), different irrigation strategies lead to different yield of cotton and maize through simulation by using a plant model EPIC (Ko et al. [2009]).

However, the effects of environmental factors in plant growth can be negative or positive. If environmental factors result in competition against plant inner factors related to development and growth, the effect is negative. The effect can be also positive, e.g. fungal provides additional phosphate for the plant in poor soils (Walter and Schurr [2005]). Moreover, the effects can be strong or non-significant (Walter and Schurr [2005]): if environmental factors change abruptly, there is a strong effect on plant growth; if environmental factors change smoothly or slowly, the plant has time to adjust

its behavior to the corresponding environmental factors. Extreme and best situations may occur that there is no changes in plant even though resources are limiting.

From the above description, the plant growth is the result of the interaction between both genotype and environment, which is widely accepted by geneticists and physiologists (Hammer et al. [2002],Yin et al. [2003],Tardieu [2003]). However, we can also find that the botanical yield components in plants compete with each other in terms of energy, for example, with severe water stress, the ratio of root to shoot increases (Trewavas [2003]). Increase in biomass allocated to a botanical yield component results in relative decrease in biomass allocated to others. The botanical yield components (plant organs) can be classified into two categories: source organs and sink organs. The organs that obtain biomass are sink organs, and the organs that are used to produce biomass are source organs. According to the definitions of the source organ and the sink organ, leaves both belong to source organs and sink organs, while the other kinds of organs generally belong to sink organs (they can also be source organs (Letort [2008])). Hence, besides plant gene and environmental conditions, different plant growth behaviors and phenotypes also result from the competition of sources and sinks, from the source-sink relation's point of view. Under the same environmental conditions, different variants of tomato have different phenotypic traits, such as fruit yield, plant height (Eshed and Zamir [1995]). Hallé [2005] mentioned that the yield of cassava for rubber is improved by grafting a small cassava to cassava for rubber. It is also the result of the competition of plant organs for biomass. The competition among plant botanical yield components is not taken into account despite its importance for the determination of the final yield. Fig.1.2 illustrates the effect of the competition of botanical yield components of a plant.

Figure 1.2: Illustration of the competition effect of botanical yield components. The fruit sink value increases from 0.

1.2.2 Objectives, contents and investigation approaches of the thesis

The main objective of the thesis is to improve plant yield of the botanical yield components through optimization and optimal control based on a functional-structural plant growth model (FSPM), called GreenLab, as the botanical yield components are economic valuable introduced above in section 1.2.1. On the basis of optimization results, we investigate mechanisms of plant growth and development mainly in terms of the source-sink dynamics. The optimization results can be considered as references to guide breeding for ideotype and to improve cultivation modes.

To achieve the aims of the thesis, we investigate the effects of the two factors on plant yield separately. First, we optimize endogenous factors given environmental conditions, and then we do optimal control on exogenous environmental factors given plant genotype.

As regards environmental biotic factors, we considered the interactions between insects pests and plant growth. As they have their own development dynamics, it requires an appropriate model of their development. A challenge is that this model must be compatible with the plant model, in terms of spatial and temporal scales.

To improve plant yield, there are two main investigation strategies: experiment based and plant growth model based approaches. The critical drawback of experiment based approach is that the time consumed for experiments is long (10 years needed by Dencic [1994] and by Lauri and Costes [2004], and 20 years by Peng et al. [2008]) and it consumes resources that are limited (field, water, labor) due to the cultivation of thousands of plants (Dencic [1994]). Moreover, the performance of the improved plant might prove disappointing in terms of grain yield compared with the original variety when the environmental conditions vary (Peng et al. [2008]). Nowadays, it becomes widely accepted that plant growth models may provide efficient tools to study plant growth behavior (Tardieu [2003], Herndl et al. [2007], Letort et al. [2008b]), since they can not only complement field experiments, but also save time and resources. However, models are generally not universal and completely accurate. They are all based on some assumptions and made some simplifications for some processes. In addition, it is not possible to take all biological processes into account, which are just needed to be unraveled. Therefore, models are generally required to be verified, which needs the help of experiments. Nevertheless, a lot of researchers dedicate themselves to study similar issues to the improvement of plant yield and plant growth behavior, e.g. ideotype breeding, based on plant models (Yin et al. [2003], Cilas et al. [2006]). Even though Cilas et al. [2006] investigated ideotype breeding from the architectural point of view, and Yin et al. [2003] from the physiological point of view using a process based plant growth model, they all agree that there exist critical relationships between plant architectures and physiological processes during plant growth, with other researchers like Rasmusson [1987], Kaitaniemi et al. [2000], Sievänen et al. [2000] , Luquet et al.

[2006], and Fourcaud et al. [2008]. Therefore, the thesis proposed a new methodology to investigate plant growth behavior: the issue of plant yield improvement is investigated based on the functional-structural plant growth model GreenLab considering both architectural and physiological aspects, and is formulated to optimization problems that are solved by optimization techniques. The problems in the thesis reveal non-convexity and multimodality, and even there is no analytical differentiate expression of objective functions of problems. For special problems like multi-objective problems, optimal solution is a set of solutions forming *Pareto front*. Therefore, a population based optimization algorithm, which can return more than one solution at each time and does not require differentiate information of objective functions, is more suitable for the optimization problems in this thesis. Due to its better performance compared with other heuristic optimization algorithms, all optimization problems are solved by a population-based, heuristic optimization algorithm, namely Particle Swarm Optimization (PSO). The principles of the original PSO and the PSO variants adopted in this thesis are introduced in chapter 4, as well as the performance comparison with other heuristic optimization algorithms.

1.2.3 Review on functional-structural plant growth model, population dynamics model and optimization application on plant models

Review on functional-structural plant growth models

Functional-structural plant growth models (FSPMs) are models that integrate submodels simulating both functional processes and architectures of a plant. They present plants as interconnected components (e.g. leaves, internodes) architecturally and model separately physiological processes involving in plant growth (e.g. carbon assimilation, photosynthesis).

On one hand, plant architecture is the result of both meristem functioning and photosynthesis, which involve complex ecophysiological processes, interacted with environmental conditions. It records growth trajectory, for instance, leaf surface area is the cumulated result of leaf expansion. On the other hand, plant architecture influences many physical, mechanical and physiological functions of plant growth (e.g. light capture, carbon gain, competitive ability, reproduction, high light and temperature stress, hydraulic constraints, biomechanical constraints). Besides prediction of the performance of plant growth under different environmental conditions, the other potential practical uses of plant models requires the integration of architecture with physiological processes, which is introduced in detail as follows. Attacks by pests on crops or trees lead to variation of architectures which in turn influence the possibility of future attack. Meanwhile, to study insect behaviour, coupling of architecture and functioning is also required, as insects feed on plants and rely on plant growth. Application of pruning strategy for

horticulture and forestry modifies plant architecture, and so influences photosynthesis. Moreover, researchers found that a number of architectures could correspond to an identical plant in terms of compartment biomass (i.e. total biomass of a population of same type of organs), which have functionally equivalent architectures (Letort [2008] p.115). In addition, organs with different ages or at different positions may have different growth behaviour at certain time, even though they are the same type. Hence, it becomes interesting research direction to develop and use functional-structural models as tools to help explain complex or unmeasured phenomenon or information, to be applied to crop management and decision support system. Recently, it becomes a popular trend to link plant growth models to genetic models for breeding species with specific objectives (Hammer et al. [2002], Tardieu [2003], Yin et al. [2004], Hammer et al. [2006]) (i.e. high yield, high resistance to pests, tolerance to water stress).

The distinctive characteristic of FSPM is the coupling of plant architecture and functioning. Among the existing FSPMs, different methods are used to perform that coupling. FSPMs have inherited from the researches on geometrical models, which aim at generating virtual realistic images but without being concerned by the physiological knowledge, and which have developed mature techniques such as the rule-based approach. Hence, at the beginning of FSPM development, plant architecture is generated by rule-based approaches, e.g. L-system (Lindenmayer [1984]), Dual-Scale Automaton (Zhao et al. [2003]). The dimension of organs (e.g. length of internodes, length and width of leaves) depends on either environmental factors (e.g. temperature, light irradiance), or source-sink dynamics, or both of them. However, number of organs initiated at each time is determinate. In some FSPMs, the number of organs initiated at each growth stage is modeled independently of the plant physiological state. Even though this kind of mechanisms could be applied to crops (e.g. maize, sugar beet), it will be unrealistic and over simple to apply to trees, which requires involving feedback between architecture and functioning. The inter-mediate approach to have nondeterministic architecture is to use stochastic methods. Whether organs initiate or become dormant or die depends on certain probability, which is defined by users. The probability values are either randomly generated or related to environmental factors. The latter implement feedback between architecture and functioning.

In the functioning part of FSPMs, the most important process is carbon allocation. According to the classification of Lacointe [2000], the approaches of carbon allocation used in the existing FSPMs are classified into four categories. At the beginning of FSPM development, empirical method was widely used. In empirical method, the relation between model inputs and model outputs is constructed by a parametric function, either the parameters of the function have no physical meaning or the function is chosen without taking into account the physical mechanism of the object that is modeled. Empirical method produces very close results with the collected data. However, it cannot be used to predict properties of interest under different environmental conditions from where the data are collected. It is the crucial limitation for its development and appli-

cation. A more flexible approach is growth rule-based method. For allometry of organs, one dimensional property is defined as a function of other dimension properties. For architectural part, the number of organs is a function of environmental factors. The third class of carbon allocation is transport-resistance method. It is derived from electrical circuit theory. Each organ is considered as an electric resistance. It is more applied to simulate hydraulic flow. The last class of the approaches, which involves physiological knowledge, is source-sink dynamic method. The organs that require biomass are sinks and the organs that produce biomass are sources. Assimilate allocation is assumed to depend on the relative ability of different sinks to import available assimilates from the source. Normally, FSPMs do not consist of only one kind of approaches, but integrate several kinds of approaches (e.g. empirical, rule-based and source-sink approaches or empirical, rule-based and transport-resistance approaches) to simulate physiological processes.

Different methods for carbon allocation have different modeling objectives. Hanan and Hearn [2003] and Renton et al. [2005] concerned organ initiation and distribution. In their models, the number of organs initiated at each time is a function of environmental factors (e.g. PAR (Photosynthetically Active Radiation) for Renton et al. [2005] and temperature for Hanan and Hearn [2003]). There is no carbon allocation involved. The criterion to evaluate the simulation result is at the architecture level, e.g. whether the number of organs at certain position is similar with the collected data. The Y_plant model (Pearcy et al. [2005]) specifically concerns the effect of crown architecture on light properties (e.g. light absorption, leaf temperature, transpiration) and the function of crown architecture (biomechanical support, hydraulic conductance). Other models concerns more widely the concept of functioning part, aiming at achieving similarity of both architecture and biomass production and allocation (Eschenbach [2005] and de Reffye et al. [2008]). They dedicate to have more realistic architecture and more similar quantitative information of phenotypic traits (e.g. biomass) for the whole plant and each organ with collected data.

Different kinds of FSPMs are developed to investigate specific objectives. According to specific requirement, we choose the appropriate one. In this section, we try to summarize characteristics, advantages and disadvantages of FSPMs.

In literature, the parameters of most of FSPMs, which cannot be measured directly, are derived empirically or by adjusting using "trial-error" approach. Sometimes, parameters cannot be identified. This shortage limits their efficiency in terms of prediction of yield or other application such as optimization. In addition, the time step of some models is too small that they are computationally expensive when simulating large plant. Table 1.2 lists the features of different FSPMs. According to the objectives of the thesis and its performance of combination of architectural and physiological knowledge, with the potential of interaction with environmental conditions, GreenLab is a prior model to choose in this thesis which aims at enhancing yield and helping ideotype design through optimization and optimal control. Further more, the mathematical formalism

of GreenLab as a dynamic system has allowed the estimation of model parameters from experimental data for a wide range of species and environmental conditions. The relative stability of parameters among seasons and treatments (Ma et al. [2007], Ma et al. [2008]) leads us to consider a possible link of model parameters to the genotype of the species (Letort et al. [2008b]), even though assessing such links would claim a considerable amount of work. Moreover, The successful application of the GreenLab model, particularly the deterministic version of GreenLab to various species of crops and trees (sunflower in Guo et al. [2003], maize in Guo et al. [2006] and Ma et al. [2008], chrysanthemum in Kang et al. [2006], cucumber in Mathieu et al. [2007], tomato in Dong et al. [2008], Wheat in Kang et al. [2008b], pine tree in Guo et al. [2008] and Wang et al. [2009], arabidopsis in Letort et al. [2008d], cotton in Zhan et al. [2008] and Li et al. [2009] and beech tree in Letort et al. [2008a]) revealed that the GreenLab model is reliable to be used to describe the plant growth behavior. It builts the foundation for the work in this thesis. Furthermore, GreenLab's mathematical formulation makes it suitable for optimization problems in the thesis.

Table 1.2: Characteristics of structural-functional plant growth models

Model	TS[1]	AM[2]	CAM[3]	SD[4]	PP[5]	ID[6]	OL[7]	Est[8]	Ref[9]
LIGNUM	yearly	RB[10]	RB	N	Y[15]	Y	Y	N[16]	Perttunen et al. [1998]
SIMWAL	hourly	PB[13]	TR[11]	Y	Y	Y	Y	N	Balandier et al. [2000]
L-OZlink	daily	L-system	RB	Y	N	Y	Y	N	Hanan and Hearn [2003]
ALMIS	_[14]	PB	SS[12]	N	Y	Y	Y	N	Eschenbach [2005]
Y_plant	daily	Digital measurement	RB	N	Y	Y	Y	N	Pearcy et al. [2005]
L-system									
+ Canonical model	Half of year	Stochastic L-system	RB	N	N	Y	Y	N	Renton et al. [2005]
ECOPHYS	hourly	Not clear	SS	Y	Y	Y	Y	N	Host et al. [2008]
L-PEACH	daily	RB	SS	Y	Y	Y	Y	N	Lopez et al. [2008]
GRAAL	daily	RB	SS	Y	Y	Y	Y	N	Drouet and Pagès [2003]
GreenLab	–	RB	SS	N	Y	Y	Y	Y	de Reffye et al. [2008]

[1] "TS" represents time step;
[2] "AM" architectural mechanism;
[3] "CAM" carbon allocation mechanism;
[4] "SD" species-dependent;
[5] "PP" physiological parameters;
[6] "ID" individual dependent;
[7] "OL" organ level;
[8] "Est" estimation;
[9] "Ref" reference;
[10] "RB" rule based;
[11] "TR" transport-resistance;
[12] "SS" source-sink dynamics;
[13] "PB" probability based;
[14] "_" non specific;
[15] "Y" yes ;
[16] "N" no;

Review on population dynamics models

Similar with the role of plant growth models for simulation of plant growth behavior, it is acknowledged that models are as important tools to analyze risk and uncertainty of the complex ecosystem, to increase understanding and to help decision making in pest management through applying control techniques (Kropff et al. [1995], Legaspi et al. [1996], Sharov [1996], Buffoni and Gilioli [2003]). Through simulations, the response to specific conditions is easily and quickly obtained. It is a complementary approach to conventional experiments with which difficulties and limitations to answer complex and specific questions, e.g. critical factors that leads to the pest outbreak, effect of the pesticide on a biological system, are encountered (Trichilo and Wilson [1993]). The methods of insect population modeling date back to a century ago, beginning with theoretical models which is the simplest (Sharov online lecture chapter 10.5). Empirical regression models are followed (Pinnschmidt et al. [1995] page 195-196: review of regression models; Marsh et al. [2000]). From the 1960s on, the models that describe different ecological processes of different development stages of population have been developed, called "life-system" models (Sharov [1996]). It corresponds with ecological reality which is missed in theoretical models, and can be used in different environmental conditions to which regression models are limited. Besides auxiliary insects as parasitoids could be deposited by man-made operations for biological control on insect pests, the fact of insect pests being parasitized naturally is a common phenomenon in reality. Without consideration of this phenomenon, study of the interaction between insect pests and plants can not progress realistically (Price et al. [1980]). In the ecosystem where plant is the kernel objects, plant needs to be protected, as it is often subject to pest attacks (Lecoustre [1988]). Moreover, it affects directly and indirectly the population development of pests and auxiliaries, since pests feed on leaves and auxiliaries live on pests. More details about the role of plant in tri-trophic ecosystems can be found in Price et al. [1980]. Therefore, to simulate the interaction between plant and insect population, plant growth needs to be more accurately described. However, among the existing pest-auxiliary models, most of them only consider two antagonists, either assuming that there are enough resource for pests (Legaspi et al. [1996], Nguyen-Huu et al. [2006], Tonnang et al. [2009] or considering a short-term development for local system Buffoni and Gilioli [2003]. Among the limited number of models that simulate the interaction among plant, pests and auxiliaries, to represent plant growth, most of them either impose a single variable (Plant et al. [1985], Brown et al. [2004]), or use crude plant growth stages (Gosselke et al. [2001]), or use regression equations (Marsh et al. [2000]). Plant growth mechanisms are simplified without physiological processes. Very few models are based on physiological plant growth mechanisms (Trichilo and Wilson [1993]). The impact of pests and auxiliaries on plant growth depends on their development stages. Furthermore, considering chemical techniques of pest management, pesticides only kill pest and auxiliary adults, whereas eggs and juveniles are protected by leaves. Therefore, in this thesis, a population dynamics model, which con-

siders different ecological processes of different development stages, is developed. The population dynamics model is linked with the functional-structural plant growth model GreenLab, which simulates plant growth on the organ level with physiological knowledge, and the tri-trophic (i.e. plant, pests and auxiliaries) ecosystem model is thus developed in this thesis. The tri-trophic ecosystem model simulates the plant growth with consideration of the interaction with biotic environmental factors, i.e. pests and auxiliaries, and simulates the population dynamics. Furthermore, sensitivity analysis method and optimization algorithms are applied to the tri-trophic ecosystem model to analyze the system behavior and to investigate the crucial factors that lead to the outbreak of pest population and the important factors that enhance the efficiency of pest management techniques. The tri-trophic ecosystem model developed in this thesis considered the partition of individuals in the population among plant organs, which is ignored in the existing models so far to the author's knowledge.

Review on optimization investigates based on plant growth models

There have been a lot of works about optimization on crops based on plant growth models in the past decades. Most of the corresponding works are related to optimization of environmental components including water (Ho et al. [2004]), temperature (Fink [1993], Cerasoli et al. [2009]); others related to optimization of cultivation modes, such as fertilizer (van Evert et al. [2006]), seeding strategy (Prasanna Kumar et al. [2009]), water supply strategy (Wu et al. [2005]). Moreover, several works focused on optimizing climate conditions especially in greenhouse (Linker et al. [1998], Dieleman et al. [2006], van Henten et al. [2006]). Morimoto et al. [1993] also did optimal control of water supply (Morimoto et al. [1995]), humidity (Morimoto et al. [1997]) and temperature control (Morimoto et al. [2003]) on plant growth. But the plant growth model he used is a kind of "black-box", which is trained and formed by neural networks, without using physiological mechanisms of plant growth. The corresponding review about optimization on environmental factors can be found in (van Straten et al. [2000], King and Sigrimis [2001], Ferentinos et al. [2006]). Technically and generally speaking, all of these optimization works belong to optimal control. Haverkort and Grashoff [2004] and Herndl et al. [2007] have already used plant models to find ideotypes of plants with the optimum product with respect to physiological parameters. However, the results they found were through trial and error method based on simulations, while the results are obtained using sensitivity analysis in Habekotté [1997]. Several works studied multi-objective optimization problems (Raju and Kumar [1999], Angelis and Stamatellos [2004], Francisco and Ali [2006], Buddadee et al. [2008]). However, these works mainly focused on agricultural systems and logistics. The objectives are land utilization, labor employment, crop production, water management, measurement techniques, for various crops (e.g. maize, wheat, tomato). The factors they considered are all related to cultivation modes (e.g. irrigation planning, water planning). For the works that were based on plant growth models, the plant growth models they used are either

process-based model without consideration of architectural information, or empirical model constructed by neural network for example. The work of optimization on plant growth based on a functional-structural plant growth model in literature to my knowledge is done by Wu et al. [2003], Wu et al. [2005], and Letort et al. [2008b], the model being GreenLab. Wu et al. [2003] optimized the fruit sink factors of maize to maximize the cob weight of maize, Wu et al. [2005] investigated the optimal strategy of water supply, and Letort et al. [2008b] linked GreenLab with a simplified genetic model and investigated the optimal plant with expected traits by optimizing the outputs of the genetic model.
The work in this thesis follows Wu [2005]. However, the physiological mechanisms of GreenLab is changed: in the work of Wu [2005], the physiological processes were modeled on the basis of hydraulic transportation by using electronic theory, the structure of plant architecture being considered as an electrocircuit; in this thesis, the physiological processes are modeled by using beer law theory (McMurtrie [1985]), which is widely accepted and used in well known process-based plant growth model. As the work in this thesis is the primary work of optimization based on GreenLab, all optimizations are based on GL1, to avoid more complex factors. Even though GL1 is the deterministic version (i.e. the number of organs generated at each time is predefined and fixed), its availability and usefulness have been proved by its successful applications to various species of plants as introduced above (for instance, Ma et al. [2008], tomato in Dong et al. [2008]).

1.2.4 Structure of the thesis

The thesis is organized as follows. In the first part, we focus on the description of modeling, consisting of GreenLab in chapter 2 and the insect population dynamics in chapter 3. In the second part, it is the subject to formulate optimization problems based on GreenLab in chapter 4, optimization and optimal control algorithms used in this thesis and the algorithm comparison being involved. Optimization applications are in the third part, beginning with the analysis of GreenLab parameters through optimization on a virtual Corner model in chapter 5. The genetic analysis of 44 genotypes of tomato on the basis of the estimated and optimized GreenLab parameters is studied in chapter 6. In chapter 7, we investigate single objective and multi-objective optimization problems on maize. In chapter 8, the optimization problem of maximization of wood yield with the biomechanics constraint of better wood quality is investigated. Optimal control of leaf harvest is introduced in chapter 9. Optimization on a ecosystem consisting of plant, insect pests and auxiliary insect is described in chapter 10. Finally, conclusions and perspectives are given.

Part I

GreenLab and population dynamics model

Chapter 2

GreenLab model

The definition of modeling objects and of work objectives are crucial to determine mechanisms that have to be included in a model. In this thesis, the objectives of the model on which the work is based are to predict plant growth in a given environment, to interpret phenotypic characteristics of plants based on physiological knowledge and to investigate the effect of cultivation modes and restrictions of certain environmental conditions (e.g. planting density, water stress, insect attack, etc) on plant growth.
The global characteristics and assumptions of GreenLab are as follows:

- it is an individual plant based model;
- it uses the so-called meso-scale, which corresponds to organ level, i.e. at an intermediate level between compartment level (i.e. a population of organs) and microscopic level (e.g. cell, xylem, phylom), to simulate plant growth;
- it outputs architectural and physiological information: number of organs at each position, dimensions of organs (e.g. lengths and diameters of internodes, surface areas of leaves), weights of organs individually, 3D image of plants at any time.
- the elementary units are plant phytomers (or metamers) consisting of organs;
- plant development is generated by a rule-based mechanism, i.e. using a sequence of rules to generate plant topology;
- biomass production and allocation are calculated by a source-sink model;
- carbon assimilation is supposed to be stored in a common pool
- all the organs of the same type have the same growth dynamics, regardless of their positions, if they were initiated at the same time;

Before we introduce the detail principles of the GreenLab model, for the sake of clarity, we define some concepts and vocabularies that will be frequently used in this chapter

here. Phytomers (or metamers) are the elementary units in GreenLab. A phytomer consists of a node to which one or several leaves are attached, a subtending internode, axillary buds at the base of leaves and flower buds if they exist. Plants grow by successive shoots of several phytomers produced by buds. The stem segment developing during an uninterrupted period of growth is called growth unit (GU) and the corresponding time period is called growth cycle (GC). Growth cycle is used as the simulation time step in GreenLab. A growth unit can consist of one metamer (mostly for crops) or several metamers (for trees with rhythmic growth). Different metamers are distinguished by two intrinsic properties: their physiological ages and the physiological ages of their potential axillary buds. Physiological age (PA) is a botanical variable characterizing the morphological differentiation of organs (Barthélémy et al. [1997], Barthélémy and Caraglio [2007]). The metamers belonging to the same growth unit have the same physiological age, while their axillary buds can have different physiological ages. If an axillary axis is rising from a bud that has the same physiological age as its bearing metamer, it is called a reiteration. Phytomers are also characterized by their chronological age (CA) that is the number of growth cycles from their emergence. The emergence sequence of metamers follows a predefined rule, modeled by a dual-scale automaton (Zhao et al. [2003]). This modeling approach of the organogenesis processes is presented in the following paragraph.

2.1 Organogenesis: production of metamers and growth units

The dual-scale automaton (Zhao et al. [2003]) is used to define the rules that control the appearance of metamers in GreenLab, as illustrated in Fig.2.1. The term dual-scale means the microstate scale representing metamers (represented by small ellipse in Fig.2.1) and the macrostate scale representing growth unites (represented by big ellipse in Fig.2.1). The different potential states of the automaton are represented with the rules defining their succession. These rules determine the number of growth cycles before transitions from each state to the following ones (arrows in Fig.2.1 represent the operation of transition). Hence, according to the dual-scale automaton, the number of each kind of metamers at any time can be calculated. In addition to axillary buds, a metamer can possibly bear different types of organs, e.g. blades, petioles and fruits. As the number of organs borne by a metamer can be known through observation of the plant investigated, the total number of organs of each type at any time can also be known. Owing to the use of substructure splitting and factorization techniques (Kang et al. [2008a]), the computing performance of growth simulations was significantly improved (Cournède et al. [2006]).

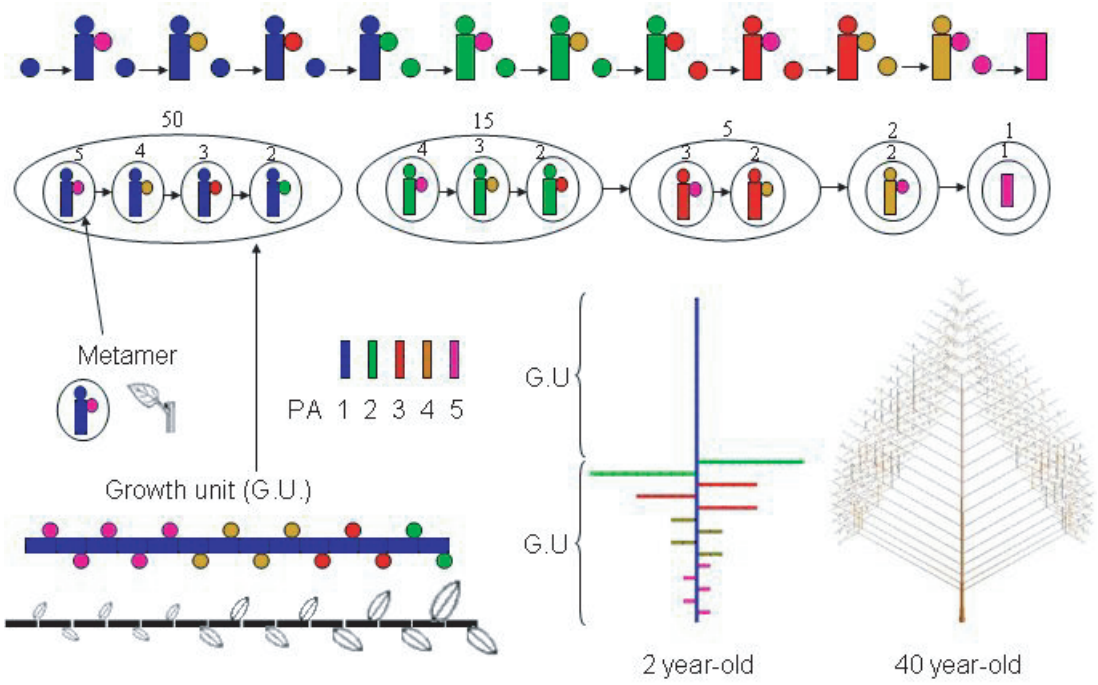


Figure 2.1: Presentation of GreenLab organogenesis mechanism of dual-scale automaton: circles represent buds; rectangles represent metamers; the different physiological ages (PAs) of buds (or metemers) are distinguished by different colors; the number above each small ellipse represents the number of metamers with the corresponding PA inside a growth unit within one growth cycle; the number above each big ellipse represents the number of growth units with the corresponding PA in the whole plant; the arrow represents the transition from each metamer or growth unit to its following one. At the right-bottom of the figure, according to the organogenesis mechanism, the skeleton of a two-year-old tree and the topological structure of a 40-year-old tree without leaves and dead branches are shown.

2.2 Structural factorization

The organogenesis mechanism can also be represented by strings or alphabet (Cournède et al. [2006]) like another famous parallel rewriting grammar named L-system (Prusinkiewicz et al. [1996]). The string or alphabet representation of organogenesis of GreenLab is convenient for numerical computation of plant structure development.
Metamers of chronological age (CA) n and physiological age (PA) p bearing axillary buds of PA q when plant CA is t are represented by $m_{pq}(n,t)$. A bud of PA p when plant CA is t is represented by $s_p(t)$. Hence, a plant of CA t can be described by the set of symbols (Cournède et al. [2006], Letort [2008]), given by

$$I(t) = \{m_{pq}(n,t), p, q \in \mathbb{P}^2, q \geq p, 1 \leq n \leq t\} \cup \{s_p(t), 1 \leq p \leq P_m\} \qquad (2.1)$$

where P_m is the maximal PA; $\mathbb{P}$ is physiological age (PA), varying from 1 to $P_m + 1$. By convention, metamers with axillary buds of PA $P_m + 1$ bear no buds.
Structures can be formed by assembling the symbols. A structure is defined as a set of organs generated from an initial bud. It includes all organs on the main axis and on axes attached to the main axis. Each structure can be decomposed hierarchically from its basal growth unit. For example, a structure can be decomposed to its basal growth unit and a substructure consisting of all metamers generated by the apical bud and by lateral buds of the basal growth unit. Hence, structures can be inferred from each other. The decomposition of a structure (Cournède et al. [2006], Letort [2008]) can be represented as

$$S_p(n,t) = \left[\prod_{p \le q \le P_m} (m_{pq}(n,t))^{u_{pq}(t-n+1)} \left(S_q(n-1,t)\right)^{b_{pq}(t-n+1)} \right] S_p(n-1,t) \qquad (2.2)$$

where $S_p(n,t)$ represents the structure of CA n with a basis metamer of PA p, when plant CA is t; $u_{pq}(n)$ is the number of metamers of PA p with lateral buds of PA q, initiated at growth cycle n; $b_{pq}(n)$ is the number of lateral axes of PA q initiated at growth cycle n on a growth unit of PA p.
By using the same principle as represented by Eq.2.2, $S_q(n-1,t)$ can also be decomposed hierarchically until $S_q(0,t)$, which is the latest bud of PA q initiated at growth cycle t, whose CA is denoted by 0. The PA of the first growth unit is assumed to be 1, and the growth unit thus can be represented by $S_1(1,1)$. The whole plant of CA t can be represented by $S_1(t,t)$ that is the structure with a basal growth unit of PA 1 and of CA t. An example is shown in Fig.2.2 to illustrate plant structure decomposition.
However, the equation 2.2 is not universal for all cases. For example, if the PA of an apical bud is changed, the axis mutates or dies. For this case, the equation 2.2 is not able to describe the structure. The solution is described in detail in (de Reffye et al. [2003]). Another particular case concerns reiteration with finite order. It should introduce brackets or another index to distinguish two structures of the same CA and PA that have different reiteration orders (Letort [2008]).
At a given growth cycle, several identical growth units will appear at different positions on a plant. For example, a structure $S_p(n-1,t)$ can be generated from apical and lateral buds of the basal growth unit of $S_p(n,t)$. More examples are illustrated in Figure 2.2: for instance, two $S_2(1,3)$ are generated from lateral buds of the basal growth unit of $S_1(2,3)$. The fact that identical structures may appear at different positions inspired the concept of substructure factorization (Yan et al. [2004], Kang [2003], Cournède et al. [2006]). Beginning from the youngest structures (i.e. the CA of the basal growth unit is 1), older structures are the result of combining their basal growth unit and pre-computed structures generated by their apical and lateral buds of the basal growth unit. Therefore, each structure type is required to be computed only once. To simulate a plant of CA N, the number of structures that need to be computed is at

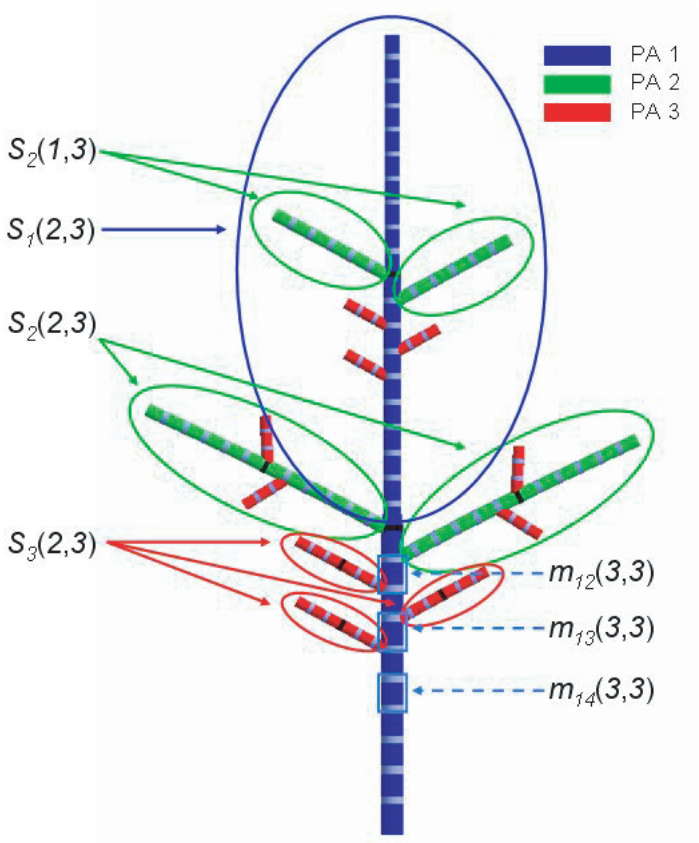


Figure 2.2: Structure decomposition illustration of plant $S_1(3,3)$, where plant CA is 3, the maximal PA (P_m) is 3. The figure is adopted and modified from Letort [2008]. $u_{14}(3) = u_{14}(2) = u_{14}(1) = u_{14}(0) = 5$, $u_{13}(3) = u_{13}(2) = u_{13}(1) = u_{13}(0) = 3$, $u_{12}(3) = u_{12}(2) = u_{12}(1) = u_{12}(0) = 2$, $u_{24}(2) = b_{24}(2) = 4 \cdot 4 = 16$, $u_{24}(1) = b_{24}(1) = 4 \cdot 2 = 8$, $u_{23}(2) = b_{23}(2) = 2 \cdot 4 = 8$, $u_{23}(1) = b_{23}(1) = 2 \cdot 2 = 4$, $u_{34}(2) = b_{34}(2) = 3 \cdot 10 = 30$, $u_{34}(1) = b_{34}(1) = 3 \cdot 3 = 9$.

most $N \times P_m$. Owing to the use of substructure splitting and factorization techniques, compared with the node by node approach for which the time consumed is proportional to the total number of nodes on a plant, the computing performance of growth simulations was significantly improved (Cournède et al. [2006]). However, the substructure splitting technique limits the model flexibility in the sense of reproducing irregular architectures, which are commonly observed for real trees. Due to the requirement of huge and tedious experimental work for precise and complete description of tree structures, exact replication of tree architectures is not realistic. A simplified description of tree architecture is sufficient for tree growth simulation, especially secondary growth simulation (i.e. radial increment) (Guo et al. [2008], Letort et al. [2008a]).

2.3 Functional processes

The functional processes of plant growth are modeled using a source-sink approach in GreenLab. Plant biomass production through photosynthesis is first gathered into a common pool, and then is used for primary growth and for secondary growth of organs. Primary growth means the expansion of organs, i.e. new organ development, leaf surface area expansion, internode expansion in length. While the radial increment of internodes and leaf thickening are called secondary growth. Hence, plant biomass can be separated into two parts: biomass for primary growth and for secondary growth. The quantity of biomass distributed to each type of growth is proportional to the demand of each type of growth. Hence, once the demand of each type of growth is known, the corresponding quantity of biomass is known, as expressed by

$$\begin{aligned} &Q_C(t) = D_C(t)\frac{Q(t)}{D(t)}, \quad C \in \{pg, sg\} \\ &D(t) = D_{pg}(t) + D^{rg}_{sg}(t) + D^{b}_{sg}(t) \end{aligned} \tag{2.3}$$

where $Q(t)$ is the biomass increment of an individual plant at growth cycle t; $D(t)$ is the total biomass demand of a plant at growth cycle t; $Q_C(t)$ represents the biomass of the compartment C; $D_C(t)$ represents the demand of the compartment C for biomass; pg represents primary growth, sg represents secondary growth of internodes with superscript rg (i.e. the radial increment of internodes), and secondary growth of leaves with superscript b (i.e. leaf thickening).

2.3.1 Primary growth

All living organs (blades, petioles, internodes, fruits, and even rings round piths) are sinks among which biomass is distributed according to their sink values. Hence, the total demand of plant for biomass for primary growth at growth cycle t is set to be the sum of sinks of all organs that need expansion except rings and leaf thickening, denoted by $D_{pg}(t)$, given by:

$$D_{pg}(t) = \sum_{o} \sum_{p=1}^{P_m} \sum_{j=1}^{\min(t,t^o_x)} N^o_p(t-j+1)p^o_p(j) \tag{2.4}$$

where $p^o_p(j)$ is the sink value of an organ o of chronological age j of physiological age p for biomass; o represents blade (b), petiole (s), internode (e), female (f) and male (fm) organs; t^o_x is the expansion duration of organ o; $N^o_p(t)$ is the number of organ o of physiological age p initiated at growth cycle t.

The biomass allocated to an organ o of chronological age j of physiological age p at growth cycle t is proportional to its sink value, as detailed in Eq.2.5.

$$\Delta q_p^o(t,j) = p_p^o(j)\frac{Q_{pg}(t)}{D_{pg}(t)} = p_p^o(j)\frac{Q(t)}{D(t)} \tag{2.5}$$

Therefore, the accumulated biomass of an organ o of chronological age i of physiological age p at growth cycle t for primary growth is given by:

$$q_p^o(t,i) = \sum_{j=1}^{i} \Delta q_p^o(t-i+j,j) \tag{2.6}$$

The sink value of organs can be variable during their expansion. The organ sink variation in time is simulated using a Beta function in GreenLab, as expressed by Eq.2.7. One of assumptions of the GreenLab model is that organs of the same kind (leaves, internodes, fruits, etc) have the same temporal dynamics of sink variations, regardless of their physiological ages.

$$\begin{aligned} & f^o(j) = \begin{cases} be^o(j)/M^o & (1 \le j \le t_x^o) \\ 0 & (j > t_x^o) \end{cases} \\ & \text{with} \quad be^o(j) = (j-0.5)^{a^o-1} \cdot (t_x^o - j + 0.5)^{b^o-1} \\ & \text{and} \quad M^o = \max_j \left(be^o(j), \quad 1 \le j \le t_x^o\right) \end{aligned} \tag{2.7}$$

where a^o and b^o are model parameters that are the coefficients of the sink variation function be^o; M^o is a normalization factor. A main advantage of the Beta functions is their flexibility, i.e. using only two parameters, diverse shapes of curves of sink variations can be obtained.
The sink value of an organ o of chronological age j of physiological age p, denoted by $p_p^o(j)$, is $f^o(j)$ multiplied by the sink amplitude (also called sink strength) P_p^o as expressed by Eq.2.8.

$$p_p^o(j) = P_p^o f^o(j) \tag{2.8}$$

2.3.2 Secondary growth

As mentioned at the beginning of section 2.3, the total biomass for secondary growth of either internodes or leaves can be known if the corresponding demand is known. However, as regards the secondary growth of internodes, the biomass for secondary growth is allocated to each metamer for its radial increment: at each growth cycle, a new ring is formed along the stem or branches. The amount of biomass allocated to each metamer to form new rings depends on the positions of metamers in plant structure, as well as the secondary growth of leaves which should be considered individually. Therefore, in GreenLab, the phenomenon of the secondary growth of internodes and

leaves is modeled in two steps: (1) global allocation of plant biomass to the secondary growth compartment and (2) biomass partitioning to each organ individual (internode or leaf). Besides internodes and leaves, other kinds of organs may have secondary growth. According to the observations involved in this thesis, only the secondary growth of internodes and leaves is concerned.

Global allocation of plant biomass to secondary growth

Internode secondary growth
There are three possible modes for calculating the demand of biomass for the secondary growth of internodes (Letort [2008]).

- First mode: *Number of leaves*. In this mode, the demand of biomass for internode secondary growth is proportional to the number of leaves alive. The idea is inspired by the Pipe model theory (Shinozaki et al. [1964]), although not exactly equivalent. More the number of leaves alive is, more the demand of biomass for internode secondary growth (D_{sg}^{rg}) is,

$$D_{sg}^{rg}(t) = \left(P_0^{rg} + P_1^{rg} \frac{Q(t)}{D(t)} \right) N^b(t) \tag{2.9}$$

  where P_0^{rg} is a parameter of constant demand and P_1^{rg} is a mass sink (g^{-1}); $N^b(t)$ is the number of leaves alive at growth cycle t.

- Second mode: *Model Q/D*. In this mode, internode secondary growth is assumed to depend on potential growth activity of a plant, which is reflected by the ratio of plant biomass Q to plant demand D. This ratio is assumed to be an index of the level of trophic competition inside plant: the lowest the ratio is, the highest the level of trophic competition is. The demand of biomass for internode secondary growth for this mode is given by

$$D_{sg}^{rg}(t) = P_0^{rg} + P_1^{rg} \left(\frac{Q(t)}{D(t)} \right)^{\gamma} \tag{2.10}$$

  where γ is a coefficient to control the importance of the ratio Q/D on internode secondary growth.

- Third mode: *Model Q*. The demand of biomass for internode secondary growth is assumed to be proportional to the plant biomass production Q, given by Eq.2.11. The highest the plant biomass production is, the largest amount of biomass is allocated to internode secondary growth (relatively to that allocated to other types of growth, i.e. primary growth or secondary growth of leaves).

$$D_{sg}^{rg}(t) = P_0^{rg} + P_1^{rg} Q(t) \tag{2.11}$$

In model applications which involve internode secondary growth, the different modes for defining the demand for the secondary growth compartment need to be tested with experimental data.
The total amount of biomass allocated to secondary growth of internodes at growth cycle t, denoted by $Q^{rg}_{sg}(t)$, is thus calculated as expressed by Eq.2.12.

$$Q^{rg}_{sg}(t) = D^{rg}_{sg}(t)\frac{Q(t)}{D(t)} \tag{2.12}$$

Leaf secondary growth
For the thickness increment of leaves, we assume that the demand for the secondary growth of leaves for biomass is only dependent on a parameter of constant demand (P^b_{sg}) and the total number of leaves alive (N^b), which is given by

$$D^b_{sg}(t) = P^b_{sg}N^b(t) \tag{2.13}$$

Therefore, the total amount of biomass allocated to secondary growth of leaves at growth cycle t, denoted by $Q^b_{sg}(t)$, is calculated as expressed by Eq.2.14.

$$Q^b_{sg}(t) = D^b_{sg}(t)\frac{Q(t)}{D(t)} \tag{2.14}$$

Biomass allocation to metamers for secondary growth

The total amount of biomass for secondary growth of internodes $Q^{rg}_{sg}(t)$ is then allocated to each metamer depending on its topological position. According to the Pressler law and to the Pipe model theory (Shinozaki et al. [1964]), biomass increment for secondary growth of a metamer depends on the number of active leaves above it in the plant architecture. However, some restrictions to this rule were observed (Deleuze and Houllier [2002]). To overcome these limitations, a mixed approach combining two submodes was proposed, which allowed allocation of the biomass for internode secondary growth both along upward and downward pathways, using a coefficient λ. In the first mode (D^{rg}_{sg1}), metamer demand for secondary growth is determined as a mere sink sub-model, while in the second mode (D^{rg}_{sg2}), the number of active leaves is taken into account, as expressed by Eq.2.15 and Eq.2.16 respectively.

$$D^{rg}_{sg1}(t) = \sum_{p=1}^{P_m}\sum_{j=1}^{t} N^e_p(t-j)\cdot p^{rg}_p \cdot l_p(t-j) \tag{2.15}$$

$$D^{rg}_{sg2}(t) = \sum_{p=1}^{P_m}\sum_{j=1}^{t} NL^b_p(t-j)\cdot p^{rg}_p \cdot l_p(t-j) \tag{2.16}$$

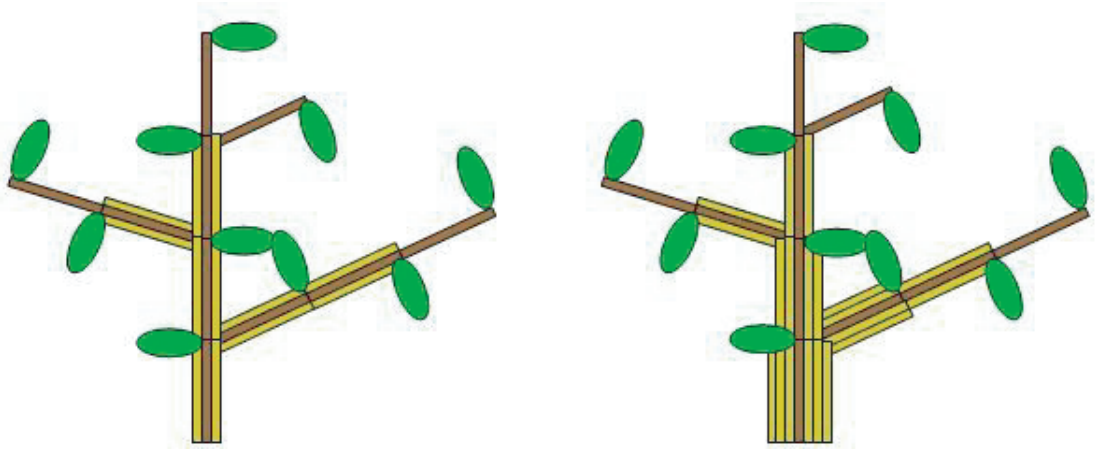


Figure 2.3: Illustration of biomass distribution into metamers with two modes for secondary growth: the left one corresponding to mode D^{rg}_{sg1} as shown in Eq.2.15 and the right one corresponding to mode D^{rg}_{sg2} as shown in Eq.2.16.

In GreenLab, we assume that the organs that develop at the current time have no secondary growth and they do not have contributions for secondary growth of other organs. Hence, only the organs developed in previous times are involved for the calculation of secondary growth. $N^e_p(t-j)$ is the number of metamers (piths) of physiological age p, chronological age j, at growth cycle $t-1$, which initiated at growth cycle $t-j$; $NL^b_p(t-j)$ is the number of living leaves above the metamer of physiological age p, chronological age j, at growth cycle $t-1$; $l_p(t-j)$ is the corresponding metamer length, which is determined by primary growth and related to $q^e_p(t-1,j)$ calculated by Eq.2.6.

In the mixed approach, the biomass allocated to secondary growth of a metamer of physiological age p, chronological age i at growth cycle t is given by

$$\Delta q^{rg}_p(t,i) = \left(\frac{1-\lambda}{D^{rg}_{sg1}(t)} + \frac{\lambda \cdot NL^b_p(t-i)}{D^{rg}_{sg2}(t)} \right) \cdot p^{rg}_p \cdot l_p(t-i) \cdot Q^{rg}_{sg}(t) \qquad (2.17)$$

where λ is a coefficient between 0 and 1.

Fig.2.3 illustrates this modeling approach under the assumption that only one metamer is generated at each growth cycle with a constant length equal to one. According to Eq.2.17, if λ takes null value, only mode D^{rg}_{sg1} takes effect. The total biomass for secondary growth is distributed equally among metamers. If λ is equal to one, only mode D^{rg}_{sg2} takes effect. With this last mode, older metamers get more biomass, as they are generally located at the bottom of the structure and thus below high number of leaves.

Biomass allocation to leaves for secondary growth

The amount of biomass used for each leaf thickening is assumed to be proportional to its surface area, given by Eq.2.18. Similar with the secondary growth of internodes, the leaf thickening is affected at the end of each growth cycle.

$$\Delta q_p^{bsg}(t,j) = \frac{BS_p(t-1,j-1)}{\sum_{i=1}^{P_m}\sum_{k=1}^{ta} N_i^b(t-k)BS_i(t-1,k)} Q_{sg}^b(t) \tag{2.18}$$

where $BS_p(t,j)$ is the surface area of the leaf of chronological age j of physiological age p, when plant age is t; ta is the blade functioning duration.
Hence, for a leaf of chronological age i of physiological age p at growth cycle t, its final biomass, denoted by $q_p^{bt}(t,i)$, is composed by the biomass for primary growth and for the secondary growth, given by

$$q_p^{bt}(t,i) = \sum_{j=1}^{i} \left(\Delta q_p^b(t-i+j,j) + \Delta q_p^{bsg}(t-i+j,j)\right) \tag{2.19}$$

In part III of this thesis, without particular explanation, the secondary growth of leaves is not concerned.

2.3.3 Plant biomass production

GreenLab simulates the plant growth from the seed stage, hence the initial plant biomass is from the seed and the initial organs' mass are driven by the seed mass. And then, at the following growth cycle t, the biomass increment of an individual plant, denoted by $Q(t)$, is calculated by Beer-Lambert's law (McMurtrie [1985]) as expressed by Eq.2.20, which describes the light interception by foliage.

$$\begin{aligned} Q(t) &= E(t)\mu Sp\left(1 - exp\left(-\frac{k}{Sp}S(t)\right)\right) \\ Q(0) &= Q_{seed} \end{aligned} \tag{2.20}$$

where Q_{seed} is the seed biomass; $E(t)$ is a variable representing the plant local environment at growth cycle t; μ represents the light use efficiency; k is a light interception coefficient (Beer-Lambert extinction coefficient); Sp is a characteristic surface area related to plant crown projection for plant modulated by the effects of self-shading and neighbor competition that is related to plant density; $S(t)$ is the total green leaf surface area at growth cycle t; hence, the ratio of $S(t)$ to Sp is a value of leaf area index (LAI) adapted to individual plant.
As plant biomass is determined by foliage area, only leaf expansion of primary growth is involved in the plant biomass production. Hence, according to Eq.2.6, the total green

leaf weight related to foliage area at growth cycle t, denoted by $QB(t)$, is given by,

$$QB(t) = \sum_{p=1}^{P_m} \sum_{i=1}^{\min(t,ta)} \left[N_p^b(n-i+1) \cdot \left(\sum_{j=1}^{\min(i,t_x^b,t)} p_p^b(j) \frac{Q(t-i+j)}{D(t-i+j)} \right) \right] \tag{2.21}$$

As a consequence, the total green leaf surface area at growth cycle t ($S(t)$) is the ratio of the total green leaf weight related to foliage area to the specific leaf weight denoted by slw, given by

$$S(t) = \sum_{p=1}^{P_m} \sum_{i=1}^{\min(t,ta)} \left[N_p^b(n-i+1) \cdot \left(\sum_{j=1}^{\min(i,t_x^b,t)} p_p^b(j) \frac{Q(n-i+j)}{D(n-i+j)} \right) \right] / slw \tag{2.22}$$

Substituting Eq.2.22 to Eq.2.20, we eventually get the complete formulation of biomass production of a plant, given by,

$$\begin{aligned} Q(t) \;=\; & E(t)\mu Sp \left(1 - exp \left(-\frac{k}{Sp \cdot slw} \right. \right. \\ & \left. \left. \sum_{p=1}^{P_m} \sum_{i=1}^{\min(t,ta)} \left[N_p^b(t-i) \cdot \left(\sum_{j=1}^{\min(i,t_x^b,t)} p_p^b(j) \frac{Q(t-i+j-1)}{D(t-i+j-1)} \right) \right] \right) \right) \end{aligned} \tag{2.23}$$

Eq.2.23 illustrates the interaction between organogenesis (defining the variations of organ number) and physiological processes in GreenLab.

2.3.4 Characteristic surface area Sp

In Eq.2.20, the item $\left(1 - exp\left(-\frac{k}{Sp} S(t)\right)\right)$ represents the fraction of intercepted light, denoted by FIL. Hence, the surface area of light interception is $FIL \cdot Sp$ (Cournède et al. [2008]).

For crops, it is reasonable to consider Sp constant. While for trees, it does not remain valid to consider Sp constant. To be valid for both crops and trees, we suppose that Sp is a function of total leaf surface areas of a tree (Cournède et al. [2008]), given by

$$Sp = Sp_0 \left(\frac{S(t)}{Sp_0} \right)^{\tau} \tag{2.24}$$

where Sp_0 and τ are model parameters that are variety dependent. Sp increases with increasing total leaf surface area. The appropriate value of τ should be estimated with observation data (see Letort et al. [2008a] for example).

Sp should be also related to planting density. If plants grow in high densities, the efficient projection area of crown is reduced due to the competition, even though total

leaf surface area is high. Imposing another variable Sd that is the inverse of planting density, we suppose that Sp is an increasing function of Sd, i.e. $Sp = Sd$, if Sd is small representing high planting densities; $Sp = f(Sd)$, if $Sd/Sp \geq 4/\pi$. For detail information of the function f is referred to (Cournède et al. [2008]).
For the theoretical application cases in this thesis, for crops, Sp is set to be constant, which is equal to Sp_0 and $\tau = 0$. For trees, Sp is only a function of Sd, τ being 0.

2.4 GreenLab versions

The number of organs is an important factor, as it is a key variable that drives the plant demand for biomass, organ dimensions and the future growth of a plant. Since GreenLab was first proposed in 1998, three versions have been implemented so far.

2.4.1 Deterministic version of GreenLab

In this earliest and simplest version of GreenLab, the number of metamers and the number of active buds at any growth cycle are predefined by the dual-scale automaton and are constant as introduced above.

2.4.2 Stochastic version of GreenLab

The transition rules between metaters of each physiological age are predefined by the dual-scale automaton. However, this automaton only determines a potential topology. Whether buds grow or not depends on certain probabilities, which determine the final number of metamers and thus plant topology. So far, five kinds of probabilities are concerned (Kang [2003], Letort [2008]). As metamers of different physiological ages have different growth characteristics, each kind of probabilities is a vector of P_m length.

- Survival probability of bud P_c: apical buds could die due to attack of insects for instance. If an apical bud of PA p died with the probability $1 - P_c(p)$, the corresponding axis stops growing. The number of growth cycles of the axis from its initiation to its death follows a truncated geometrical law (P_c, M_a), the elements of vector M_a being the maximal number of growth units in the axis of each PA.

- Growth probability of apical bud P_a: a survived apical bud can be either active to grow or keep dormant till certain growth cycle. The number of growth units follows a binomial law $B(M_a, P_a)$.

- Growth probability of metamer P_i: the number of metamers inside a growth unit is not constant. It follows a binomial law $B(\rho, P_i)$, the elements of vector ρ being the maximal number of metamers per growth unit of each PA.

- Starting probability of axillary bud P_b: axillary buds borne by a metamer can either produce branches or die. Hence, the number of branches at certain place of their main axis is not constant. It follows a binomial law $B(N_B, P_b)$, the elements of N_B being the number of axillary buds borne by a metamer of each PA.
- Appearance probability of fruit P_f: the probability of the abortion of fruits is $1 - P_f$.

Probability values can be either constant or vary in time. The final number of metamers on a plant is the result of compositions of all these probabilities. Even though it is more complex, the mean and variance of number of metamers can be computed using substructure technique (see Kang et al. [2008a] for details). The number of structures that need to be computed is $n \times N \times P_m$, n being the number of samples for each type of structure, N being the plant CA, and P_m being the maximal PA of the plant.

2.4.3 Mechanistic version of GreenLab

In this version, a feedback influence is introduced between the physiological processes (e.g. assimilation) and the plant development (i.e metamer production). The number of metamers appearing at each growth cycle is a linear function of the ratio of plant biomass Q to plant demand D. According to this mechanism, the topological structure of a plant is more flexible. The influence of stressful environmental conditions on plant growth is more realistic. More details are referred to (Mathieu [2006]).

2.5 Calculation of stem mechanical stability

Due to growth phenomena, modeling of plant (especially tree) biomechanical responses to external loads is a complex problem. Stem and branches are inhomogeneous elements composed of central piths surrounded by stacks of rings that were added at each time step. Moreover, mechanical stresses in tree axes can originate from external, e.g. wind or gravity, and internal, i.e. maturation stresses, actions. In addition, the variable wood density and mechanical properties are factors that enhance the complexity to model tree biomechanical responses to external loads. In the thesis, we aim at assessing the influence of growth model parameters on stem mechanical stability by introducing a simple mechanical criterion in GreenLab. For our primary work in the thesis, we calculated the top stem deflection at the end of the plant growth simulation, considering that the total self weight was applied in one stage only on a slightly leaning stem, i.e. buckling. Such a mechanical criterion, which does not account for progressive growth and gravitropism phenomena due to the formation of reaction wood, is commonly used to assess tree stability (see the use of Euler's buckling criteria by Spatz and Bruechert [2000] for instance), and it was considered as sufficient for the theoretical purpose of the work in this thesis.

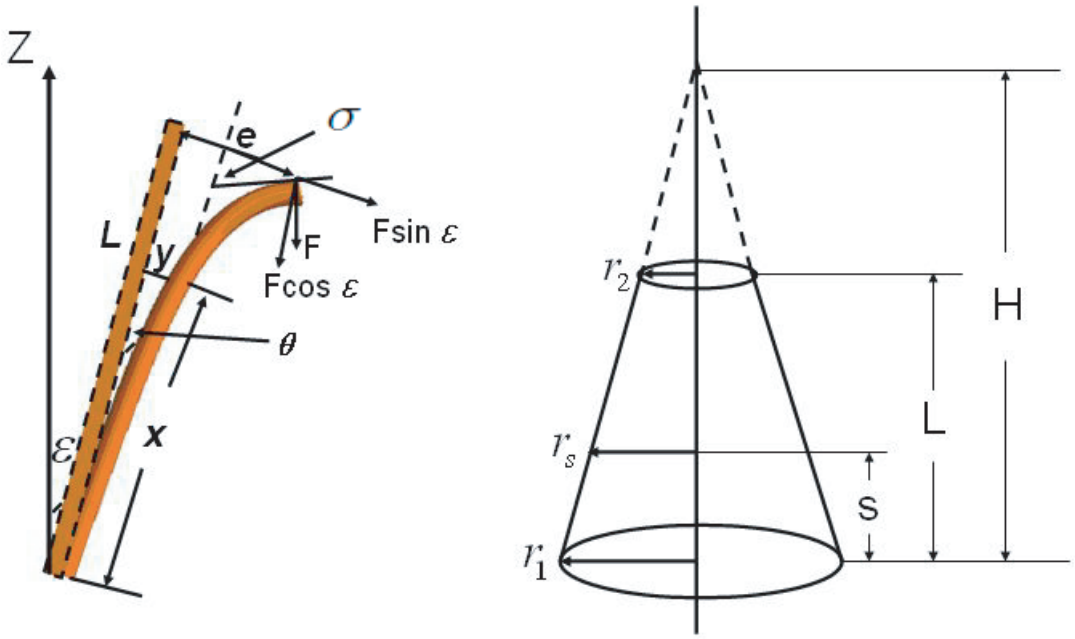


Figure 2.4: Illustration of variable definitions in Eqs. 2.25−2.33. Stem is assumed to be conic as a first approximation. The part of cone with solid line is the tree stem.

Hereafter, we take a tree as example to explain the modeling approach of calculating stem mechanical stability in GreenLab. To simplify the calculations, all the loads are supposed to be integrated and applied at the stem top. Furthermore tree stems were considered slender enough to allow using the beam theory. We used the equilibrium equations adapted to conic beams and proposed by de Reffye [1979] in order to calculate the deflection, i.e. displacement angle (σ), of the stem tip (Eqs.2.25-2.33). Fig.2.4 illustrates the corresponding variable definitions.
When the stem leans with an angle θ with regard to the vertical, the equilibrium of moments of the force at position s from the stem base is given by Eq.2.25 (AMAP tutorial lesson 7 written by de Reffye (not published)):

$$\begin{aligned} E_y I_s \frac{d\theta}{ds} &= F\cos\epsilon\cdot(e-y) + F\sin\epsilon\cdot(L-x) \\ dx &= \cos\theta\cdot ds \\ dy &= \sin\theta\cdot ds \end{aligned} \tag{2.25}$$

where I_s is the secondary moment of the circular cross section at position s from the stem bottom, $I_s = \pi\frac{r_s^4}{4}$, $r_s = r_1\left(1-\frac{s}{H}\right)$, describing the property of the shape and being used to predict the resistance to bending and deflection; r_1 is the radius at the bottom of the cone, and r_s is the radius at position s; H is the length of the cone, which equals to $L(r_1/(r_1-r_2))$; L is the length of the stem, which is determined by q_p^e as expressed

by Eq.2.6; r_2 is the radius at the top of the stem; E_y is structural Young's modulus, a measure of the stiffness of an elastic material, less Young's modulus, less stiffness; θ is the displacement angle with regard to the stem without load; F is the force applied on the stem top; ϵ is an initial displacement angle at the stem bottom in the vertical plane; x is the projection distance to the plane of the stem without load; y is the deflection distance which is vertical to the stem without load; e is the final deflection distance which is vertical to the stem without load.
We differentiate Eq.2.25 with respect to s,

$$E_y I_s \frac{d^2\theta}{ds^2} = F\cos\epsilon\cdot(-\sin\theta) + F\sin\epsilon\cdot(-\cos\theta) \tag{2.26}$$

We multiply Eq.2.26 by $d\theta/ds$ and integrate it with θ that varies from 0 to σ. The result is obtained as expressed by Eq.2.27.

$$\frac{1}{2}E_y I_s\left(\frac{d\theta}{ds}\right)^2 = F\cos\epsilon\,(\cos\theta-\cos\sigma) - F\sin\epsilon\,(\sin\theta-\sin\sigma) \tag{2.27}$$

Suppose

$$K_s = \sqrt{2\frac{F}{E_y I_s}} = \sqrt{\frac{2\dfrac{F}{E_y\pi\dfrac{r_1^4\left(1-\dfrac{s}{H}\right)^4}{4}}}} = \sqrt{\frac{2F}{E_y I_1}}\frac{1}{\left(1-\dfrac{s}{H}\right)^2} = K_0\frac{1}{\left(1-\dfrac{s}{H}\right)^2} \tag{2.28}$$

where I_1 is the second moment of the circular cross section at the stem base $\pi r_1^4/4$.
Substituting Eq.2.28 to Eq.2.27, we get Eq.2.29,

$$K_0\frac{ds}{\left(1-\dfrac{s}{H}\right)^2} = \frac{d\theta}{\sqrt{\cos(\theta+\epsilon)-\cos(\sigma+\epsilon)}} \tag{2.29}$$

The final shape of the stem is given by integrating Eq.2.29, as expressed by Eq.2.30.

$$\int\frac{1}{\left(1-\dfrac{s}{H}\right)^2}ds = L = \int_0^\sigma\frac{d\theta}{K_0\sqrt{\cos(\theta+\epsilon)-\cos(\sigma+\epsilon)}} \tag{2.30}$$

When θ reaches the value σ, the sum of the ds equals to the length of the beam L.
Solving Eq.2.30, the final displacement angle of the stem σ is as Eq.2.31-2.33 (AMAP tutorial lesson 7 written by de Reffye (not published)).

$$\sigma = \frac{sin(\epsilon)\cdot(1-\cos(g))}{\cos(\epsilon)\cdot\cos(g)} \tag{2.31}$$

$$g = \sqrt{\frac{2F}{E_y I_1}}\cdot H\cdot\sqrt{|\cos(\epsilon)|} \tag{2.32}$$

$$F = 9.8\cdot\left(\pi\cdot d\cdot L\cdot\frac{r_1^2+r_2(r_1+r_2)}{3}+\sum_{t=1}^{n}Q(t)\right) \tag{2.33}$$

where I_1, r_1, r_2 are determined by the summation of q_p^i and q_p^{rg} shown in Eq.2.6 and Eq.2.17; d is the density of the stem; σ is the final rotation angle at the top of the deflected stem when a small initial leaning angle ϵ is applied to the tree.

2.6 Summary of the GreenLab model

GreenLab is discretized in both spatial and temporal scales. In spatial scale, the discretized unit is metamers which a growth unit is composed of. In temporal scale, the discretized unit is growth cycle. As the calculation of organogenesis and physiological processes are synchronous as expressed by Eq.2.20, GreenLab can be considered as a dynamic discrete model. Techniques for dynamic discrete system in the domain of optimization and optimal control can be applied on GreenLab.
GreenLab is an intermediate functional-structural model, which simulates ecophysiological processes of plant growth taking into account the dynamics of organogenesis. It is an organ based model at individual level, which can output geometrical dimensions and biomass of organs at each position of a plant. It helps investigate the effect of the change of local environmental conditions on plant growth, for instance, insect attack on leaves. Moreover, GreenLab simulates ecophysiological processes using a source-sink model. It makes the model possible to interpret phenotypic characteristics from the physiological, especially source-sink, point of view.

Chapter 3

Interaction of GreenLab with population dynamics model of insect

3.1 Population dynamics model – generic characteristics

Besides endogenous factors of plants, plant growth behavior is affected by external environmental conditions. It is widely accepted by biologists, physiologists, geneticists, agronomists that morphogenesis and architectures of plant are the results of genotype x environment (G x E) (Klèová et al. [2004], Dingkuhn et al. [2005]). One of external environmental conditions that have negative effects on plant growth is insect pest attack. As damages of plant resulted from insect pest attack can be very significant, integrated pest management has been developed after World War II (Dent [1995]). In order to analyze factors that result in the outbreak of insect pest population, and to analyze crucial factors that enhance the efficiency of insect pest management techniques, we develop a insect population dynamics model and link it to GreenLab. In this thesis, we take insect pests and auxiliary insects of oil palm (*Elaeis guineensis*) as example to model the population dynamics. For the sake of clarity, we abbreviate insect pest to pest and abbreviate auxiliary insect to auxiliary.

The population dynamics of pests and auxiliaries are simulated using a dynamic discrete model. We separate the whole population dynamics of pests into four stages: egg, larva, nymph and adult. Each stage has its own impact on plant growth: eggs occupy leaf area; larvae dig galleries on leaf areas and make that part useless for photosynthesis; nymphs are transition form from larvae to adults; adults feed on leaves and lay eggs on leaf areas. As it is difficult to distinguish larvae and nymphs and the longevity of nymphs is very short, larva and nymph are lumped to an identical stage in the model, called juvenile. The auxiliary population dynamics are similar with pests, with also

three stages: egg, juvenile and adult. Auxiliaries feed on either pest eggs or juveniles according to the species of auxiliaries, and pest eggs or juveniles die due to parasitism. Population dynamics of pests and auxiliaries share the same development processes, e.g. age distribution of the initial population, female adults laying eggs, parasitism of hosts, stage growth (i.e. population development within stages) and stage transition (i.e. population development from one stage to another). Hence, they are modeled in a generic frame. Because of different food they feed on, specific items for pests and auxiliaries will be introduced after the description of the generic frame.
Before we introduce the principles of the population dynamics modes, the assumptions on which the insect population dynamics model is based are summarized as follows:

- larvae and nymphs of insects are not distinguished in the model, and are lumped to an identical stage, called juvenile;
- female adults of pests lay eggs only on the green leaf area, which can produce biomass through photosynthesis;
- pest eggs do not destroy leaf area. They occupy certain area on the leaf area, in order to assure that there are enough resources for evolved juveniles to survive under the ideal condition;
- pest juveniles and adults uniformly destroy leaf area during their whole life;
- pest eggs and juveniles die within one time step once they are parasitized by auxiliaries, while auxiliaries survive;
- pests and auxiliaries share the same development processes.

3.1.1 Age distribution of the initial population

Normally, the population of pests or auxiliaries consists of adults of different ages. The age distribution of adults in the initial population is modeled by a discretized Beta distribution, given by Eq.3.1. Various age distributions of adults can be obtained by changing the coefficients ai and bi.

$$\begin{aligned} F0\left(i,T\right) &= \left(g\left(i\right)/S\right)\left|F0(T)\right|, \quad i=1,2,\ldots,t_{ad} \\ g\left(i\right) &= \left(\left(i-0.5\right)/t_{ad}\right)^{ai-1}\left(1-\left(i-0.5\right)/t_{ad}\right)^{bi-1} \\ S &= \sum_{i=1}^{t_{ad}} g\left(i\right) \end{aligned} \tag{3.1}$$

where $F0(i,T)$ is the number of adults of age i in the initial population; T is the coming time of pests or auxiliaries; $|F0(T)|$ is the initial population size; the notation $||$ represents the summation of all elements in the vector; t_{ad} is maximal life span for adults; g is discretized Beta function; S is the normalization factor.

3.1.2 Egg laying

The number of eggs laid by female adults depends on their reproduction rate.
Reproduction rate
The ability of laying eggs for female adults varies with their ages. The number of eggs laid by female adults of different ages is also modeled by a discretized Beta distribution, which has the same formula as for the age distribution of the initial population (Eq.3.1) except that the coefficients are denoted by ar and br.
Suppose that the fecundity per female adult (i.e. the total number of eggs laid per female adult during its whole life) is M, the number of eggs laid per female adult of age i, denoted by $G(i)$, is given by Eq.3.2.

$$G\left(i\right) = \left(g\left(i\right)/S\right)M, \quad i = 1, 2, \ldots, t_{ad} \tag{3.2}$$

Egg laying
Suppose that the number of adults of age i at time t is denoted by $NT_{ad}(i,t)$ and the sex-ratio is 0.5, the number of eggs laid by the corresponding female adults is $NT_{ad}(i,t)/2 \cdot G(i)$. Hence, the total number of eggs laid by female adults in the population at time t, denoted by $N_{eg}^{'}(t)$, is given by Eq.3.3.

$$N_{eg}^{'}\left(t\right) = \sum_{i=1}^{t_{ad}} \frac{NT_{ad}\left(i,t\right)}{2} G\left(i\right), \quad t \geq T \tag{3.3}$$

3.1.3 Parasitism of hosts

The hosts of pests are leaves, and the hosts of parasitoids are pest eggs (or juveniles). The model assumes that the number of pest eggs laid on leaves or auxiliary eggs parasitizing pest eggs (or juveniles) follows a Poisson distribution, given by Eq.3.4.

$$P_{he}(i) = \left(\eta^i/i!\right) \cdot exp\left(-\eta\right) \tag{3.4}$$

where $P_{he}(i)$ is the proportion of the hosts i eggs are laid on; η is the mean number of the eggs laid on each host.
According to Eq.3.4, the proportion of survived hosts that get zero eggs is $exp\left(-\eta\right)$. Hence, the number of hosts that get eggs, denoted by $A^{'}$, is given by Eq.3.5.

$$A^{'} = A\left(1 - exp\left(-\omega N_{eg}^{'}(t)/A\right)\right) \tag{3.5}$$

where A is the total number of hosts; ω is a coefficient that controls the mean value of the Poisson distribution.
We suppose that if a host is parasitized by more than one eggs, only one egg survives, and the others die due to lack of resource. Therefore, the final number of eggs that parasitize hosts and survive is equal to $A^{'}$.

3.1.4 Population dynamics

Besides parasitism, pests and auxiliaries may die due to other factors, which are integrated to a viability coefficient. Stage growth, i.e. population dynamics within stages, is calculated by Eq.3.6.

$$NT_s(i,t) = NT_s(i-1,t-1) \cdot C_s(t-1) \quad (3.6)$$
$$\text{if} \begin{cases} \text{egg stage}, & 1 \le i \le t_{eg} \\ \text{juvenile stage}, & 1 < i \le t_{ju} \\ \text{adult stage}, & 1 < i \le t_{ad} \end{cases}$$

where $NT_s(i,t)$ is the number of pests (or auxiliaries) at stage s (i.e. egg (eg), juvenile (ju), adult (ad)) of age i at time t; t_{eg} is egg longevity; t_{ju} is juvenile longevity; t_{ad} is adult longevity; $C_s(t-1)$ is the viability rate of an individual at stage s at time $t-1$. Stage transition, i.e. population dynamics from one stage to another, is calculated by Eq.3.7, where $NT_{eg}(0,t)$ will be introduced in section 2.3.2 for pests and section 2.3.4 for auxiliaries.

$$\begin{aligned} NT_{ju}(1,t) &= NT_{eg}(t_{eg},t-1) \cdot C_{eg}(t-1) \\ NT_{ad}(1,t) &= NT_{ju}(t_{ju},t-1) \cdot C_{ju}(t-1) \end{aligned} \quad (3.7)$$

In the model, the viability rate C_s is a time-variation variable. It could be either constant all the time or be variable with environmental conditions. In the thesis, we suppose that it may increase as the amount of resource increases, as expressed by Eq.3.8.

$$C_s(t) = c \cdot exp(-\lambda(|NT_s(t)|/B)) \quad (3.8)$$

where $|NT_s(t)|$ is the total number of pests or auxiliaries at stage s at time t; c is the maximal viability rate at stage s; B represents the amount of resource, e.g. leaf biomass for pests, number of pests for auxiliaries; λ is a coefficient. If λ is zero, the viability rate is constant; otherwise, it varies.

3.2 Population dynamics model – specifics

Besides leaf biomass is involved in the calculation of the amount of resource of pests, leaf biomass is involved indirectly in the calculation of the auxiliary population, as auxiliary hosts are insect pests. Hence, in this subsection, we first recall the procedures of leaf development in GreenLab.
Here, we define another variable to represent the accumulated biomass of a leaf initiated at time k when plant age is t, denoted by $X(k,t)$. For the sake of clarity, we take leaves of PA 1 as example. The procedure of a leaf development is given by

$$X(k,t+1) = X(k,t) + p_1^b(t-k+1)(Q(t)/D(t)) \quad (3.9)$$

The plant biomass production is thus given by Eq.3.10.

$$Q(t) = E(t)\mu Sp\left(1 - exp\left(-\frac{k}{Sp \cdot slw}\sum_{k=t-ta}^{t-1} X\left(k,t\right)\right)\right) \tag{3.10}$$

For the survival of pest juveniles during the whole life t_{ju}, α cm^2 of leaf area is required and is destroyed, which is equivalent to $\alpha \cdot slw$ g of leaf biomass. Different from the pest juveniles, pest eggs do not destroy leaf area and have no effect on plant photosynthesis. Even though pest eggs do not destroy leaf area, the model assumes that female pest adults lay eggs with the constraint to guarantee that there is enough food for evolved juveniles to survive in ideal conditions. Hence, pest eggs occupy certain area on the leaf and thus affect the partition of pest eggs on leaves. Therefore, two kinds of variables of leaf biomass should be defined. One represents leaf biomass available to accept new eggs, denoted by X_B, and the other represents the leaf biomass that is used for photosynthesis, denoted by X_f.

3.2.1 Amount of resource for pests

The factor A in Eq.3.5 represents the number of hosts, which is the potential number of eggs plant can accept. Hence, X_B is involved. A is calculated by Eq.3.11. As only green leaf area is attractive to pests and ta is leaf life-span, the lower bound of the summation in Eq.3.11 is $t - ta$. The setting of the lower bound of the summation in Eq.3.12, Eq.3.13 and Eq.3.14 is in a similar way.

$$A = \sum_{l=t-ta}^{t-1} X_B\left(l,t\right) / \left(\alpha \cdot slw\right) \tag{3.11}$$

According to the model assumptions, for pest eggs, the amount of resource is the available leaf biomass X_B, whereas for pest adults, it is the total green leaf biomass that is not eaten by juveniles or adults. Hence, the factor B for eggs in Eq.3.8 is the same as Eq.3.11, whereas for adults (or juveniles), it is given by Eq.3.12.

$$B = \sum_{l=t-ta}^{t-1} X_f\left(l,t\right) \tag{3.12}$$

3.2.2 Pest egg partition

The model assumes that the attraction of pests to a leaf for laying eggs is proportional to the corresponding leaf area available to accept new eggs. Hence, the number of eggs laid on the leaf initiated at time k when plant age is t is calculated by Eq.3.13.

$$PI_{eg}\left(k,0,t\right) = \left(X_B\left(k,t\right) / \sum_{l=t-ta}^{t-1} X_B\left(l,t\right)\right) \cdot NI_{eg}\left(t\right) \tag{3.13}$$

where $PI_{eg}(k,0,t)$ is the number of pest eggs laid on the leaf initiated at time k when plant age is t, whose age is denoted by zero; $NI_{eg}(t)$ is the final number of pest eggs plant accepts, calculated by Eq.3.5 and Eq.3.11.

3.2.3 Amount of resource for auxiliaries

For the species of oil palm (*Elaeis guineensis*) that we study, six main species of auxiliaries are observed (Lecoustre [1988]): two species of egg auxiliary, denoted by PE and four species of larva auxiliary, denoted by PW. Egg auxiliaries feed on pest eggs, and larva auxiliaries feed on pest larvae. Hence, the resource for auxiliaries is pest eggs or pest juveniles. We take egg auxiliary as an example to introduce the auxiliary development and the interaction between pests and auxiliaries. In the following part, egg auxiliary is abbreviated to auxiliary.
Suppose that pest eggs die once they are parasitized within one time unit, the amount of resource in Eq.3.5 and Eq.3.8 are the same, which is total number of pest eggs on plant, given by Eq.3.14.

$$A = \sum_{k=t-ta}^{t-1} |PI_{eg}(k,t)| \tag{3.14}$$

3.2.4 Auxiliary egg partition

The number of parasitized pest eggs on each leaf is assumed to be proportional to the ratio of the number of pest eggs on each leaf to total number of pest eggs, as given by Eq.3.15.

$$PE_{eg}(k,0,i,t) = NPE_{eg}(t) \cdot PI_{eg}(k,i,t)/A \tag{3.15}$$

where $PE_{eg}(k,0,i,t)$ is the number of auxiliary eggs parasitizing pest eggs of age i which are on the leaf initiated at time k when plant age is t, the age of the new auxiliary eggs being 0; $NPE_{eg}(t)$ is the final number of auxiliary eggs on plant, calculated by Eq.3.5 and Eq.3.14; A is given by Eq.3.14.

3.2.5 Interaction between pests and auxiliaries

In the thesis, we suppose that the number of survived auxiliaries is equal to the number of parasitized pests. Hence, the number of pest eggs survived from parasitism at time t is $PI_{eg}(k,i,t) - PE_{eg}(k,0,i,t)$, and the number of survived pest juveniles is $PI_{ju}(k,i,t) - PW_{eg}(k,0,i,t)$, $PI_{ju}(k,i,t)$ being the number of pest juveniles of age i on the leaf initiated at time k when plant age is t and $PW_{eg}(k,0,i,t)$ being the number of eggs of larva auxiliaries of age zero parasitizing pest juveniles of age i on the leaf initiated at time k when plant age is t.

3.3 Interaction with population dynamics model

It involves tri-trophic components to simulate plant growth with consideration of the interaction with population dynamics: plants, insect pests and auxiliary insects. In the thesis, we call the model that can simulate the interactions among the tri-trophic components tri-trophic ecosystem model. Plant growth and population dynamics have been modeled using GreenLab and a dynamic model described previously. To model tri-trophic ecosystem is to link population dynamics model to GreenLab. However, the following issues should be considered and be handled, in order to successfully link the two models.

- Time discretization.

  In GreenLab, time is discretized by growth cycle corresponding to the time span of the appearance of two successive phytomers, which is a function of thermal time. While, for population dynamics of insect pests and auxiliary insects, individual development at each stage is also related to thermal time. Hence, we consider the shortest time span between the growth cycle in GreenLab and the durations of each stage of pests and auxiliaries as time unit, for the tri-trophic ecosystem model.

- Event order.

  As the model is discrete, within one time unit, many sub-ecological processes are involved, e.g. egg laying of insects, biomass production and partition, parasitism of plant and pests. Different orders of all the sub-ecological processes within one time unit lead to significant different results (Gosselke et al. [2001]). Therefore, it is important to define the order. In the present model, the order of the sub-ecological processes is presented by the flowchart shown in Fig.3.1.

3.3.1 Decrement of leaf area due to pest attacks

From the number of survived pest eggs, juveniles and adults on each green leaf, we can deduce the leaf area occupied by pest eggs and eaten by juveniles and adults. We assume that juveniles and adults eat leaf area uniformly during their whole lives. The leaf area available for accepting new eggs and for photosynthesis after being eaten by pests are given by Eq.3.16 and Eq.3.17.

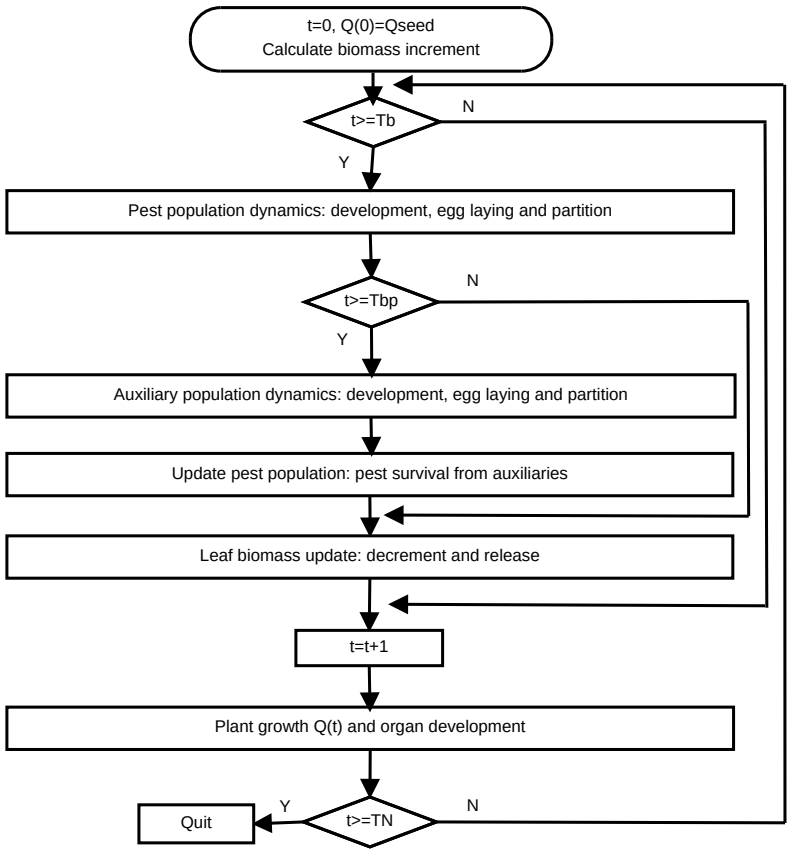


Figure 3.1: Flowchart of the tri-trophic ecosystem. T_b is the time when pests come and begin to attack plant; T_{bp} is the time when auxiliaries come and begin to attack pests; $Q(t)$ is plant biomass production at time t; Q_{seed} is the seed biomass; T_N is plant longevity.

$$\begin{aligned} X_B\left(k,t+1\right) = X_B\left(k,t\right) + \underbrace{p_1^b\left(t-k+1\right)\frac{Q(t)}{D(t)}}_{I} \\ \underbrace{-\left[PI_{eg}\left(k,0,t\right) - PE_{eg}\left(k,0,0,t\right)\right]\cdot\alpha\cdot slw}_{II} \end{aligned} \tag{3.16}$$

$$\begin{aligned} X_f\left(k,t+1\right) = X_f\left(k,t\right) + \underbrace{p_1^b\left(t-k+1\right)\frac{Q(t)}{D(t)}}_{I} \\ \underbrace{-\frac{\beta\cdot slw}{t_{ad}}\cdot\left(\left|PI_{ad}(k,t)\right|\right)}_{III} \\ \underbrace{-\sum_{j=1}^{t_{ju}}\left(PI_{ju}\left(k,j,t\right) - PW_{eg}\left(k,0,j,t\right)\right)\cdot\frac{\alpha\cdot slw}{t_{ju}}}_{IV} \end{aligned} \tag{3.17}$$

where β is the leaf area eaten by an adult during its whole life t_{ad}; $|PI_{ad}(k,t)|$ is the total number of pest adults on the leaf initiated at time k when plant age is t; the item I in Eq.3.16 and Eq.3.17 represents the biomass partitioned by plant biomass; the item II represents the biomass that is equivalent to the area occupied by pest eggs; the item III represents the biomass eaten by adults; the item IV represents the biomass eaten by juveniles.
From Eq.3.17, adults first eat leaves where juveniles, from whom they evolved, stayed, and then they fly away to choose leaves as big as possible to lay eggs. However, adults could choose bigger leaves to eat. To implement this strategy, the item III in Eq.3.17 is replaced by Eq.3.18.

$$NI_{ad}(t) \cdot \frac{\beta \cdot slw}{t_{ad}} \cdot \frac{X_f(k,t)}{\sum_{l=t-ta}^{t-1} X_f(l,t)} \tag{3.18}$$

where $NI_{ad}(t)$ is the total number of pest adults at time t.

3.3.2 Release of leaf area

As detailed in section 3.2.1, eggs are laid on leaf area with the constraint that there is enough food for future juveniles to survive under ideal conditions. Therefore, if eggs or juveniles die, the leaf biomass available to accept new eggs should be corrected.
Egg death results from parasitism and from other factors which are integrated to the viability rate C_{eg}. Therefore, the leaf area released due to the death of eggs at the end of time t, denoted by $R_{eg}(t)$, is given by Eq.3.19.

$$R_{eg}(t) = \sum_{i=0}^{t_{eg}} \left[\left(PI_{eg}(k,i,t) - PE_{eg}(k,0,i,t) \right) \cdot \left(1 - C_{eg}(t) \right) \cdot \alpha \right] \tag{3.19}$$

A juvenile of age i has eaten a quantity of leaf area equal to $\alpha \cdot i/t_{ju}$. Hence, the area released due to juvenile deaths at time t, denoted by $R_{ju}(t)$, is given by Eq.3.20.

$$R_{ju}(t) = \sum_{i=1}^{t_{ju}} \left[\left(PI_{ju}(k,i,t) - PW_{eg}(k,0,i,t) \right) \cdot \left(1 - C_{ju}(t) \right) \cdot \frac{\alpha}{t_{ju}} \left(t_{ju} - i \right) \right] \tag{3.20}$$

Therefore, the biomass available for accepting new eggs due to the deaths of eggs and juveniles is given by Eq.3.21.

$$\begin{aligned} X_B(k,t+1) &= X_B(k,t) + p_1^b(t-k+1)\frac{Q(t)}{D(t)} \\ &- \left(PI_{eg}(k,0,t) - PE_{eg}(k,0,0,t) \right) \cdot \alpha \cdot slw \\ &+ R_{eg}(t) \cdot slw + R_{ju}(t) \cdot slw \end{aligned} \tag{3.21}$$

3.4 Conclusion

The assumptions and the principles of the insect population dynamics model developed in this thesis is introduced in this chapter. The population dynamics model is linked with the GreenLab model, and a tri-trophic ecosystem model, which is able to simulate plant growth with consideration of the interaction of population dynamics, is thus modeled. The analysis and optimization applications of the tri-trophic ecosystem model will be introduced in part III.

Part II

Formulation of optimization problems based on GreenLab

Chapter 4

Formulation of optimization problems based on GreenLab

In the modeling process, optimization techniques are required at several steps, in particular for parametric identification and for practical model applications, such as designing ideotypes or improving cultivation management. It is common that some parameters of a model cannot be experimentally measured and need to be estimated from indirect observations. Without the procedure of parameter identification (or parameter estimation), a model remains purely theoretical and cannot be validated nor be used for application of prediction, optimization or optimal control. The aim of parameter identification is to find the optimal set of model parameter values that minimize the difference between model outputs and the corresponding observed experimental data. Once model parameters are identified with experimental data, the model can be applied to predict and to optimize yield within the appropriate range of environmental conditions. According to the definition of optimization concept in mathematics: optimization refers to choosing the best set of variables from available alternatives that satisfy certain objectives, parameter identification is an optimization problem, from a mathematical point of view. However, in this thesis, the optimization problems we focus on concern the investigation of the optimal plant under certain environmental condition without reference to any experimental data that would be considered as objective. Hereafter, we use the concept of *parameter identification* (or *estimation*) to represent the procedure of minimizing the difference between simulated outputs of model and experimental data. And *optimization* only refers to finding the optimal set of parameters with which plant can achieve the optimal objective, e.g. optimal yield, no matter whether this optimal plant is realistic or not. In this thesis, two categories of problems are involved: optimization and optimal control. Optimization issues deal with optimizing system inner variables, while optimal control issues deal with optimizing system external variables. Taking plant as system, inner factors of the system are the endogenous factors related to plant genetics, and external factors of the system are the exogenous factors related to environment or cultivation mode, water supply strategy or pruning strategy for in-

stance. Let us now introduce how to formulate such optimization and optimal control problems in the frame of the functional-structural plant growth model GreenLab.

4.1 Dynamic discrete system

According to the mathematical formulism of GreenLab presented in chapter 2, GreenLab can be considered as a dynamic discrete system. A dynamic discrete system involves two types of unknowns: parameter $P \in \mathbb{R}^p$ and state variables $X_n \in \mathbb{R}^x$, $n = 1, 2, \ldots, t_n$, t_n being the finite sequence of successive times. The parameter and the state variables of the system are subject to the state equation given by Eq.4.1, with X_0 given.

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

where U_n is exogenous variables at time t_n, e.g. water supply strategy, pruning strategy. Model output Y corresponding to observation data $\hat{Y}$, are related to state variables X, X_n being the nth coordinate of X. The relation between Y and X is as expressed by Eq.4.2.

$$Y = G(X, P) \tag{4.2}$$

where Y are vectors in time series: $Y_n, n = 1, 2, \ldots, t_n$ is the nth coordinate of Y, model outputs at each given time.
Hence, if the initial state variable X_0 and U_n are known, model output Y is a function of model parameter P, denoted by $Y(P)$.

4.2 Parameter identification

Parameter identification problems deal with the reconstruction of unknown parameters in systems of differential equations by model inversion (Burger [2002]). The purpose of parameter identification is to find the parameters that minimize a so called fitness function for given observation data. If the problem has a solution, parameters are identifiable. Generally, the fitness function for parameter identification is chosen as expressed by the least square criterion as the following equation (Letort [2008], Cournède [2009]):

$$P^* = argmin_P((\hat{Y} - Y(P))\Omega^{-1}(\hat{Y} - Y(P))) \tag{4.3}$$

where Ω is the covariance matrix consisting of the error vector $\hat{Y} - Y$. This expression gives the best unbiased estimator of P in the case of a linear model. Although our model is not linear, it can be locally approximated by a linear model by applying the same method with several iterations. Generally, Ω is unknown. We assume that errors are independent, thus, Ω is diagonal, given by

$$\Omega_i = \frac{\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})}{n_i} \cdot \frac{n}{n-p} \tag{4.4}$$

where $\hat{y}_{ij}$ is the ith observation data of the jth output of interest, $i = 1, 2, \ldots, n_i$; n is the total number of observations; p is the number of parameters to be identified.

4.3 Optimization and optimal control

For a dynamic discrete system as expressed by Eq.4.1, our objective is to optimize either endogenous parameters P or exogenous factors U, or both of them in order to achieve optimal objective values. According to the types of variables to be optimized, different types of optimization problems can be distinguished: (1) direct optimization of system endogenous parameters for ideotype design and (2) optimal control on exogenous factors to guide optimal cultivation modes.

4.3.1 Optimization for ideotype design

For this type of optimization problem, the objective is to find the optimal set of endogenous parameters to achieve optimization objectives, under certain environmental conditions. Generally speaking, the optimization objectives are yield of plant or organs. As environmental conditions are given, the exogenous parameter U is known. Therefore, the criterion of the optimization problem is given by

$$P^* = argmin_p J(X, P) \tag{4.5}$$

where J is an objective function.
Even though the environmental condition with which plants grow is given, plants grow also under other constraints, e.g. amount of pollen for fruit development. Hence, constraints have to be concerned and be integrated into optimization problems. The corresponding optimization problem with constraints is given by

$$\begin{aligned} &P^* = argmin_p J(X, P) \\ &\text{subject to} \quad g(X, P) < 0 \end{aligned} \tag{4.6}$$

where $g(X, P)$ represents the set of constraint functions.
For some plant species, several parts of the plants can have an interest, each of them having different economical values. In order to benefit from all of them, multi-objective optimization problems are encountered. They can be expressed by

$$P^* = argmin_p \hat{J}(J_1, J_2, \ldots, J_m) \tag{4.7}$$

where J_i represents the ith objective function, $i = 1, 2, \ldots, m$, m being the number of objective functions.
The multiobjective optimization problem with constraints is thus given by

$$\begin{aligned} &P^* = argmin_p \hat{J}(J_1, J_2, \ldots, J_m) \\ &\text{subject to} \quad g(X, P) < 0 \end{aligned} \tag{4.8}$$

4.3.2 Optimal control

For a given species, besides optimization on the endogenous factors of a plant, another way to improve plant growth in the domain of agriculture and forestry is to improve cultivation modes, e.g. water supply strategy, pruning strategy. This kind of optimization problem belongs to the field of optimal control. Its objective is to find the best sequence of $U \in \Gamma_U$ to optimize the objective function $J(X, P)$, Γ_U corresponding to the available set of control variables. For GreenLab, the initial biomass X_0 coming from the seed is given as input variable, and thus according to the state equation given by Eq.4.1, the sequence (X_n) is fully determined by X_0, P and U. As the plant species is given, P is known. Hence, X_n is a function of U denoted by $\phi(X_0, U)$. The optimal control problem is as follows (Cournède [2009]).

$$U^* = argmin_U J(\phi(X_0, U), U) \tag{4.9}$$

We apply the variational approach and the Lagrange theory to solve the optimal control problem. As the optimal set of control variables must insure that the state deduced from the control variables satisfy the state equation, the state equation is the constraint of the optimal control problem, given by

$$\begin{aligned} &U^* = argmin_u J(X, U) = argmin_u \sum_{t=0}^{N-1} G_t(X_t, U_t) + \Phi(X_N) \\ &X_{t+1} = F_t(X_t, U_t) \end{aligned} \tag{4.10}$$

We impose a Lagrange coefficient Λ to transform the constrained problem to unconstrained problem. The objective function is given by

$$\hat{J} = \Phi(X_N) + \sum_{t=0}^{N-1} \left[G_t(X_t, U_t) + {\Lambda_{t+1}}^T (F_t(X_t, U_t) - X_{t+1}) \right] \tag{4.11}$$

We define the Hamiltonian function in time series, as expressed by

$$H(X_t, U_t, \Lambda_t) = G_t(X_t, U_t) + \Lambda_{t+1}^T F_t(X_t, U_t) \tag{4.12}$$

Substitute Eq.4.12 into Eq.4.11, we get

$$\hat{J} = \Phi(X_N) + \sum_{t=0}^{N-1} \left[H_t(X_t, U_t, \Lambda_{t+1}) - \Lambda_{t+1}^T X_{t+1} \right] \tag{4.13}$$

The derivative of $\hat{J}$ is given by

$$\begin{aligned} d\hat{J} &= \frac{\partial \Phi^T(X_N)}{\partial X_N}\partial X_N + \sum_{t=0}^{N-1}\left[\frac{\partial H_t^T}{\partial X_t}\partial X_t + \frac{\partial H_t^T}{\partial U_t}\partial U_t + \frac{\partial H_t^T}{\partial \Lambda_{t+1}}\partial \Lambda_{t+1}\right] \\ &\quad - \sum_{t=0}^{N-1}\left[X_{t+1}^T \partial \Lambda_{t+1} + \Lambda_{t+1}^T \partial X_{t+1}\right] \end{aligned} \tag{4.14}$$

Regrouping Eq.4.14, we get

$$\begin{aligned} d\hat{J} &= \left(\frac{\partial \Phi^T}{\partial X_N} - \Lambda_N^T\right)\partial X_N + \frac{\partial H^T}{\partial X_0}\partial X_0 + \sum_{t=1}^{N-1}\left[\frac{\partial H^T}{\partial X_t} - \Lambda_t^T\right]\partial X_t \\ &\quad + \sum_{t=0}^{N-1}\left[\frac{\partial H^T}{\partial U_t}\partial U_t + \left(\frac{\partial H^T}{\partial \Lambda_{t+1}} - X_{t+1}^T\right)\partial \Lambda_{t+1}\right] \end{aligned} \tag{4.15}$$

As the initial state X_0 is given, $\partial X_0 = 0$. Hence, the necessary conditions for a minimum of the optimal control problem are given by the following equations. The optimal control problems can thus be solved.

$$X_{t+1} = \frac{\partial H}{\partial \Lambda_{t+1}}, \quad t \in [0, N-1] \tag{4.16}$$

$$\Lambda_t = \frac{\partial H}{\partial X_t}, \quad t \in [1, N-1] \tag{4.17}$$

$$\frac{\partial H}{\partial U_t} = 0 \tag{4.18}$$

$$\Lambda_N = \frac{\partial \Phi}{\partial X_N} \tag{4.19}$$

According to Eq.4.18, the gradient of the objective function Eq.4.11 with respect to U is given by

$$0 = \frac{\partial G}{\partial U_t} + \Lambda_{t+1}^T \frac{\partial F}{\partial U_t} \tag{4.20}$$

4.4 Numerical optimization method

Optimization problems can be classified into several classes: constrained and unconstrained problems, linear and nonlinear problems, continuous and discrete problems, single objective and multi-objective problems. Generally speaking, each class of problems require specific optimization algorithms used to search solutions. Optimization algorithms can be classified as local optimization algorithms and global optimization

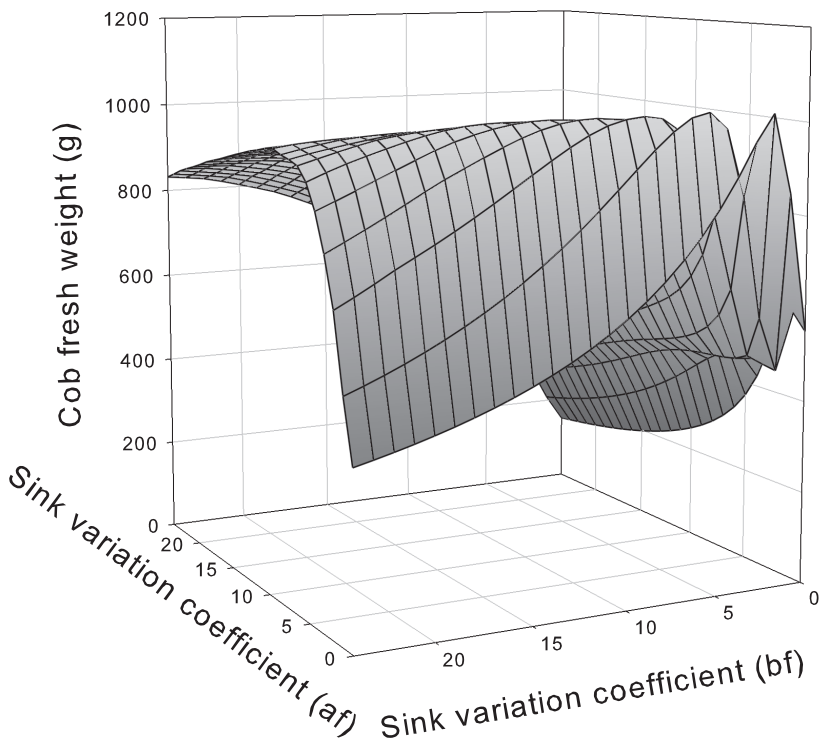


Figure 4.1: Simulation results of cob yield of maize with respect to coefficients of cob sink variation function (a^f, b^f).

algorithms, classical and heuristic optimization algorithms, direct and iterative (numerical) optimization algorithms. The optimization problems that we study revealed non-convex and multimodal as shown in Fig.4.1, particularly there is no unique solution for multi-objective problems. Therefore, the iterative, population-based heuristic optimization algorithm, namely Particle Swarm Optimization, was used in our study, thanks to its high convergence rate and generalization ability compared with other algorithms.

4.4.1 Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) was first proposed by Kennedy and Eberhart [1995], which originally simulated the behaviour of bird flocking. It is an iterative, population-based method. The particles are described by their two instinct properties: position and velocity. The position of each particle represents a point in the parameter space, which is a possible solution of the optimization problem, and the velocity is used

to change the position. The particle properties are time-variant. They are updated by Eq.4.21 and Eq.4.22 in the standard PSO.

$$v_{ij}^{k+1} = \omega^k \cdot v_{ij}^k + c_1 r_1 \left(B_{ij} - x_{ij}^k\right) + c_2 r_2 \left(B_{gj} - x_{ij}^k\right) \tag{4.21}$$

$$x_{ij}^{k+1} = x_{ij}^k + v_{ij}^{k+1} \tag{4.22}$$

$$i = 1, 2, \ldots, Np$$

$$j = 1, 2, \ldots, Nd$$

where Np is the number of particles in the population; Nd is the number of variables of the problem (i.e. dimension of a particle); v_{ij}^k is the j^{th} coordinate component of the velocity of the i^{th} particle at iteration k; B_{ij} is the j^{th} coordinate component of the best position recorded by the i^{th} particle during the previous iterations; B_{gj} is the j^{th} coordinate component of the best position of the global best particle in the swarm, which is marked by g; x_{ij}^k is the j^{th} coordinate component of the current position of particle i at the k^{th} iteration; ω^k is the inertia weight at iteration k, which decreases linearly as iteration increases as expressed by Eq.4.23; c_1, c_2 are the acceleration coefficients; r_1, r_2 are the uniformly distributed random values between 0 and 1. The last two items on the right side of Eq.4.21 are considered as cognition knowledge and social knowledge of a particle. From Eq.4.21 and Eq.4.22, the direction and the distance controlling how individuals move are determined by their velocities and their experiences during the search. With the help of social and cognition knowledge of each individual, the population (also called swarm) converges to the optimal solution (or position).

$$\omega^k = \frac{MAXITER - k}{MAXITER} \cdot (\omega_{start} - \omega_{end}) + \omega_{end} \tag{4.23}$$

where ω_{start} and ω_{end} are the initial and the final values of inertia weight respectively; $MAXITER$ is the maximal number of algorithm iterations.

During the decades of development, many variants of PSO have been proposed by researchers, in order to enhance the convergence accuracy or adapt it to specific problems. A review of PSO is given in Song and Gu [2004]. In this thesis, PSO with passive congregation (He et al. [2004]) is used to solve single objective optimization problems, thanks to its generalization capacities and robust performance. The equations used to calculate velocities and the new positions in the PSO with passive congregation are given by:

$$\begin{aligned} v_{ij}^{k+1} &= \omega^k v_{ij}^k + c_1 r_1 \left(B_{ij} - x_{ij}^k\right) + c_2 r_2 \left(B_{gj} - x_{ij}^k\right) + c_3 r_3 \left(B_{rj} - x_{ij}^k\right) \\ x_{ij}^{k+1} &= x_{ij}^k + v_{ij}^{k+1} \end{aligned} \tag{4.24}$$

The difference between standard PSO and PSOPC is the fourth item on the right side of Eq.4.24. B_{rj} is the jth coordinate of the best position recorded by a random selected particle r during the previous iterations. It is used to avoid converging to the local optimum.

The specific algorithm that we used for the multi-objective optimization problem is the mixture of the algorithms proposed by Mostaghim and Teich [2003] and by Tripathi et al. [2007]. To extend the original PSO to solve multi-objective problems and to find the optimal solutions, known as *Pareto front*, the equations for changing the velocity and position of each particle are improved slightly, as given by

$$\begin{aligned} v_{ij}^{k+1} &= \omega^k v_{ij}^k + c_1 r_1 \left(B_{ij} - x_{ij}^k\right) + c_2 r_2 \left(B_{lj} - x_{ij}^k\right) \\ x_{ij}^{k+1} &= x_{ij}^k + v_{ij}^{k+1} \end{aligned} \tag{4.25}$$

The aim of multi-objective optimization problems is to find all the optimal solutions that form the *Pareto front*. Therefore, to obtain various solutions at a given iteration, the algorithm is changed by replacing the unique global best position with a local guide best position for each particle, denoted by Bl_{ij} for the jth coordinate of particle i in Eq.4.25. For the problems with constraints, there are two criteria to decide whether the best position of each particle B_i is updated by the new position x_i^{k+1}: if x_i^{k+1} satisfies the constraints while B_i does not, or if one of the objective function value with respect to x_i^{k+1} is better than the one with respect to B_i, no matter whether the constraints are satisfied, replace B_i with x_i^{k+1}.
All the optimal solutions are recorded in an archive with limited size. The total number of solutions inside the *Pareto front* is thus controlled by the archive size. If the number of optimal solutions does not achieve the archive size, all of them are accepted and added to the archive; otherwise, the most similar optimal solutions inside the archive, which is evaluated by the criteria of the nearest distance between each two solutions, will be eliminated. The Sigma method (Mostaghim and Teich [2003]) is used to determine the local guide best position of each particle. The solution in the archive which has the nearest distance from a given particle is decided to its local guide best position. For more details, we refer to (Mostaghim and Teich [2003]).

4.4.2 Performance comparison among Particle Swarm Optimization and other optimization algorithms

Besides PSO, there are other heuristic algorithms that were developed earlier and are also widely used, e.g. Simulated Annealing (SA) (Poupaert and Deville [2000]), Genetic Algorithm (GA) (Houck et al. [1996], Sastry et al. [2005]) and Ant Colony Optimization (ACO) (Dorigo and Blum [2005]). To be fair, the performances of PSO and other heuristic optimization algorithms should be compared using the same PC and the same platform of PC and programming language. In the latest version of Scilab (Web), two heuristic optimization algorithms are integrated as inner function: for instance, function *optim_ga* of Genetic Algorithm (GA) and function *optim_sa* of Simulated Annealing (SA), for single objective problems. These inner Scilab functions should be more matured than the ones implemented by other authors, as Scilab is released offi-

cially. In addition, PSO has been implemented in Scilab as toolbox namely PSOTS by the author, which can solve continuous problems, discrete problems and mixed integer problems, single and multi-objective problems. Hence, we compared the performances of GA, SA, and PSO by using the Scilab inner functions and the toolbox PSOTS and by employing two widely accepted criteria. The first criterion is the error between the optimal function value f_{obj} and the known analytical optimal value f_{anal}, which satisfies $|f_{obj} - f_{anal}| < 0.001$ in our experiment and the second is the maximum number of objective function evaluation, which is set to be 100000 and 2000000 respectively. When the first criterion is satisfied, we think that this algorithm is successful in this experiment. When the second criterion is satisfied, the consumed time is recorded and returned right way and the current optimal solution is recorded, no matter whether the algorithm converges to the global optimal value.
First, we compare the performance of PSO with GA and SA by minimizing DeJong function as expressed by Eq.4.26 with increasing number of dimensions. In the thesis, PSO with passive congregation (PSOPC) is used. The searching range for each dimension is between -100 and 100, and the population size is set to be 100 for GA and PSO. To compare the consuming time of each algorithm, the number of objective function evaluation is used as the criterion. The number of objective function evaluation for GA and PSO is population size times number of iteration, while for SA, it is equal to the multiplication of number of temperature decrease and number of iterations during each temperature stage. Hence, if the maximal number of iterations for GA and PSO is set to be 100, the number of temperature decrease is set to be 100 and the number of iterations during each temperature stage is 100; if the maximal number of iterations for GA and PSO is set to be 2000, the number of temperature decrease and the number of iterations during each temperature stage are set to be 200 and 1000 respectively. The other parameter values of GA and SA are the default values in Scilab, and the other parameter values of PSO are chosen as recommended by He et al. [2004]: the acceleration coefficients (c_1, c_2 and c_3) are 0.5, 0.5 and 0.6 respectively, and the inertia weight decreases from 0.9 to 0.7 linearly. For each algorithm, we run it 10 times independently. The average value of the consumed time, success rate and convergence rate of all 10 independent runnings are listed in Table 4.1 and Table 4.2.

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i^2 \tag{4.26}$$

Time cost

We adopt CPU time to compare the time consumed by each algorithm. From Table 4.1, we found that the time consumed by using PSO is less that that by using GA and SA. However, the time consumed by using PSO is not stable, increasing as the problem dimension increases (i.e. number of variables). It is because in the PSO code, there is

a loop that is related to the problem dimension.

Convergence accuracy

No matter what the complexity of the optimization problem is, the optimal solutions found by PSO are all better than that found by GA and SA, as listed in Table 4.1 and 4.2. In addition, comparing the success rate, we found that PSO successfully solves more complex problems that GA and SA deal with with low successful rate or even they completely failed to solve. PSO is more able to deal with computationally expensive and complex optimization problems.

Convergence rate

Even though GA and SA found the optimal solution of the problem when the number of objective function evaluation is large, 200000 for instance as shown in Table 4.2, the solutions returned by GA after 10000 iterations are much worse, as listed in Table 4.1. Even though the solutions returned by SA are acceptable when the number of objective function evaluation is 10000, the optimal solutions found by PSO are all better. Compared with GA and SA, PSO has a faster convergence rate.

Table 4.1: Comparison of optimal results of DeJong function over 10 independent runnings, by GA, SA and PSO, number of objective function evaluation being 10000. The optimal value of DeJong function is 0, obtained at the origin coordinates.

Dimension	Time consuming (s)			Success rate			Optimal solution[1]		
	PSO	GA	SA	PSO	GA	SA	PSO	GA	SA
2	0.35	10.90	2.74	100%	100%	100%	0	0	$2.40e-6$
							0	0	0
4	0.60	10.96	2.77	100%	100%	100%	$4.32e-5$	$1.67e-4$	$1.44e-4$
							$2.30e-5$	$2.00e-5$	$3.20e-5$
6	0.86	10.99	2.80	100%	30%	30%	$2.92e-4$	6.04	$1.28e-3$
							$6.90e-5$	$2.39e-4$	$6.58e-4$
8	1.11	11.07	2.78	100%	0%	0%	$9.42e-4$	73.94	$3.42e-3$
							$7.13e-4$	2.07	$2.01e-3$
10	1.35	10.95	2.78	0%	0%	0%	$3.31e-3$	176.76	$5.96e-3$
							$1.08e-3$	26.07	$2.75e-3$

[1] the number in the first row represents the average value; the number in the second row represents the minimum value.

Besides continuous problems, we take Traveling Salesman Problems (TSP) as example to compare the performances of PSO, GA and SA for discrete optimization problems. The objective of TSP is to find the shortest path which connects all cities, each city

Table 4.2: Comparison of optimal results of DeJong function over 10 independent runnings, by GA, SA and PSO, number of objective function evaluation being 200000. The optimal value of DeJong function is 0, obtained at the origin coordinates.

Dimension	Time consuming (s)			Success rate			Optimal solution		
	PSO	GA	SA	PSO	GA	SA	PSO	GA	SA
2	6.97	216.65	54.99	100%	100%	100%	0	0	0
							0	0	0
4	12.12	216.90	55.09	100%	100%	100%	0	$4.60e-5$	$4.08e-5$
							0	$1.00e-6$	$6.00e-6$
6	17.33	220.84	55.33	100%	100%	100%	0	$3.69e-4$	$3.74e-4$
							0	$8.10e-5$	$1.75e-4$
8	22.09	219.47	55.19	100%	60%	0%	0	$1.05e-3$	$1.52e-3$
							0	$6.70e-4$	$1.13e-3$
10	26.85	220.72	55.34	100%	10%	0%	0	$2.41e-3$	$2.99e-3$
							0	$6.21e-4$	$2.34e-3$

[1] the number in the first row represents the average value; the number in the second row represents the minimum value.

being passed only once. TSP is a kind of combinatory optimization problem. The TSP instance used in the thesis is chosen from the public TSP library (TSPLIB95) named Burma14 (TSPLIB95). Even though the number of cities for Burma14 is 14, there are many paths with similar path length with the shortest path (Angus and Hendtlass [2005]), hence it is sufficient to test the performance of optimization algorithms for TSP. In the latest version of Scilab, the integrated GA and SA cannot solve TSPs, i.e. they have not been adapted or extended to TSPs yet. Hence, we use the GATS toolbox developed by Li and Hu [2005] for GA, and use the SA toolbox provided by Collette [2009] who is the developer of the integrated function of GA and SA in Scilab. The PSO for TSPs is integrated in the toolbox PSOTS and is referred to Clerc [2002]. Even though the geographical distance is adapted in TSPLIB95, we calculate the euclidean distance of the TSP Burma14, as SA provided by Collette [2009] is only suitable for the euclidean distance computation so far. The maximal number of the objective function evaluation is considered to be the stop criterion. If the stop criterion is satisfied, the time consumed by the algorithm is recorded. If the algorithm finds the optimal solution, the number of the objective function evaluation is recorded, and it is considered as the factor to evaluate the convergence rate of the algorithm. Among 10 independent runnings, all the alogrithms find the optimal solution, i.e. the success rate is 100% for all the algorihtms. The comparison results of the performance of PSO, GA and SA in terms of the consumed time and the convergence rate are shown in Table 4.3.

From the results listed in Table 4.3, the convergence rates of GA and SA for TSPs are faster than PSO. One of the reasons may be that PSO is first proposed and tested

Table 4.3: Comparison of the performance of PSO, GA and SA on the TSP Burma14 over 10 independent runnings.

Algorithm	Time consuming (s)	Number of function evaluation
GA	75.20	10670
SA	31.31	10200
PSO	55.22	26670

for continuous problems, while GA and SA are first proposed for discrete problems. Hence, compared with continuous problems, GA and SA are more suitable for discrete problems. Considering the consuming time, the time consumed by PSO is intermediate. Compared the performance from the Table 4.1, Table 4.2 and Table 4.3, we found that the performance of PSO is better than GA and SA regardless of continuous and discrete problems as a whole.
Besides the above outstanding performance of PSO compared with GA and SA, compared with other optimization algorithms, PSO has advantages in terms of the following issues.

Complexity for implementation

Simply speaking, there are only two processes for PSO and SA: determining the searching direction and step length along this direction at current point, updating the current point. While for GA, besides coding and decoding solutions, selection, crossover and mutation processes must be applied to each individual for generating new generation.

Complexity for adjusting the algorithm parameters

For each algorithm, there are several parameters to adjust as listed in Table 4.4. The number of parameters to adjust for GA is a little larger than the other three algorithms.

Besides heuristic optimization algorithms, classical gradient based optimization algorithms (Snyman [2005]) are also useful approaches to solve optimization problems, e.g. Steepest Descent Method, Conjugate Gradient Method, Penalty Function Method, and widely used Levenberg-Marquardt method. All of them are local algorithms, i.e. the optimal solutions found by them are local optimum. Due to it, they are easy to trap into local optimum and are very sensitive to initial values of variables to be optimized. Another disadvantage is that they have difficulties to solve discrete or mixed optimization problems where part of or the whole variable set that need to be optimized are integers, as these methods require differentiation information of optimization functions. In contrast, PSO does not require differentiation information and can deal with any type of data. It is suitable to solve continuous, discrete or mixed optimization problems. In addition, it is not sensitive to initial values of variables and thus does not need

Table 4.4: Parameters for each heuristic optimization algorithm

Algorithms	Parameters to adjust
PSO	size of population
	acceleration coefficient
	initial and end weight
SA	neighborhood generation strategy
	initial and end temperature
	cooling strategy
	times of temperature cooling
	number of iterations at certain temperature stage
GA	size of population
	code
	selection method,
	crossover method and parameter
	mutation method and parameter
ACO	size of neighborhoods
	jumping length
	jumping length decrease coefficient

priori knowledge of optimization problems.

4.5 Conclusion

This thesis dedicates to apply optimization techniques in the domain of agriculture and forestry, i.e. to apply optimization techniques to solve special issues in agronomy, so as to optimize the yield, taking into account of plant growth and plant architecture for the selected compartment. It is not the research objective to develop or improve optimization algorithms As PSO is versatile for a wide area application with good performance (Kennedy and Eberhart [2001]), all the optimization problems are solved by PSO in this thesis.

Part III

Optimization applications

Chapter 5

Model analysis on a virtual Corner model through optimization

Hallé and Oldeman [1970] classified plant architectures into 23 classes with the criteria concerning inflorescence position, axis growth pattern and differentiation, and branching patterns. Corner model (or Holtum model) is one of the 23 classed of plant architectures, which concerns unbranched, single vertical axis plants with lateral inflorescences, as shown in Fig.5.1. The crops that belong to the Corner model are sunflower, tomato, maize, wheat with no tiller, etc. These crops are cultivated mainly for the purpose of fruits for human or animal consumption. Therefore, in this chapter, we investigate an optimization problem with the objective of maximization of fruit yield of the Corner model. The variables that we optimize are the endogenous factors that affect fruit yield. Fruit yield is the result of the cooperation of many complex processes (e.g. source-sink competitions), involving almost all plant endogenous factors. Moreover, the topological information of fruits, e.g. number of fruits on plant and their positions, is also a crucial factor. Hence, to investigate factor effects on fruit yield more clearly, we separate the factors into different groups according to the organ type, and study their effects group by group, as shown in Fig.5.2 that illustrates the flow of the optimization investigation and the factors investigated in each step. Even though we investigate the optimization problem of maximization of fruit yield based on the Corner model that has the simplest architecture among the 23 plant architecture classes, the results derived from simulation and optimization give us insights of source-sink competitions during plant growth, in order to obtain optimal fruit yield.
The topological and physiological information of the Corner plant that we investigate in this chapter is as follows.

- Plant has 20 metamers consisting of a leaf and a fruit if it exists.
- Plant growth terminates at the 26^{th} growth cycle.
- Expansion duration of leaves and internodes is seven growth cycles, and fruits expand for six growth cycles. It guarantees that at the end of plant growth, all

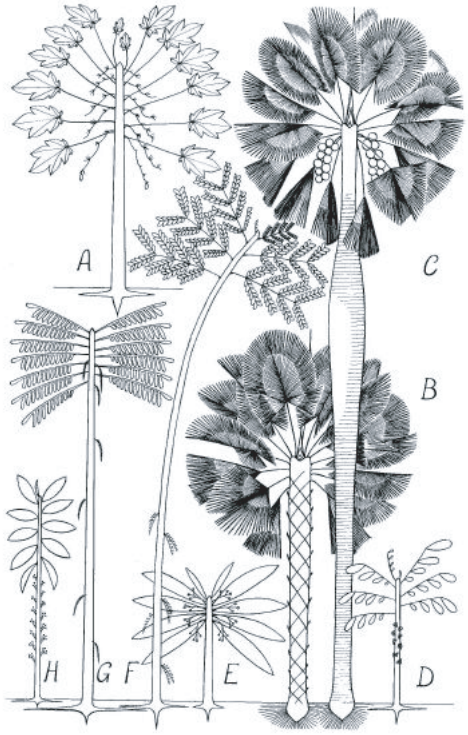


Figure 5.1: Corner model architecture (from Hallé [2005]).

organs stop expansion and at least one organ individual of each organ type has the possibility to expand at the previous growth cycle.

- Functioning duration of a leaf is 10 growth cycles.
- Fruits appear continuously, i.e. if the first fruit appears at certain metamer, other fruits will appear at the following metamers continuously.
- The secondary growth of internodes is ignored.
- The sink strength of internode is 0.2, relative value to the sink strength of leaf which is 1.
- The maximum physiological age of the plant is 1, as the Corner model has a single stem without mutation. The notation of fruit sink strength P_1^f is thus abbreviated to P^f in this chapter.

Fig.5.3 illustrates the plant that we investigate, called control plant hereafter, and the definition of variables that will be used frequently in this chapter.

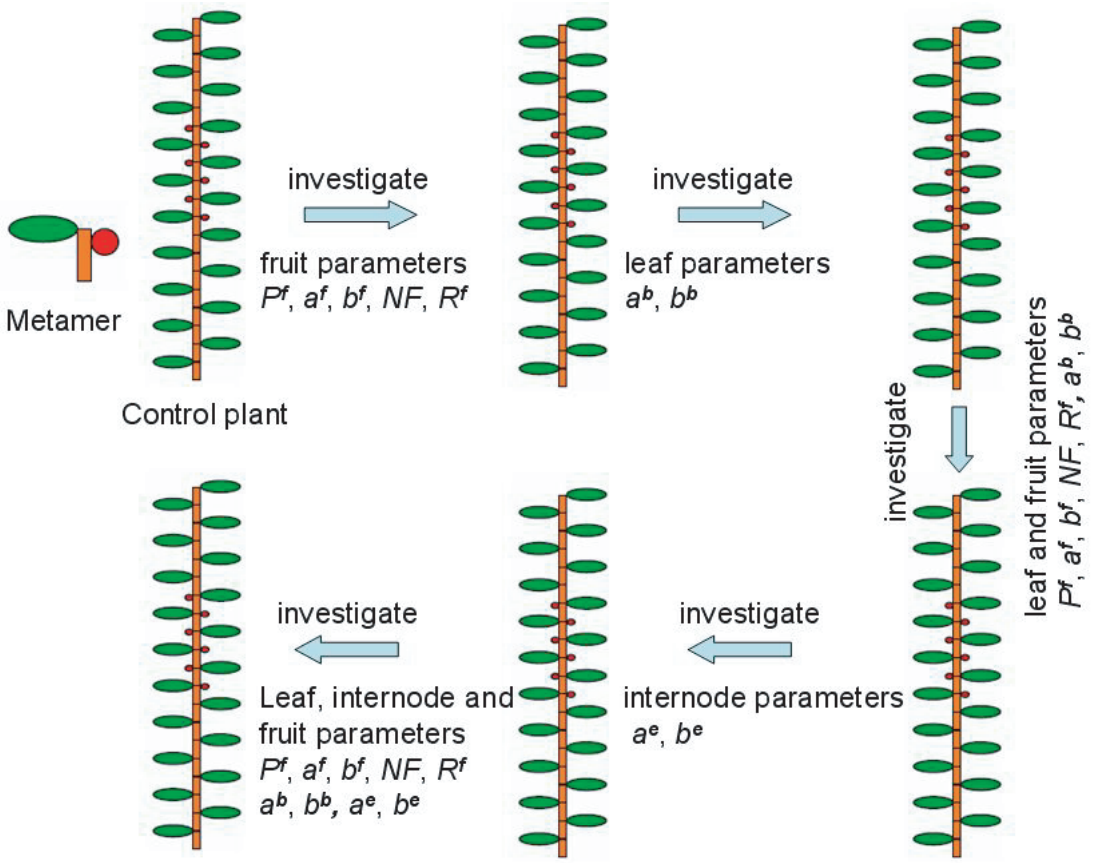


Figure 5.2: Flow of the optimization investigation. P^f is fruit sink strength; a^o and b^o are the coefficients of the sink variation function (Beta function), o being leaf (b), internode (e), and fruit (f); NF is the number of fruits on plant; R^f is the first fruit position from the stem bottom;

5.1 Effect of fruit parameters on fruit yield

The factors that affect fruit yield and are related to fruits are fruit sink strength and sink variation, number of fruits on plant, and the position of the first fruit (hereafter, the position of the first fruit is abbreviated to fruit position). In this section, we investigate the effect of these factors on fruit yield, and fix other factor values. Beginning from the relative simple plant, we first assume that the sink variations of all organs are constant. And then, we investigate fruit factors with variable fruit sink variation.

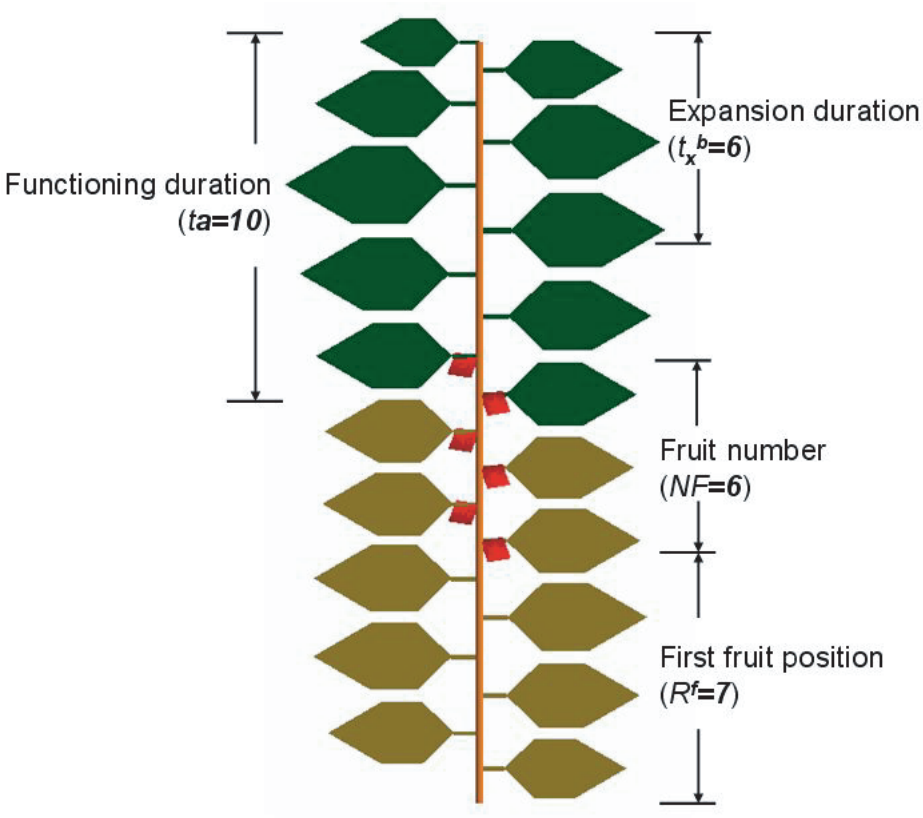


Figure 5.3: Control plant and the variable definitions that will be used frequently in this chapter.

5.1.1 Constant fruit sink value

We choose three different fruit sink strengths to investigate the effect of the fruit sink strength on fruit yield, with respect to fruit number and their positions. The simulation results are shown in Fig.5.4. As we assume that fruits appear continuously on metamers, the feasible range of the position of the fruit that first initiates on a plant is from 1 to $NGU - NF + 1$, NGU representing the maximal number of metamers in the plant and NF representing the number of fruits. For example, if there are 19 fruits in the plant, there are only two available positions for the first fruits: the first metamer and the second metamer from the bottom of the main axis. Hence, the simulation results of fruit yield for the plant with 19 fruits contain only two points, as shown in the curve marked in full diamond with dash line in Fig.5.4.
We analyze the simulation results shown in Fig.5.4 from three aspects as follows.

- Given number of fruits in the plant. If the fruit sink strength is not big, $P^f = 1$ or 10 for instance, the optimal fruit yield is obtained when fruits are at the top of the plant (i.e. the last fruit is at the last metamer of the main axis), as shown in

Fig.5.5 where the relation between the fruit position corresponding to the maximal fruit yield and fruit number is linear when $P^f = 1$ or $P^f = 10$. The crop species that has this inflorescence is cotton (Hanan and Hearn [2003]). While if the fruit sink strength is too big, to get the maximal fruit yield, the position where the first fruit appears depends on the number of fruits that the plant has. When $P^f = 200$, the relation between the fruit position corresponding to the maximal fruit yield and the number of fruits are not linear. The crop species that coincide with this phenomenon are maize and sunflower. For maize, the fruits are around the middle of the main axis and the last metamer does not bear any fruit (Guo et al. [2006]). Whereas for sunflower which has only one fruit, the fruit is at the top (Monograph [2001]).

- Given fruit sink strength. If the fruit sink strength is too big, $P^f = 200$ for instance, more fruits make the plant grow under more stressed condition due to fruit competition against source organs (leaves) and make the plant die, as shown in Fig.5.4(c) where the curves are very close to each other when fruit number is more than five, representing that more fruits do not enhance fruit yield. Three 3D examples of plants with different fruit numbers are shown in Fig.5.6, from where we see that the plant with the largest fruit number terminates growth.

- Considered the plant with only one fruit, the optimal fruit yield can be obtained if the fruit is at the top and its sink strength is big. The corresponding crops are sunflower (Monograph [2001]). The simulation results of the fruit yield with respect to the fruit position and the fruit sink strength are shown in Fig.5.7.

5.1.2 Variable fruit sink value

Besides fruit sink strength, fruit number and their positions, the effect of fruit sink variation on fruit yield is studied in this section. Three cases of fruit sink variation are chosen, as shown in Fig.5.8. For the first case, fruits mainly absorb biomass at the beginning of the expansion duration. For the second case, fruits obtain the biomass mainly in the middle period of the expansion. And for the last case, fruits obtain the biomass with a long time delay.
We analyze the simulation results of the maximal fruit yield and the corresponding fruit position with respect to fruit number as shown in Fig.5.9 from the following aspects.

- Relation between the maximal fruit yield and fruit position

  Generally speaking, to get the maximal fruit yield, fruits should be at the top, no matter what the fruit sink strength is and no matter how fruits expand, as shown in Fig.5.9(b),5.9(d) and 5.9(f) where the relation between the number of fruits and the first fruit position corresponding to the maximal fruit yield is linear. The

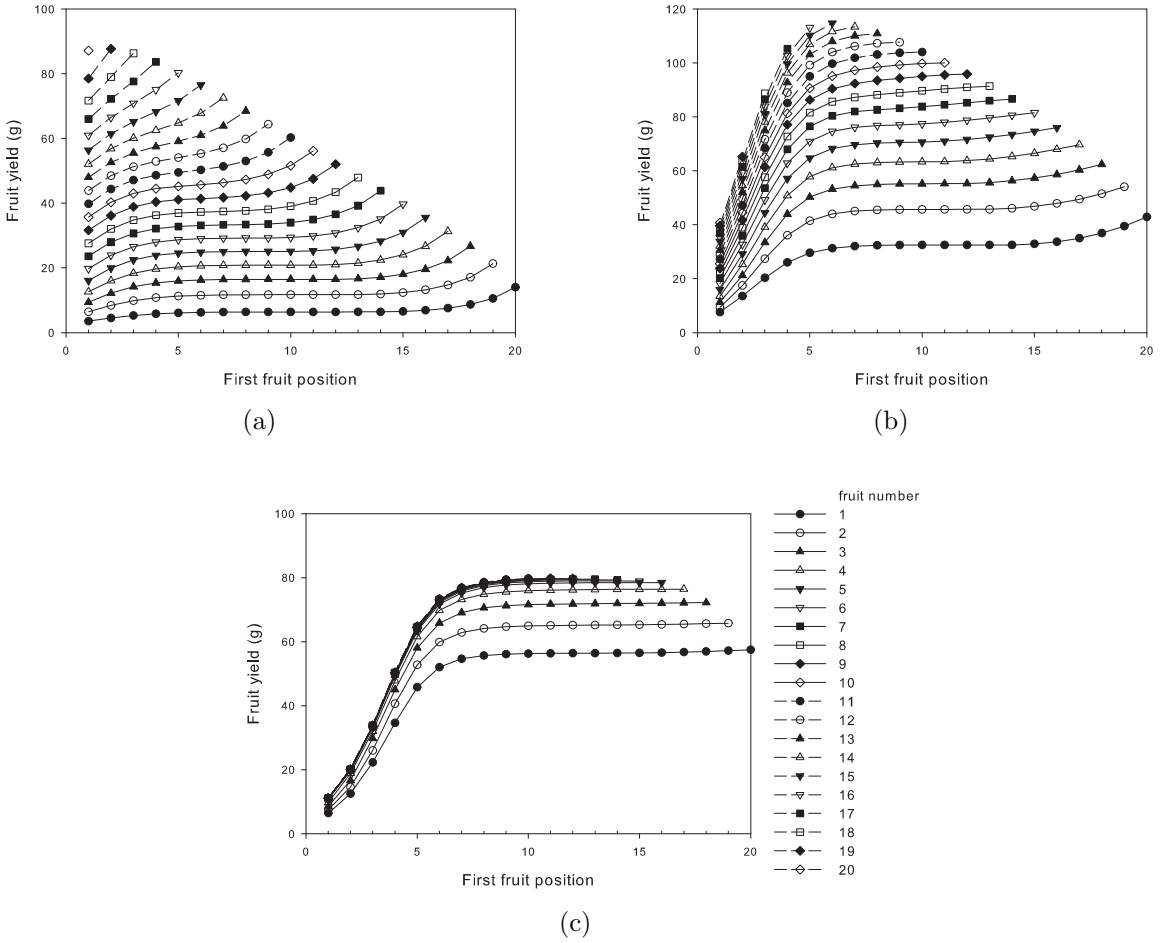
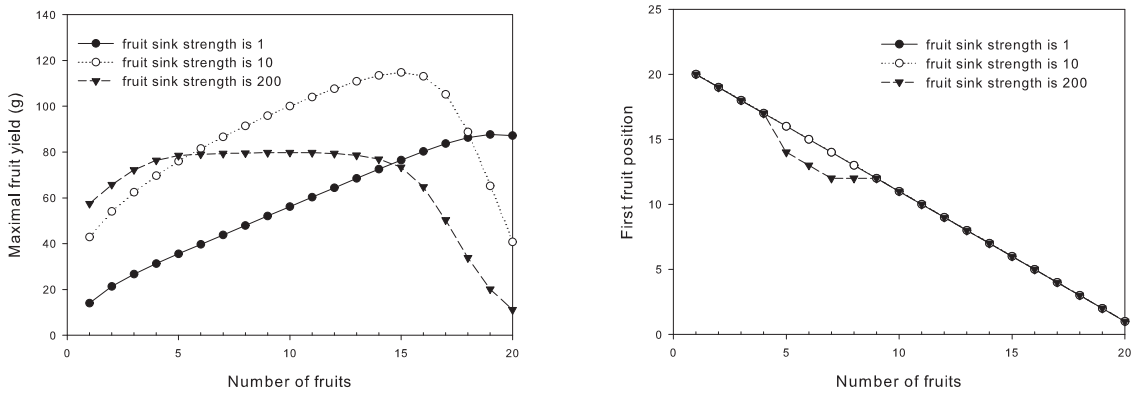


Figure 5.4: Simulation results of fruit yield with respect to fruit sink strength (P^f), fruit number (NF) and fruit position (R^f). (a) $P^f = 1$ (b) $P^f = 10$ (c) $P^f = 200$.

exceptional cases happen when the fruit sink strength is 200 and the number of fruits is from 7 to 10. For these cases, the maximal fruit yield is obtained without fruits at the top, as shown in the curve marked with full triangle in dash line in Fig.5.9(f) where the relation between the number of fruits and the first fruit position corresponding to the maximal fruit yield is not linear.

According to the GreenLab principles as expressed by Eq.2.5, fruit yield is not only dependent on its sink value, but also on plant biomass. Big fruit sink strength, on one hand, enhance the ability of fruits to obtain biomass. On the other hand, it has negative effect on the development of other organs, especially source organs (e.g. leaves). The maximal fruit yield and the corresponding fruit position are the result of the balance between sources and sinks.



(a) (b)

Figure 5.5: Results of the maximal fruit yield and the corresponding fruit position with respect to fruit number.

- Effect of fruit sink variation on fruit yield

  For a given fruit strength, the maximal fruit yields of plants with different fruit sink variations are similar. However, the fruit yield for the plant with fruits having long time delay is a little more for a given number of fruits. Moreover, for the plants with large fruit sink strength, fruit yields are very similar within certain range of fruit numbers, as shown by the plateaus in Fig.5.9(e). As analyzed in the previous section, fruits with large sink strength inhibit plant growth. Compared with sink strengths of other organs ($P^a = 1$ for leaves and $P^i = 0.2$ for internodes), fruit sink strength is so big that it makes fruits the dominant organs, and the development of other organs, especially leaves, is inhibited.

  Considered the plant with many fruits, to get the maximal fruit yield, fruits should expand with a delay, as shown in Fig.5.9(a), 5.9(c) and 5.9(e) where fruit yields with the coefficients of sink variation like $a^f = 5$ and $b^f = 2$ are better than the others. It is to insure that plant grows without more biomass consumers (sinks) at the beginning of plant growth, in order to let leaves develop and then to produce more biomass with more leaf surface area. It is also a reflection of the trade-off between sources and sinks.

5.1.3 Optimization of fruit yield on fruit factors

The optimal results of the optimization problem of maximization of fruit yield by the Particle Swarm Optimization algorithm are given in Table 5.1. The variables that are optimized are fruits factors : number of fruits and their positions, fruit sink strength and

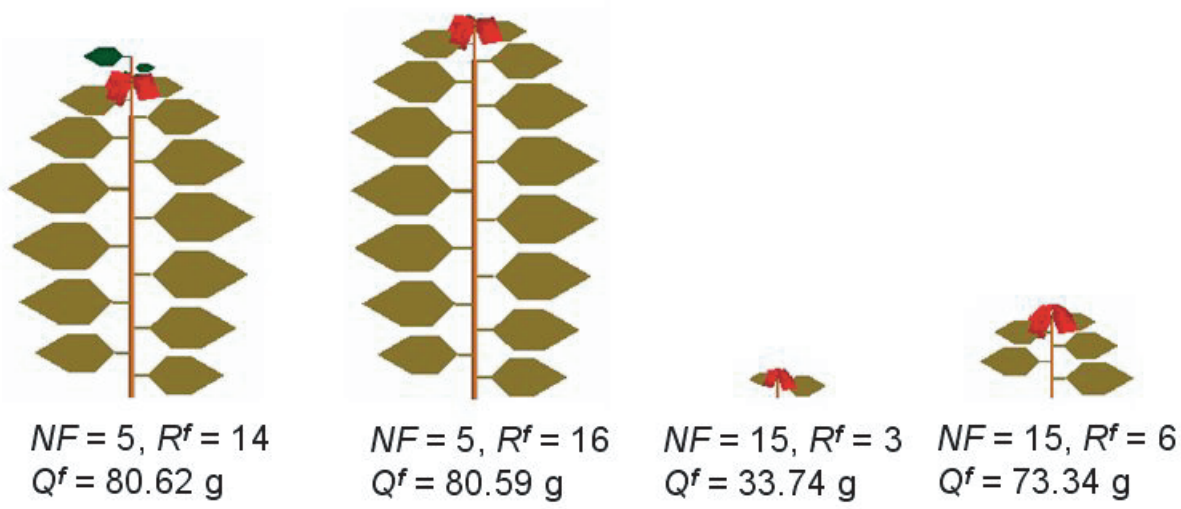


Figure 5.6: 3D images of the control plant with different fruit numbers (NF) and different fruit positions (R^f), fruit sink strength (P^f) being 200. Q^f represents fruit yield.

Table 5.1: Optimal results of fruit factors for the optimization problem of maximization of fruit yield.

Parameter	Searching space	Optimal value
Fruit sink strength (P^f)	$[0, 1000]$	16.70
Coefficients of fruit sink variation function (a^f, b^f)	$[0, 25]$	11.88, 1.97
Number of fruits (NF)*	$[1, 20]$	20
First fruit position (R^f)*	$[1, 20]$	1

* the parameter type is integer.

the sink variation. The optimal sink variation of fruits corresponding to the optimal results listed in Table 5.1 is shown in Fig.5.10. With the optimal parameter values, the optimal fruit yield is 134.61 g. The optimal results coincide with the analysis conclusions through simulation in the previous sections: with the consideration of the balance between sources and sinks, to get the maximal fruit yield, fruits are at the top (i.e. the last fruit is at the last metamer of the main axis) and fruits expand with a delay.

5.2 Effect of leaf factors on fruit yield

The endogenous parameters related to leaves in GreenLab are leaf sink strength and its sink variation. In GreenLab, the sink strengths of organs are relative values, taking

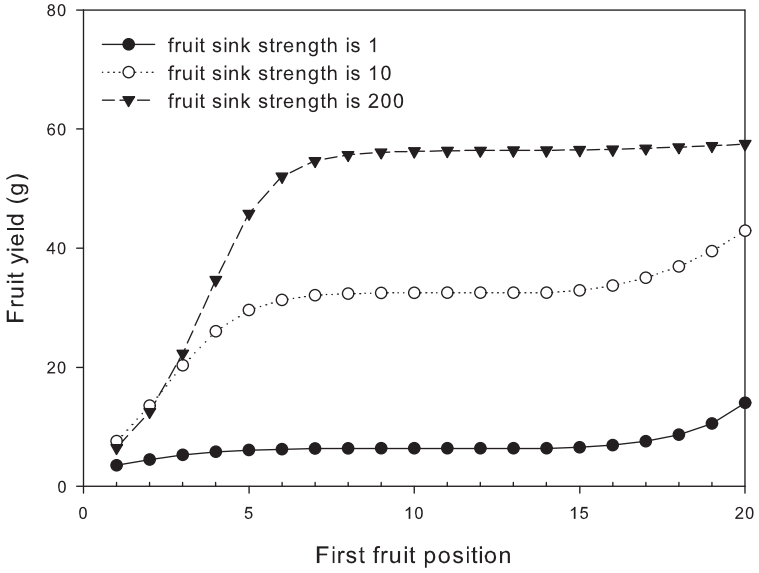


Figure 5.7: Simulation results of the maximal fruit yield with respect to the fruit position and fruit sink strength. Only one fruit is in the plant.

leaf sink strength as reference. Leaf sink strength is thus set to one, which is fixed. Hence, in this section, we study the effect of leaf sink variation on fruit yield. The other parameters of the control plant are the same as described at the beginning of the chapter, especially the sink value of internodes are assumed to be constant in this section. First, the effect of leaves on the growth behavior of plant with no fruits is analyzed based on mathematic equations in GreenLab. And then the effect of leaves on fruit yield is analyzed through optimization.

5.2.1 Mathematical analysis

To study the effect of leaves on the growth behavior of a plant, we choose two expansion modes for leaves: (1) immediate expansion within one growth cycle and (2) the leaf sink value during the whole expansion duration is constant.
As the control plant belongs to the Corner model, the maximal physiological age P_m is 1. Moreover, there is no secondary growth for the control plant. Hence, the plant demand at growth cycle t is equal to the plant demand for the primary growth, which is given by Eq.2.4 in Chapter 2. We rewrite it here.

$$D_{pg}(t) = \sum_{o} \sum_{j=1}^{\min(t,t_x^o)} N_1^o(t-j+1)p_1^o(j) \tag{5.1}$$

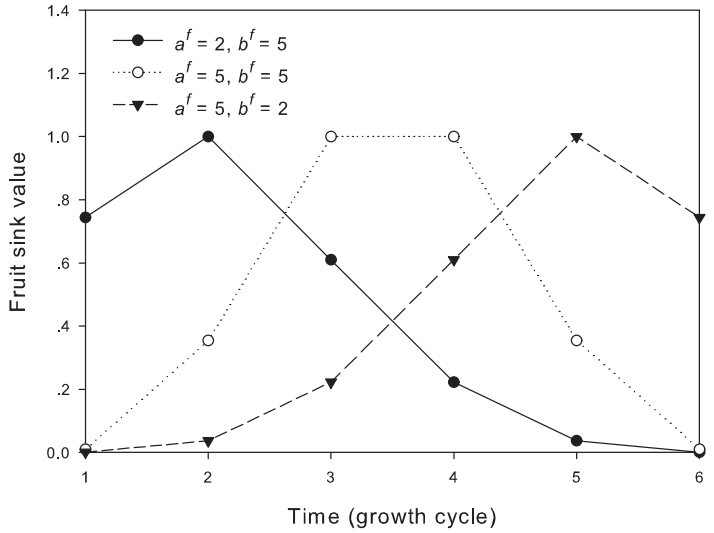


Figure 5.8: Three cases of fruit sink variation.

As only one metamer appears at each growth cycle, each metamer bears only one leaf and the sink variation of internodes is constant, $N_1^o(t) = 1$ and $p_1^o(j) = P_1^o$. Hence, the plant demand for the two modes of leaf expansion mentioned above are given by

$$D_{pg1}(t) = P_1^b + t_x^e \cdot P_1^e \quad (5.2)$$

$$D_{pg2}(t) = t_x^b \left(P_1^b + P_1^e\right) \quad (5.3)$$

where D_{pg1} represents the demand of plant with leaf expansion mode 1, and D_{pg2} the demand of plant with leaf expansion mode 2; as the description of the control plant, the expansion duration of leaves and internodes are the same. For the leaves with immediate expansion, its sink value is $p_1^b = P_1^b$, and $p_1^b(j) = 0, j > 1$.
The plant biomass production at growth cycle t for the control plant is calculated by Eq.5.4.

$$Q(t) = E(t)\mu Sp\left(1 - exp\left(-\frac{k}{Sp \cdot slw}\sum_{i=1}^{ta}\sum_{j=1}^{i} p_1^b(j)\frac{Q(t-i+j-1)}{D_{pg}(t-i+j-1)}\right)\right) \quad (5.4)$$

Suppose $l = i - j + 1$, Eq.5.4 is rewritten as expressed by Eq.5.5.

$$Q(t) = E(t)\mu Sp\left(1 - exp\left(-\frac{k}{Sp \cdot slw}\sum_{l=1}^{ta}\frac{Q(t-l)}{D_{pg}(t-l)}\sum_{i=l}^{t_x^b} p_1^b(i-l+1)\right)\right) \quad (5.5)$$

Substitute Eq.5.2 and Eq.5.3, which replaces D_{pg}, to Eq.5.5, we get the biomass of plant with two leaf expansion modes, denoted by Q_1 and Q_2 respectively.

$$Q_1(t) = E(t)\mu Sp\left(1 - exp\left(-\frac{k}{Sp \cdot slw}\sum_{l=1}^{ta}\frac{Q(t-l)}{P_1^b + t_x^e P_1^e}P_1^b\right)\right) \quad (5.6)$$

$$Q_2(t) = E(t)\mu Sp\left(1 - exp\left(-\frac{k}{Sp \cdot slw}\sum_{l=1}^{ta}\frac{Q(t-l)}{P_1^b + P_1^e}P_1^b\right)\right) \quad (5.7)$$

According to Eq.5.6 and Eq.5.7, we found that the plant with leaves that expand uniformly during the expansion duration produces more biomass than the plant with leaves that expand immediately within only one growth cycle, except the first growth cycle when both of plants produce the same amount of biomass.

5.2.2 Analysis through optimization

Beginning from the relative simplest plant with only one fruit, we study the effect of leaf sink variation on fruit yield through optimization. Referred from the conclusions in section 5.1, if there is only one fruit on the plant, to obtain the maximal fruit yield, the fruit should be at the top, with large fruit sink strength. Hence, the fruit sink strength of the plant that we will optimize is set to be 200, and the fruit is at the top. The sink value of internodes is also identical during the whole expansion duration. We optimize fruit yield on two coefficients of leaf sink variation function, associated with different expansion modes of fruits. From the optimal results of leaf expansion modes as shown in Fig.5.11, we found that to obtain the maximal fruit yield, leaves should obtain biomass mainly at the beginning of their expansion duration, no matter how the fruit expand.

According to the mathematical analysis, under the condition of uniform expansion of internodes, the biomass of plant with uniform expansion of leaves is more than the one with leaves that expand immediately within only one growth cycle. If biomass production of plant is bigger, fruits with a given sink value can obtain more biomass. But for the control plant that we studied, the optimal result derived from the optimization procedure is that leaves obtain biomass mainly at the beginning of their expansion duration, no matter how the fruits expand. Does it conflict with the results from the mathematical analysis? The answer is no. By studying plant biomass increment at each growth cycle, we found that before the appearance of fruits, biomass increment for all plants with different leaf expansions reaches the same limit value, as shown in Fig.5.12. Under this condition, if the sink value of leaves is small, more biomass can be allocated to fruits. Hence, for the control plant that we investigated, the optimal results are that leaves expand quickly within a short duration.

Table 5.2: Optimal results of factors of leaves and fruits for the optimization problem of maximization of fruit yield.

Parameter	Searching space	Optimal value
Coefficients of leaf sink variation function (a^b, b^b)	$[0, 25]$	$2.18, 8.77$
Fruit sink strength (P^f)	$[0, 1000]$	7.02
Coefficients of fruit sink variation function (a^f, b^f)	$[0, 25]$	$14.69, 4.68$
Number of fruits $(NF)^*$	$[1, 20]$	19
First fruit position $(R^f)^*$	$[1, 20]$	2

* the parameter type is integer.

Now, we study the effect of leaf expansion on the growth of the plant with more fruits that appear before the plant biomass increment achieve the limit value. Referred from the conclusions in section 5.1, we set the fruit sink strength to be 10, and the number of fruits on the plant is 18. We optimize two coefficients of leaf sink variation function (a^b and b^b) for a plant with given fruit sink variation in order to obtain maximal fruit yield. The optimal results of leaf sink variation associated with each fruit sink variation are shown in Fig.5.13. The optimal results revealed that if fruits appear before plant biomass increment achieve a limite value (a sufficient leaf surface area), to get the maximal fruit yield, leaves should expand during their whole expansion duration to have more leaf surface area and thus to produce more biomass.

5.2.3 Optimization on leaf and fruit factors

In this section, we optimize factors of leaves and fruits together, in order to get the maximal fruit yield. The optimal results are listed in Table 5.2, and the corresponding expansion modes of leaves and fruits are shown in Fig.5.14.
The conclusion is similar: to get the maximal fruit yield, fruits expand with a long delay. Leaves expand during the whole expansion duration if biomass increment of plant has not arrived the limit value, otherwise, they expand mainly at the beginning of their expansion duration to let more biomass allocated to other organs. In a word, the optimal parameter values are the results of the trade-off between sources and sinks.

5.3 Effect of internode factors on fruit yield

In this section, we study the effect of internode factors on fruit yield. While other factors are given to the optimal value derived from the previous sections.
According to Eq.2.6 and Eq.2.23, we know that more fruit biomass can be obtained if the sink value of internodes is less. However, internodes normally expand during a period. Hence, in this section, we investigate the optimal expansion mode of internodes, in order to obtain optimal fruit yield.

Table 5.3: Optimal results of factors of all organs for the optimization problem of maximization of fruit yield.

Parameter	Searching space	Optimal value
Coefficients of leaf sink variation function (a^b, b^b)	$[0, 25]$	$2.44, 10.44$
Coefficients of internode sink variation function (a^e, b^e)	$[0, 25]$	$6.54, 22.03$
Fruit sink strength (P^f)	$[0, 1000]$	5.16
Coefficients of fruit sink variation function (a^f, b^f)	$[0, 25]$	$13.87, 10.32$
Number of fruits $(NF)^*$	$[1, 20]$	19
First fruit position $(R^f)^*$	$[1, 20]$	2

* the parameter type is integer.

According to the optimal results given by the previous sections, the plant having more fruit yield has more fruits with small sink strength and with a delay expansion, and leaves that expand during the whole expansion duration. Hence, the control plant that we studied in this section has 18 fruits with sink strength 10 and with a delay expansion ($a^f = 5$ $b^f = 2$), and leaves that expand uniformly. The optimal results of internode expansion is shown in Fig.5.15, which expand with a long delay. It is to make more biomass allocated to other organs, especially leaves, and to let leaves develop first, which will result in more leaf surface area and thus produce more plant biomass at the following time. The crop species that coincides with this phenomenon that internodes expand with a long delay is sunflower (Monograph [2001]). At the beginning of sunflower growth, only leaves appear, and after a long time, the internodes begin to elongation. Moreover, internodes expanding with a delay and with a short period will result in less biomass allocated to themselves, as normally their sink values are relatively small. If internodes are solid, their lengths will be short and the plant height is small, which is not good for leaves to intercept light. However, plants are intelligent. They produce internodes like pipe with long length and thin wall. According to the mechanical theory, this shape is good for the stiffness of the stem.

5.4 Optimization on all endogenous factors of organs

In this section, we optimize all the factors that are already studied separately in the previous sections, in order to obtain maximal fruit yield. The optimal results are listed in Table 5.3 and the corresponding optimal expansion of organs are shown in Fig.5.16. The optimal fruit yield is 155.50 g.

5.5 Conclusion

To get the most yield of fruits, the expansion rate of organs, sink value of fruits, number and positions are summarized as follows.
For fruits:

- if there are several fruits in the plant, the sink value of fruits should not be too large; and there is a delay for fruits to appear. This kind of crops are maize (Guo et al. [2006]) and cotton (Hanan and Hearn [2003], Ritchie and Bednarz [2007]).
- if there is only one fruits in the plant, the sink value of the fruit should be as large as possible, and it is on the top of the plant. This kind of crops are sunflower (Monograph [2001]).
- fruits should not appear too early or should expand with a delay. It is consistent with the results found by Jones et al. [1996], who investigated cotton and found that early flower removal does not reduce the yield while later flower removal does.

For internodes:

- there is a delay for internodes to expand, to make more biomass allocated to other organs, especially leaves on one hand. On the other hand, expansion with a delay will enhance the stiffness of the stem by stretching the length and thinning the wall of the pipe shaped internode.

For leaves:

- expansion modes of leaves depend on the following condition: if biomass production of plant is similar, no matter how leaves expand, leaves should absorb at the beginning of the expansion duration. Otherwise, leaves should absorb biomass during the whole expansion duration.

Through analysis of the optimal results based on GreenLab, we understand how plant endogenous factors affect fruit yield. We found that the results derived in this chapter are consistent with the growth behavior of some crops in reality. On one hand, it revealed that after nature selection and a long time breeding, crops tend to achieve optimal status; on the other hand, it indicated that the GreenLab model is reliable and reasonable for simulating plant growth.
It gave us insights of the relation between sources and sinks during plant growth by analyzing the optimization results on Corner model. The plant growth behavior and the effect of the source-sink dynamics on plant growth will be investigated deeply for different plant species and different optimization objectives in the following chapters.

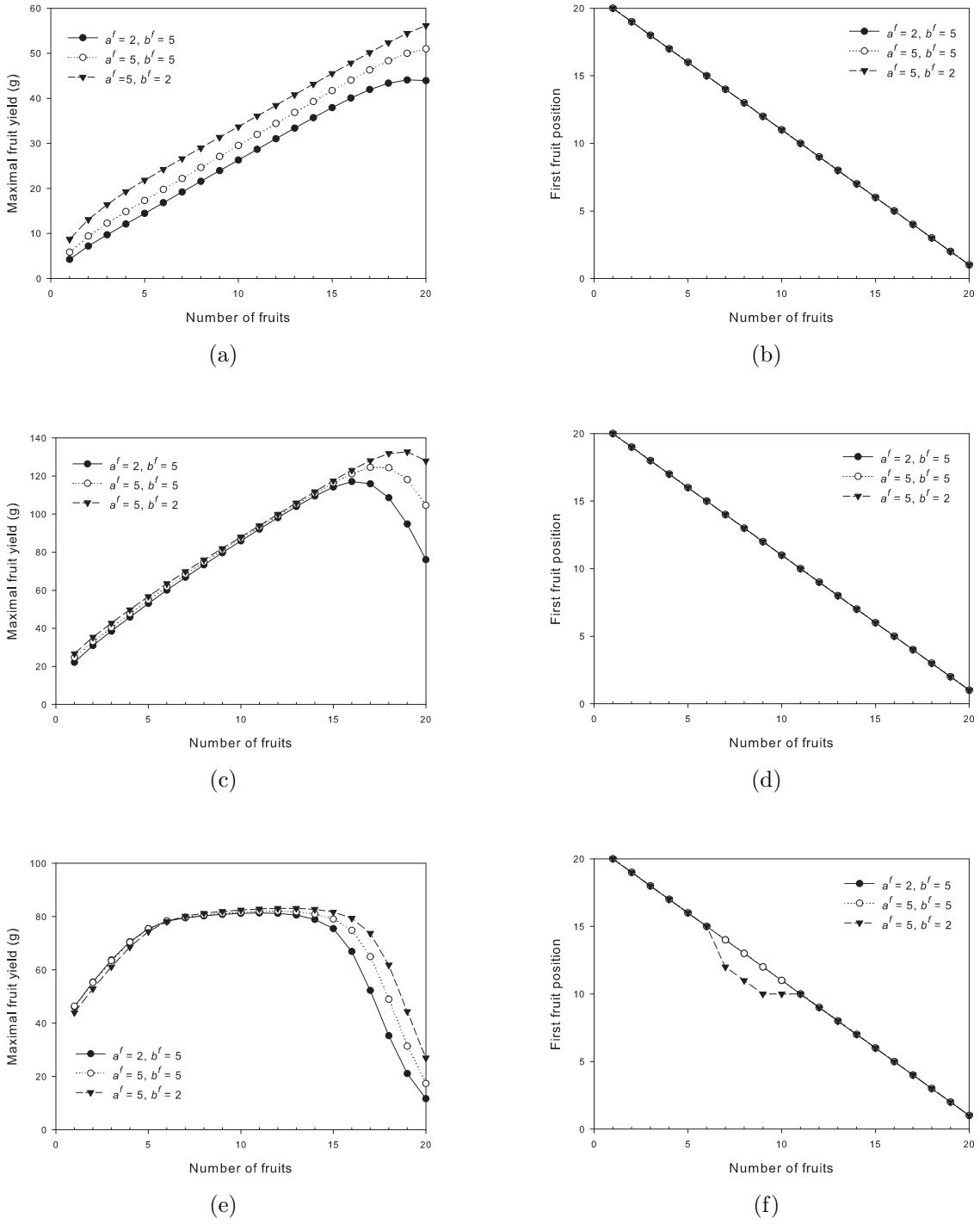


Figure 5.9: Results of the maximal fruit yield and the corresponding first fruit position with respect to number of fruits, where fruit sink strength is (a) (b) 1 (c) (d) 10 (e) (f) 200.

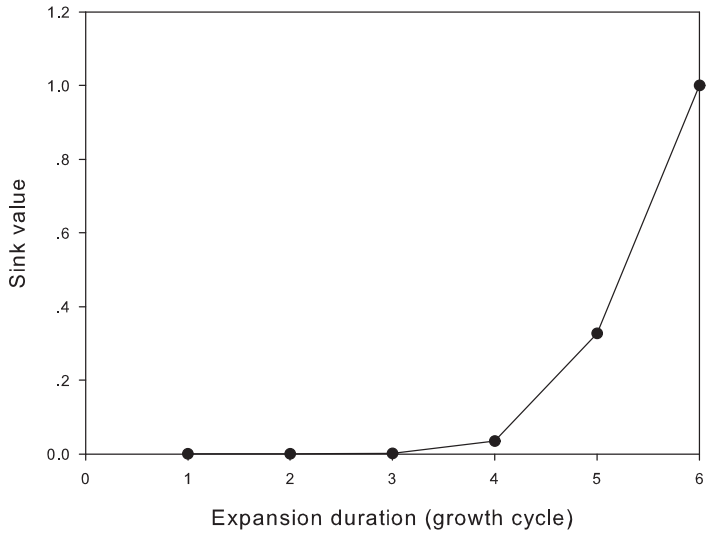
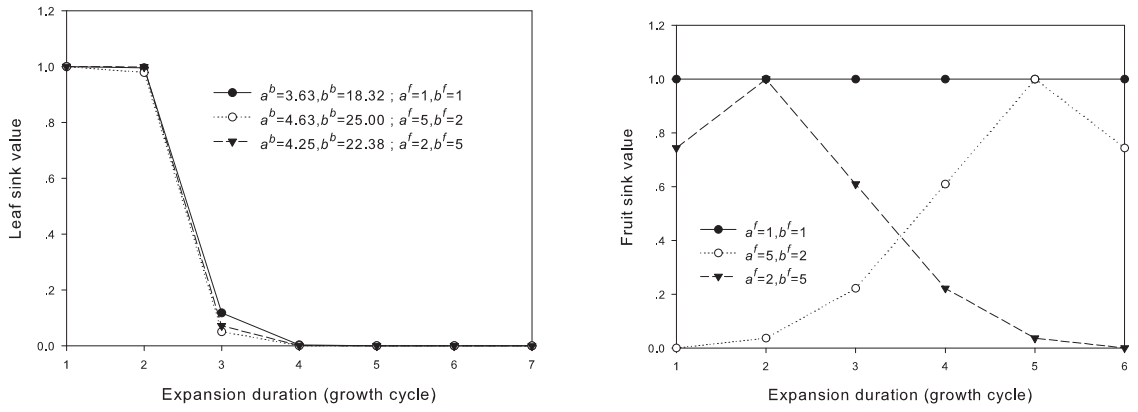


Figure 5.10: Optimal sink variation of fruits.



(a) (b)

Figure 5.11: Optimal sink variation of leaves associated with different fruit sink variations, for the optimization problem of maximization of fruit yield of the plant with only one fruit at the top. (a) Optimal sink variation of leaves (b) The corresponding sink variation of fruits. The optimal fruit yield of the plant is 59.21 g with leaf expansion $a^b = 3.63, b^b = 18.32$ and fruit expansion $a^f = 1.00, b^f = 1.00$, 48.68 g with leaf expansion $a^b = 4.63, b^b = 25.00$ and fruit expansion $a^f = 5.00, b^f = 2.00$, and 52.34 g with leaf expansion $a^b = 4.25, b^b = 22.38$ and fruit expansion $a^f = 2.00, b^f = 5.00$.

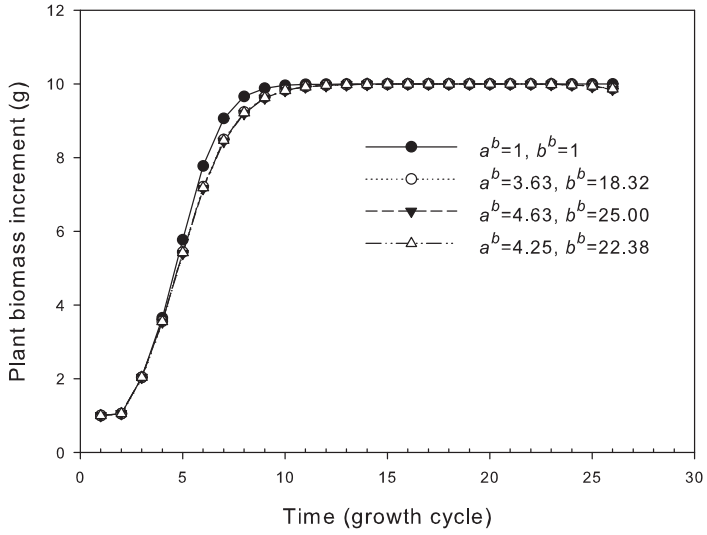


Figure 5.12: Biomass increment of the plant without fruits associated with different modes of leaf expansion.

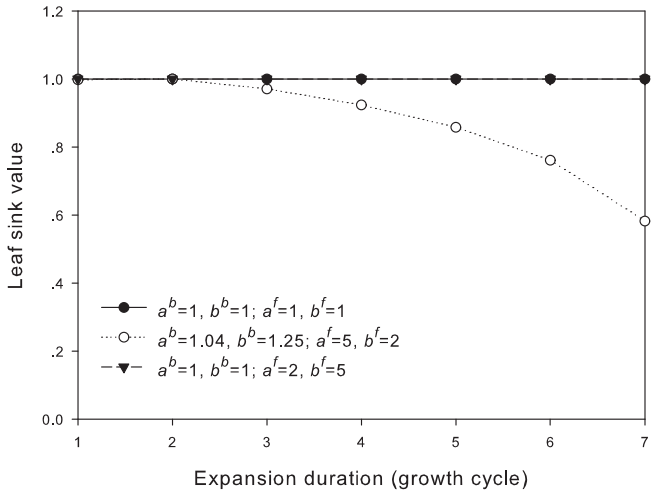


Figure 5.13: Optimal sink variation of leaves associated with different fruit sink variations as shown in Fig.5.11(b), for the optimization problem of maximization of fruit yield of the plant with 18 fruits at the top. The optimal fruit yield of the plant is 88.75 g with leaf expansion $a^b = 1.00, b^b = 1.00$ and fruit expansion $a^f = 1.00, b^f = 1.00$, 131.87 g with leaf expansion $a^b = 1.04, b^b = 1.25$ and fruit expansion $a^f = 5.00, b^f = 2.00$, and 108.60 g with leaf expansion $a^b = 1.00, b^b = 1.00$ and fruit expansion $a^f = 2.00, b^f = 5.00$.

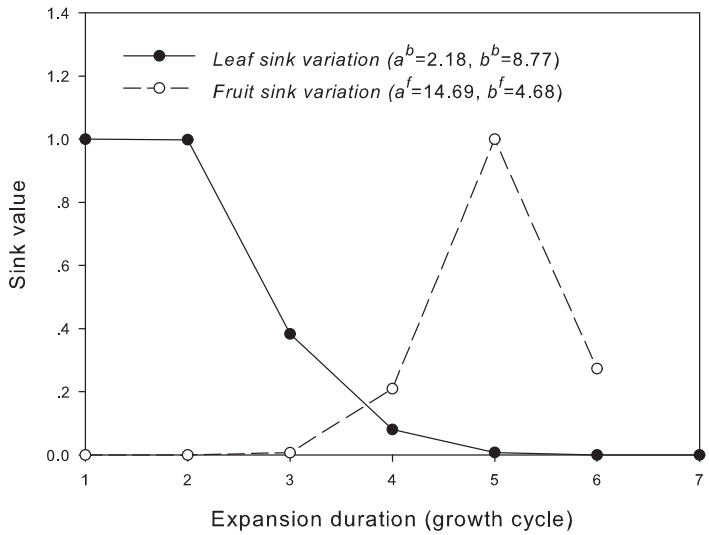


Figure 5.14: Optimal sink variation of leaves and fruits, for the optimization problem of maximization of fruit yield. The optimal fruit yield of the plant is 133.02 g.

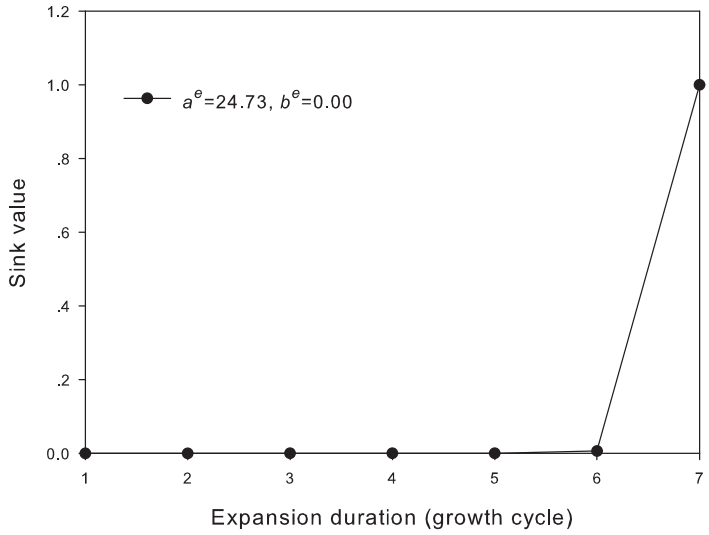


Figure 5.15: Optimal sink variation of internodes, for the optimization problem of maximization of fruit yield. The optimal fruit yield of the plant is 140.99 g.

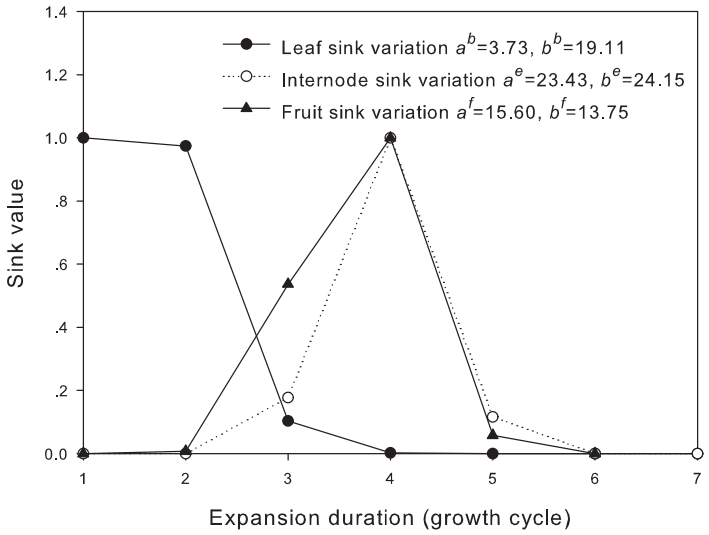


Figure 5.16: Optimal sink variation of all organs, for the optimization problem of maximization of fruit yield. The optimal fruit yield of the plant is 155.50 g.

Chapter 6

Genetic analysis on the GreenLab parameters

Geneticists dedicate themselves to produce new genotype of species with expected performance (i.e. high tolerance of water deficit, insect resistance, high yield, etc), by crossing genes of species variants. However, in millions years of domestication, the genetic basis of species is narrowed, which reduce genetic variation of species (Tanksley and Susan [1997]). One way to overcome this shortage is to use genes of wild ancestor plants, to avoid missing unused genes that are favorable for enhancing quantitative traits (e.g. yield) (Tanksley and Susan [1997], Zamir [2001]).

Introgression line (IL) population is a novel approach for the quantitative trait loci (QTL) studies, which can overcome the limitations and drawbacks of other approaches, e.g. backcross, F_2/F_3, recombinant inbred populations: overshadowing effect on QTL, time-consuming, laborious and even need chance (Eshed and Zamir [1995], Zamir [2001]). ILs are a set of nearly isogenic lines to the recipient genotype, each line containing a single homozygous restriction fragment length polymorphism (RFLP)-defined chromosome segment. The wild species genome can be completely represented by gathering all the lines. Any phenotypic difference between an IL plant and the parent only results from the introduced single chromosome segment.

So far, the relation between genetic and phenotypic variation is generally built in a statistical way without using plant growth models based on physiological knowledge. Researchers begin to realize that plant models can be essential tools to bridge genes with quantitative phenotypes. Quantitative traits such as yield are results of complex dynamic processes including source-sink dynamics. Plant architecture influence yield both in terms of carbon assimilation and allocation to different parts (de Reffye et al. [2008]). Meanwhile, the quantity of assimilation determines the activity of organs (either active, or dormant, or dead). Hence, plant models combining plant architecture and ecophysiological processes are suitable for the study of the relation between genetic and phenotypic variation.

In this chapter, we first estimate the GreenLab parameters that drive the growth of

individual organs with the observation data of 44 genotypes of tomato associated with 44 ILs. And then, we analyze the factors that influence the differences of quantitative traits, especially fruit yield, on the basis of the estimated parameters. At last, we optimize the factors to find the optimal tomato genotype by using the optimization algorithm PSO.

6.1 Plant material and measurements

The IL population as described in Eshed and Zamir [1995] were used in this study. The population is composed of 50 lines, the parental lines being the processing tomato inbred variety M82 (*L. esculentum*) and the inbred accession of *L. pennellii* (LA 716). Each line contains a single RFLP-defined chromosome segment of *L. pennellii*; overlapping regions between neighboring lines were selected to ensure complete representation of the wild species genome. 44 genotypes were used in this work, due to the missing of the seeds of the other 6 genotypes. For the sake of description clarity, we number 44 genotypes of tomato as follows. There are 12 chromosomes in tomato, which are labeled from 1 to 12. For each chromosome, there are several ILs, which are labeled from 1 to m, m being the number of ILs in the chromosome. If the IL is at the chromosome labeled by 1 and it is the second ILs in the chromosome, the genotype is named IL1-2. The experiment was conducted in a solar greenhouse at LangFang experiment site, HeBei province China ($116°35' E$, $39°36' N$). Seeds of the 44 ILs and their parents were collected for germination. Seedlings (having about $4-6$ leaves) were transplanted at 6 Sep 2006 (S1) into 20 cm-diameter pots, filled with organic soil. Around 30 pots for each genotype were arranged in two paralleled north-south oriented rows that were 0.5 m apart and 0.8 m between different genotypes. M82 plants were repeated nine times homogeneously arranged in East-West direction in greenhouse to evaluate the site effect on growth. Irrigation and nutrition were provided when needed, and routine plant management and greenhouse control were applied to protect crops from pest, weed or low temperature. All side shoots were pruned to eliminate the effect of branching and ease data collection and analysis, except two M82 repetitions where a branch is left. Pruned M82 variety is the control plant.
Besides seedlings, for each genotype, three plants were sampled destructively at three stages of planting: at 28 September 2006 approximately corresponding to the end of development of the main stem (S2); at 31 October approximately corresponding to the flowering stage (S3) and at 21 December corresponding to the ripening of most fruits (S4), growth duration thus lasting 107 days. Only plants in the center area of north-south direction were taken where light condition is most homogeneous. Once a plant was removed, the same position is replaced by another pot from the tips of rows to keep the same population density. The new plants were never sampled. The collected data includes fresh weight of individual leaves, internodes and truss, total dry weight of each type of organs, position of each truss and its fruit number, size of internodes

and leaves.
The phytomer appearance is controlled by thermal time in the GreenLab model for tomato. The model time step, growth cycle (GC), is thus equivalent to thermal time requirement for each phytomer appearance. The thermal time required by each phytomer appearance for each genotype is different from each other, and the environmental value for each day is the same for all the genotypes. Hence, the environmental factor in Eq.2.20 in the scale of growth cycle is different for each genotype, which is normalized by the ratio of the longevity of the genotype with the highest growth rate in growth cycle to the longevity of the considered genotype, the environmental factor for the genotype with the highest growth rate in growth cycle being 1.

6.2 Parameter estimation

In GreenLab, the parameters are classified into two categories: (i) measurable parameters, i.e. functioning duration of blades, number of organs emerged at each growth cycle, and (ii) hidden parameters that cannot be measured directly in the field, i.e. organ sink. To guarantee that GreenLab can describe the dynamics of plant growth well, it is necessary to estimate the hidden parameters through minimizing the difference between the measured data and the corresponding simulation results of GreenLab. The same set of parameters that drive the dynamic development of individual organs is estimated simultaneously by fitting with several plants of a genotype at different development stages, which is called multi-fitting. The parameters are estimated by a generalized non-linear least square method (Zhan et al. [2003]). It is observed that the leaves have a thickening process during the stage S3 and S4. Hence, in this chapter, we consider the process of leaf thickening, as well as internode secondary growth for radial increment. We use the mode *Number of leaves* to calculate the demand of biomass for internode secondary growth, the mass sink P_1^{rg} being 0. In addition, we assume that the flowering period lasts 6 growth cycles.
The fresh weight of each type of organs at each phytomer and the total weight of a population of organs for each type in a plant are used as the detail and the compartment target data respectively for parameter estimation. The simulation results of the growth dynamics for the genotype IL1-1 with the estimated parameters of GreenLab are shown in Fig.6.1. The comparison of the observation data and the simulation results with the estimated parameters of the fresh weight of leaves, internodes and fruits for all genotypes is shown in Fig.6.2.

6.2.1 Genetic analysis based on the estimated parameters

M82 is the parental genotype and has the maximal fruit yield among all 50ILs at the average level. Hence, M82 is taken as the reference genotype to compare and analyze the results of parameter estimation. As the objective of this section is to find the factors

that affect the fruit yield, the plants sampled at the final stage (S4) are considered. Even though three plants were sampled and their observation data were used to estimate the model parameters for each genotype, only one of them, whose simulation results of the organ weight with the estimated results are the closest to the observation ones, is chosen to be used for the future analysis.

We first analyze the difference of phenotypic traits and the estimated model parameters of the ILs from the M82 by the group of each chromosome, and then we analyze the specific ILs individually.

There is no significant difference of the expectation of the leaf yield for all ILs on each chromosome from M82. There is no significant difference of the expectation of the internode yield for all ILs on each chromosome from M82, except IL3 ($p < 0.05$). Considering the fruit yield, IL1, IL3, IL4, IL5, IL8, IL11 and IL12 have significant difference from M82 ($p < 0.05$), while the fruit yield of other ILs have no significant different from the fruit yield of M82. The fruit yield of each IL and M82 is shown in Fig.6.3.

Now, we only focus on the estimated parameters of the ILs whose fruit yield has significant different from M82, in order to find the reason and the factors that result in the significant difference of the fruit yield.

There are four ILs on the first chromosome. The estimated parameter values for the IL1s on the first chromosome and for M82 are listed in Table 6.1. From Table 6.1, we found that the first three IL1s (IL1-1, IL1-2 and IL1-3) possess similar parameter values, whose ranges are different from the last IL (IL1-4). Compared the means for each parameter between M82 and the first three ILs, the characteristic projection area (Sp) has the significant difference ($p < 0.05$), while the others have no significant difference. According to the equation for calculating the biomass production of a plant in GreenLab (Eq.2.20), one of the factors that determine the limitation value of the biomass production of a plant is Sp. The other factors are the light use efficiency (μ) and the environmental factor (E_t), which are similar for all genotypes, as shown in Fig.6.4. In GreenLab, the biomass obtained by each organ depends on the plant biomass production. If the plant biomass production is small, the biomass obtained by organs is also small. Even though the leaf area index (LAI) of the IL1s is bigger than the LAI of M82 as shown in Fig.6.5, the LAI saturates and the plant biomass production reaches the limitation value that is determined by Sp. As Sp for IL1s is smaller than M82, the limitation value of the plant biomass production of the IL1s is smaller than M82, as shown in Fig.6.6. Hence, the fruit yield of the IL1s is smaller than M82 under the condition that the light use efficiency and the environmental factors are similar for all genotypes. Considering the last IL (IL1-4), its Sp and the fruit sink strength (P_1^f) are bigger than the values of M82. Even though Sp and P_1^f have positive effect on fruit yield, the constant demand parameter value for secondary growth (P_0^{rg}, $P^b sg$) is bigger than the value of M82. The function of the secondary growth of plant in terms of biomass production is to obtain biomass, not to produce biomass. Hence,

Table 6.1: Estimated parameter values and the observation data of fruit yield for the IL1s and M82.

Genotype	b^b	b^e	b^f	P_1^e	P_1^f	Sp	μ	P_0^{rg}	P_{sg}^b	Q^f
M82	2.67	3.13	2.50	0.39	19.42	791.21	0.0419	0.465	0.024	590.05*
IL1-1	3.53	4.90	5.20	0.28	12.06	443.17**	0.0401	0.627	0.226	255.50
IL1-2	3.672	4.91	1.16	0.33	2.04	392.37**	0.0415	0.642	0.146	295.50
IL1-3	2.75	4.28	10.11	0.29	10.07	381.92**	0.0367	0.329	0.099	284.80
IL1-4	3.98	5.52	35.08	0.31	82.87	1528.33	0.0383	1.194	0.128	439.55

** represents that the mean for the parameter of the ILs is significant different from M82 at the 0.05 level (2-tailed);

* the value is the average value.

Table 6.2: Estimated parameter values and the observation data of fruit yield for the IL3s and M82.

Genotype[1]	b^b	b^e	b^f	P_1^e	P_1^f	Sp	μ	P_0^{rg}	P_{sg}^b	Q^f
M82	2.67	3.13	2.50	0.39	19.42	791.21	0.0419	0.465	0.024	590.05*
IL3-1	3.34	3.90	19.92	0.28	21.34	1611.83	0.0460	0.997	0.120	481.40
IL3-2	3.74	4.60	2.43	0.26	4.60	224.43	0.0640	0.610	0.099	226.60
IL3-4	3.38	3.28	1.25	0.22	0.42	559.48	0.0320	0.510	0.077	88.04
IL3-5	3.69	4.77	2.25	0.26	39.45	324.05	0.0650	0.574	0.095	432.00

[1] The seed of IL3-3 is missing.

* the value is the average value.

it is a kind of sink organ, which involves in the competition of biomass against other organs. Increases in secondary growth will result in the decrease of biomass allocated to other organs. Therefore, P_0^{rg} and P_{sg}^b have the negative effect on fruit yield. With the parameters that have both negative and positive effect on fruit yield, the fruit yield is still smaller than the one of M82 finally.

Statistically, there is no significant difference between the means for the parameters of the IL3s on the third chromosome and M82, no matter considering the group of the IL3s on the third chromosome or each IL. However, we found that the values of Sp for the IL3s are smaller than M82 except the IL3-1, as listed in Table 6.2. Moreover, for the IL3-2 and IL3-4, the fruit sink strength (P_1^f) is smaller, which also leads to the reduction of the fruit yield. Similar with the IL1-4, even though Sp and P_1^f for the IL3-1 and IL3-5, which have positive effect on fruit yield, are bigger, P_0^{rg} and P^bsg are also bigger, which are negative for fruit yield. The fruit yield for the IL3-1 and IL3-5 are still smaller than the M82 finally.

For IL4s, the mean of Sp is significantly different from M82 at the 0.05 level, and the other parameters have no significant difference. For IL5s, the parameters whose means

Table 6.3: Estimated parameter values and the observation data of fruit yield for the IL5s and M82.

Genotype	b^b	b^{e**}	b^f	P_1^{e**}	P_1^f	Sp	μ	P_0^{rg}	P_{sg}^{b**}	Q^f
M82	2.67	3.13	0.39	2.50	19.42	791.21	0.0419	0.465	0.024	590.05*
IL5-1	3.21	4.77	4.11	0.28	31.71	550.95	0.0428	0.379	0.125	478.16
IL5-2	2.95	4.66	5.43	0.32	17.42	397.97	0.0470	0.373	0.108	327.34
IL5-3	3.60	4.70	3.82	0.27	2.99	492.30	0.0405	0.896	0.140	282.90
IL5-4	5.04	4.80	6.54	0.21	49.24	651.69	0.0460	0.413	0.161	531.50
IL5-5	2.83	3.74	4.55	0.25	6.18	506.26	0.0407	0.175	0.165	307.50

** represents that the mean for the parameter of the IL5s is significant different from M82 at the 0.05 level (2-tailed);

* the value is the average value.

Table 6.4: Estimated parameter values and the observation data of fruit yield for the IL12s and M82.

Genotype	b^b	b^e	b^f	P_1^e	P_1^f	Sp	μ	P_0^{rg}	P_{sg}^b	Q^f
M82	2.67	3.13	0.39	2.50	19.42	791.21	0.0419	0.465	0.024	590.05*
IL12-1	3.51	4.65	0.38	4.79	9.53	667.87	0.0431	0.736	0.099	509.60
IL12-2	3.69	4.13	0.28	2.63	16.36	581.28	0.0427	0.361	0.091	469.70
IL12-3	3.02	4.33	0.35	1.08	2.52	850.46	0.0304	0.337	0.268	473.70
IL12-4	2.56	3.12	0.32	1.23	9.46	607.07	0.0407	0.462	0.096	478.52

* the value is the average value.

are significant different from M82 are the internode sink strength (P_1^e), the coefficient of the internode sink variation function (b^e) and the constant demand parameter for the leaf secondary growth (P_{sg}^b), listed in Table 6.3. The smaller P_1^e has the positive effect on the fruit yield, as smaller P_1^e will lead to the increase of biomass allocated to other organs. The bigger b^e and P_{sg}^b inhibit the growth of other organs, as shown in Fig.6.7. For IL5-2 and IL5-5, the inhibition effect of the parameters b^e and P_{sg}^b are not obvious, however, Sp is smaller for the IL5s. The limitation plant biomass production is thus reduced, as shown in Fig.6.8, which results in the decrease in biomass partition to fruits.

The parameters whose means are significant different from M82 are P_1^e, P_1^f and P_{sg}^b for the IL8s, P_1^f and P_0^{rg} for the IL11s. Even though the difference of the means of the parameters for the IL12s from M82 is not significant, we can find that P_1^f is smaller than M82 and P_{sg}^b is bigger as listed in Table 6.4, which both have negative effect on the fruit yield.

6.3 Parameter optimization

In this section, we optimize the model parameters, in order to obtain a tomato genotype with maximal fruit yield. All the parameters that are estimated in the previous section are optimized, except the characteristic projection area Sp and the light use efficiency μ. Zheng [2009] found that the light use efficiency of certain plant species or genotype is stable and is difficult to enhance according to field experiments. Hence, in this section, the light use efficiency for each plant genotype is kept the same as the estimated one. The characteristic projection area is related to planting density and the crown structure. It could be changed according to the crown structure and thus could be changed by gene improvement. However, in this section, we consider its value for each genotype the same as the estimated one in the first place. Its effect on plant yield will be discussed on the end of the section.

On the basis of the results of parameter estimation, each parameter has several values corresponding to genotypes. As the optimization objective is to find the parameters that most affect the fruit yield and to obtain a genotype with maximal fruit yield by gene improvement, the feasible region of each parameter is discrete, and the feasible solutions in the feasible region of each parameter are the parameter values estimated from the previous section. The optimization procedure is applied to each plant genotype, as the topological structure of each plant genotype is different. In order to record the dynamic growth information of plant, the weights and sizes of organs at different growth stages are measured. As the measurements are destructive, the plants measured are different. However, for the parameter estimation, we assume that for given genotype, the plant measured at a growth stage has the same growth behavior completely with the plant that measured at previous growth stage. Hence, the plants of certain genotype possess the same set of parameter values, even though the sampled plants are different. Moreover, at certain growth stage, several plants are sampled and measured to avoid missing some information of the genotype. Even though plants belong to the same genotype, there are differences individually. The criterion to accept the estimated parameters is minimization of the difference between the simulation results and the observation data on the average level. Hence, the simulation results of some plants of certain genotype are close to the observation data, while others are little far from the observation data. We assume that the estimated parameters can represent the growth dynamics of plants well if the simulation results with the estimated parameters are very close to the observation data. Therefore, in this section, the plant of certain genotype on which the optimization is applied is the one whose simulation results are closer to its observation data. The Particle Swarm Optimization algorithm is used to solve the optimization problem.

6.3.1 Statistical analysis of the optimal results

By analyzing the optimal results as shown in Fig.6.9 globally, we found that the distribution of the optimal fruit sink strength is significantly different from the distribution of the estimated one ($p < 0.05$), where the optimal fruit sink strength values for most of genotypes took the maximal value among the estimated values; the optimal constant demand parameter for the internode secondary growth took the minimal value for all plant genotypes; the optimal coefficient of the internode sink variation function and the optimal constant demand parameter for the leaf thickening are similar with the estimated one ($p > 0.1$); the distribution of the optimal coefficient of the leaf sink variation function and the distribution of the optimal internode sink strength are significantly different from the distribution of the estimated ones ($p < 0.05$). Even though the distribution of the optimal coefficient value of the fruit sink variation function is not significantly different from the estimated one statistically, the optimal values are more converged to its mean value.

6.3.2 Case analysis

Now we take some plant genotypes to analyze the optimal results in detail, to find the reason why the optimal results will lead to the improvement of fruit yield. Two plant genotypes are randomly selected, which are IL2-5 and IL8-1.
The leaf area index (LAI) of the estimated IL2-5 (i.e. the plant with the estimated parameter values) is shown in Fig.6.10. Its value is small during the whole time, even though its value keeps increasing. On the contrary, the value of the LAI of the optimal IL2-5 (i.e. the plant with the optimal parameter values) is large and reaches a limitation by mainly decreasing the coefficient value of the fruit sink variation function (b_1^f) and the constant demand parameter for internode secondary growth (P_0^{rg}), as shown in Fig.6.10 and Table 6.5. Smaller b_1^f is, bigger the delay of fruit expansion is. Increase of the delay of fruit expansion can stay the competition for biomass against other organs and let the LAI increase. When the LAI is large enough, fruits begin to absorb biomass. The demand of biomass for internode secondary growth is continuous during the whole time, and it increases with increasing total number of leaves alive in the plant. Hence, the reduction of its corresponding parameter (P_0^{rg}) leads to the reduction of the competition for biomass of internode secondary growth against other organs. More biomass will be allocated to leaves and piths, as shown in Fig.6.11. Considering the optimal plant, even though at certain time fruits become dominant (i.e. more plant biomass is obtained by fruits) as shown in Fig.6.11(b), the LAI is still large enough due to the long functioning time of leaves and will not affect negatively the biomass production of plant. Compared with the estimated leaf expansion strategy, the leaves of the optimal plant tend to absorb biomass at the beginning of their expansion duration. Under the condition where the LAI is large enough, the decrease of leaf sink helps fruit yield increase.

Table 6.5: Comparison of the estimated and optimal parameter values for tomato IL2-5

Parameter	b^b	b^e	b^f	P_1^e	P_1^f	P_0^{rg}	P_{sg}^b
Estimated value	2.73	3.54	20.51	0.29	8.79	1.072	0.204
Optimal value	4.45	2.90	12.46	0.25	82.87	0.00949	0.130

Taking into account tomato IL8-1, the estimated LAI value is large and reaches a limitation as shown in Fig.6.12. However, more biomass are allocated to secondary growth than leaves as shown in Fig.6.13(a). By decreasing the constant demand parameter for internode secondary growth (P_0^{rg}), the biomass allocated to leaves significantly increases, and the LAI value also increases. Similar with the results for tomato IL2-5, thanks to the long functioning time of leaves, increase in the fruit sink strength does not affect negatively the LAI value, as shown in Fig.6.13(b) and Fig.6.12. The fruit sink strength should thus be as large as possible in order to obtain more biomass. In addition, fruits begin to expand at the time when the LAI reaches the limitation. Hence, the coefficient of the fruit sink variation function decreases.

6.3.3 Correlation of parameters to yield

The correlation result of each optimized parameter value with the optimal fruit yield is listed in Table 6.6. We found that the amount of fruit yield that can be enhanced is significantly correlated to the characteristic projection area of plant (Sp), as well as the coefficient of the fruit sink variation function (b^f). The relationship between the gain of fruit yield and Sp is shown in Fig.6.14.

6.3.4 Analysis of the relation between optimal parameter values and ILs

The objective of this chapter is to find the ILs that affect the fruit yield based on GreenLab. Hence, the range of parameter values are restricted to the discrete parameter space formed by the estimated parameter values for each IL, and it is a one-to-one mapping between parameter values and ILs. Therefore, the IL associated with the optimal parameter values can be found. The distribution of the ILs with respect to the optimal parameter values of GreenLab is shown in Fig.6.15. From Fig.6.15, we found that the optimal fruit yield is not determined by a single IL, while it is the result of the interaction among several ILs. In addition, summarizing and analyzing the results shown in Fig.6.15, we found that IL6-1 and IL12-4 have no positive effect on fruit yield, as there is no optimal parameter values associated with these two ILs.

Table 6.6: Correlation of each optimized parameter to the optimal fruit yield by Pearson correlation method.

Parameter	Pearson correlation[1]	Significance
b^b	-0.066	0.644
b^e	0.056	0.696
b^f	0.321(*)	0.020
P_1^e	-0.162	0.251
P_1^f	0.175	0.215
Sp	0.939(**)	0.000
μ	-0.135	0.339
P_0^{rg}	–	–
P_{sg}^b	-0.159	0.259
Q_{opt}^f	1	–

* represents that the correlation is significant at the 0.05 level (2-tailed);

** represents that the correlation is significant at the 0.01 level (2-tailed);

[1] "–" represents that the correlation cannot be computed because at least one of the variables is constant;

Q_{opt}^f represents optimal fruit yield.

6.4 Conclusion

Under given environmental condition, if the variation of the fruit sink value does not affect negatively the LAI, which is an important factor to determine the plant biomass production, the fruit sink strength should be as large as possible in order to obtain maximal fruit yield. Moreover, the increase of the LAI value can be achieved by decreasing the competition of biomass against the source organ (leaves), e.g. decreasing the constant demand parameter for internode secondary growth (P_0^{rg}) and even expanding at the beginning of the expansion duration of leaves. The optimal fruit yield is the result of considering the optimal balance between sources and sinks.

The optimal results of the constant demand parameter for internode secondary growth (P_0^{rg}) for all tomato genotypes are the same, which is the minimal value among the estimated values. It aims to reducing the competition of biomass and to making more biomass allocate to other organs. In this chapter, we did not take the biomechanical constraint into account. If so, the optimal results of P_0^{rg} would be bigger than the one found in this chapter. However, for tomato, it is applicable that the biomechanical constraint does not be taken into account. The internode of tomato could twine kind of fence.

The optimal fruit yield for each plant genotype is obtained under the condition that the characteristic projection area (Sp) is given as the estimated one. If Sp is not fixed, the optimal value of Sp should be as large as possible, as the optimal fruit yield is significantly related to Sp. Sp is a parameter related to plant architectural. In order to improve Sp, geometrical properties of plant should be considered, e.g. the spatial distribution of leaves, leaf shape. Moreover, another factor that affects Sp is the planting density. If the planting density is high, Sp is reduced. The optimization objective in this chapter is to maximize fruit yield of an individual plant. Taking into account the effect of Sp on the total fruit yield in certain planting area, on one hand, increase in Sp will enhance the fruit yield of individual plant. On the other hand, increase in Sp results in small number of individual plants in the planting area. Therefore, if the total fruit yield in certain planting area is considered, the optimal Sp may not be as the one given in this chapter.

The work in this chapter is a primary work for the application of GreenLab to analyze the difference among plant genotypes by using the observation data of different plant genotypes. Even though the results in this chapter need the experiment validation in further step, the optimal plant growth mechanism is analyzed from the model parameter's point of view, which is consistent with the results found by researchers based on experiment observations. The optimal growth mechanism will be investigated in detail based on GreenLab for different plant species through optimization in the following chapters.

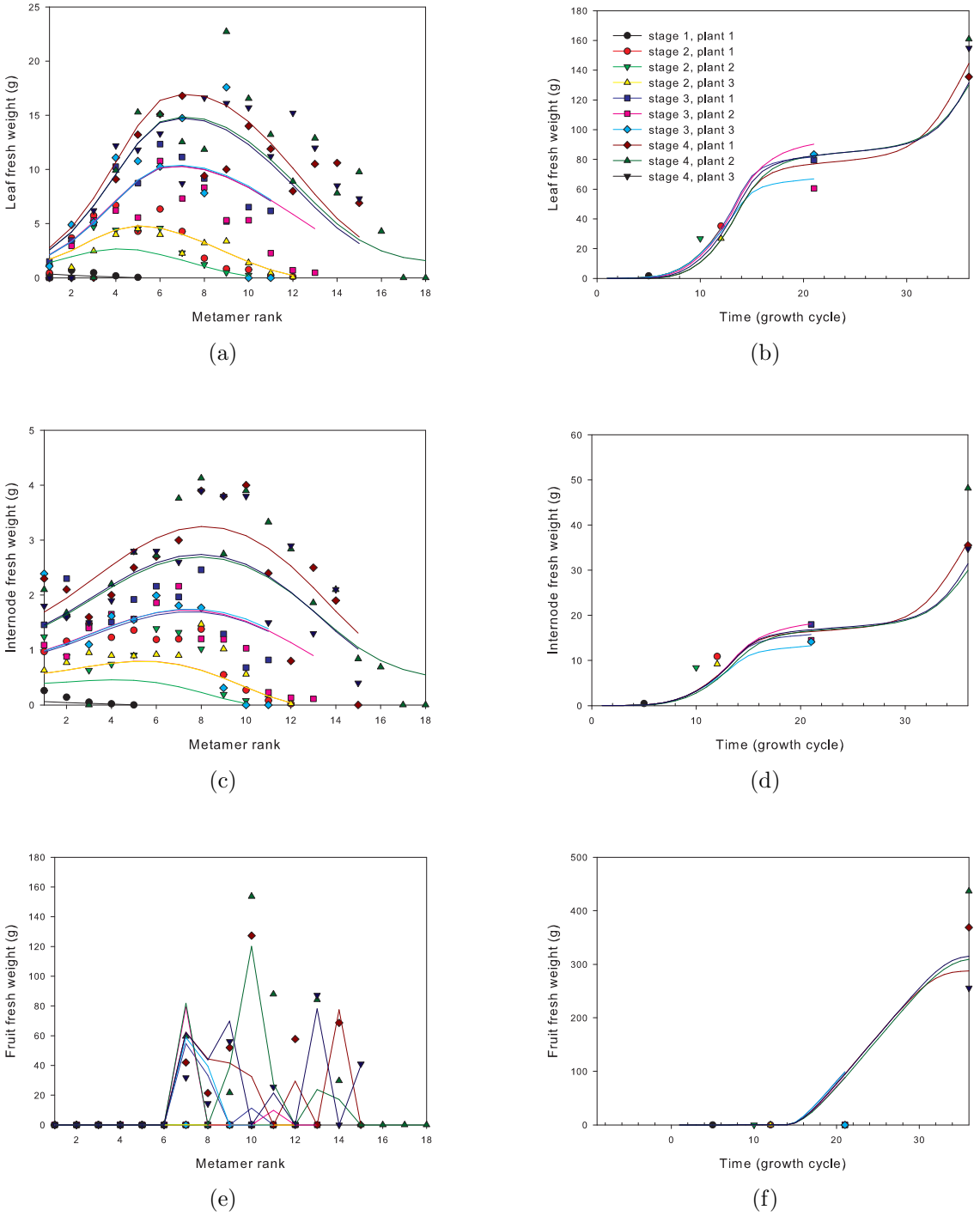
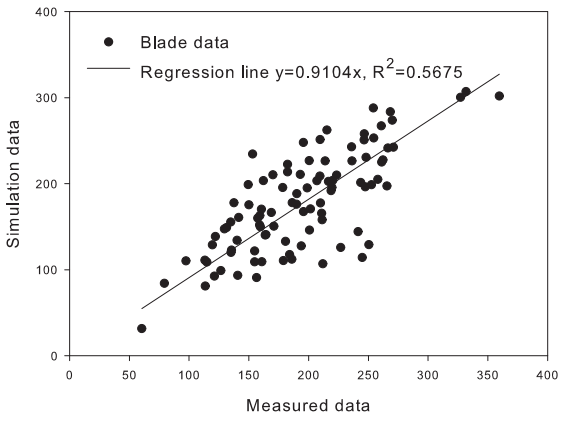
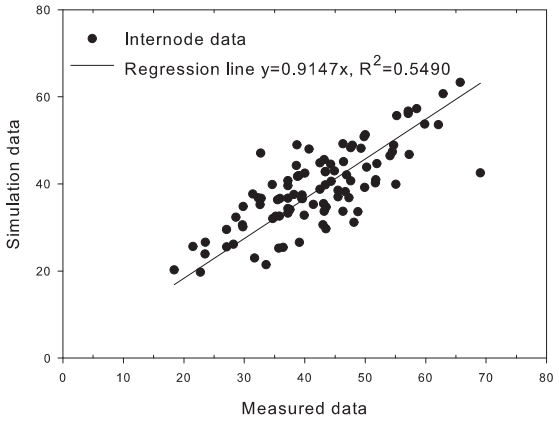


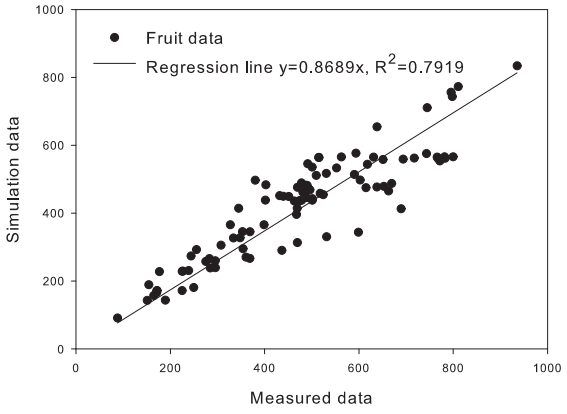
Figure 6.1: Comparison of the observation data and the simulation results with the estimated parameters of the fresh weight of (a) individual leaves (b) population of leaves (c) individual internodes (d) population of internodes (e) individual fruits (f) population of fruits, for IL1-1. Symbols represent observation data at each stage; lines represent the simulation results with the estimated parameter values.



(a)



(b)



(c)

Figure 6.2: Comparison of the observation data and the simulation results with the estimated parameters of the fresh weight of (a) leaves (b) internodes (d) fruits. "•" represents observation data at stage S4 corresponding to the ripening of the fruits; "—" represents the linear regression line.

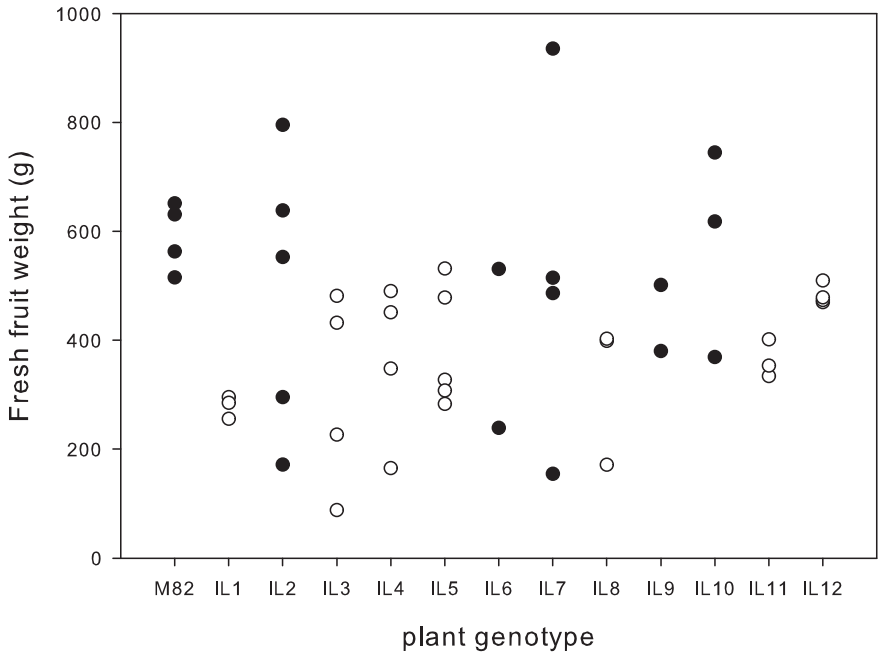
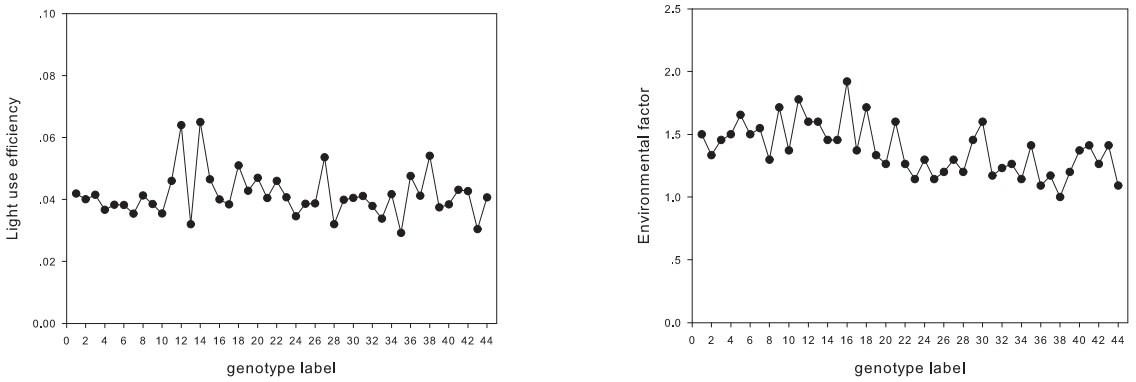


Figure 6.3: Fruit yield of all ILs and M82 at the harvest time. The empty circle represents the fruit yield that has the significant difference from the fruit yield of M82 at the 0.05 level.



(a) (b)

Figure 6.4: Results of (a) the light use efficiency (μ) and (b) the environmental factor (E_t) for all genotypes.

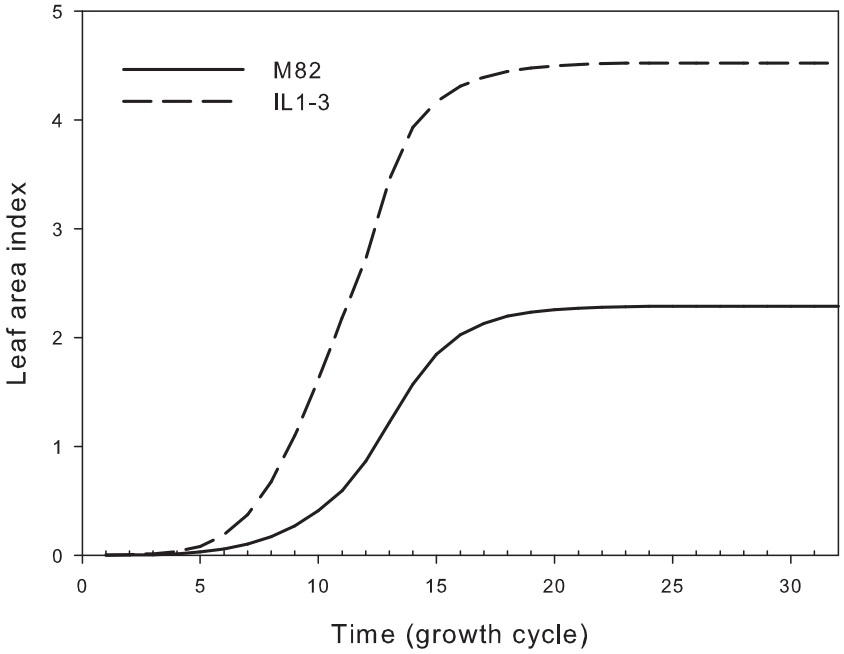


Figure 6.5: Leaf area index (LAI) of the IL1-3 and M82.

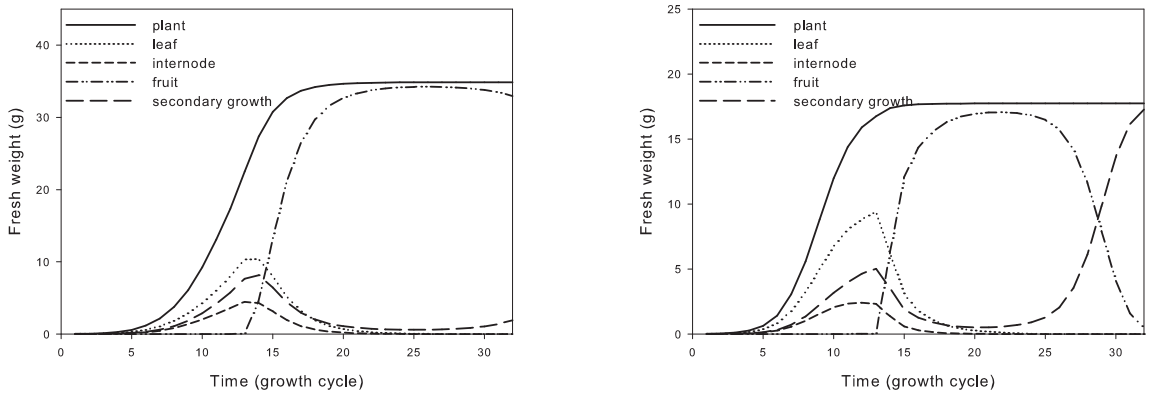
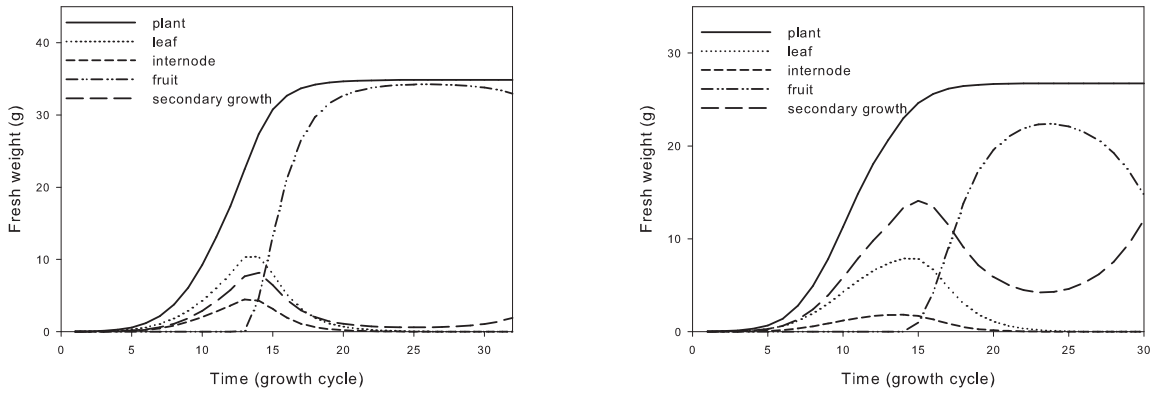
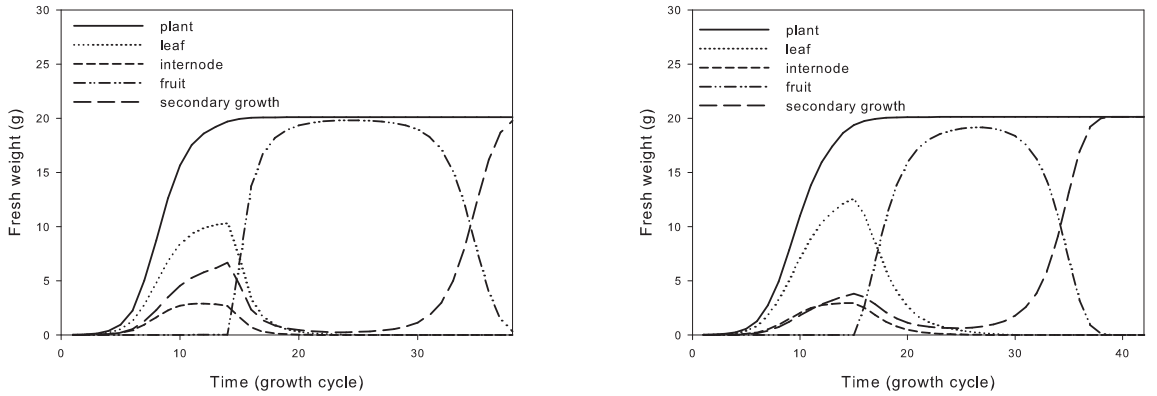


Figure 6.6: Biomass partition for (a) M82 and (b) IL1-3.



(a) (b)

Figure 6.7: Biomass partition for (a) M82 and (b) IL5-3 with the estimated parameter values.



(a) (b)

Figure 6.8: Biomass partition for (a) IL5-2 and (b) IL5-5 with the estimated parameter values.

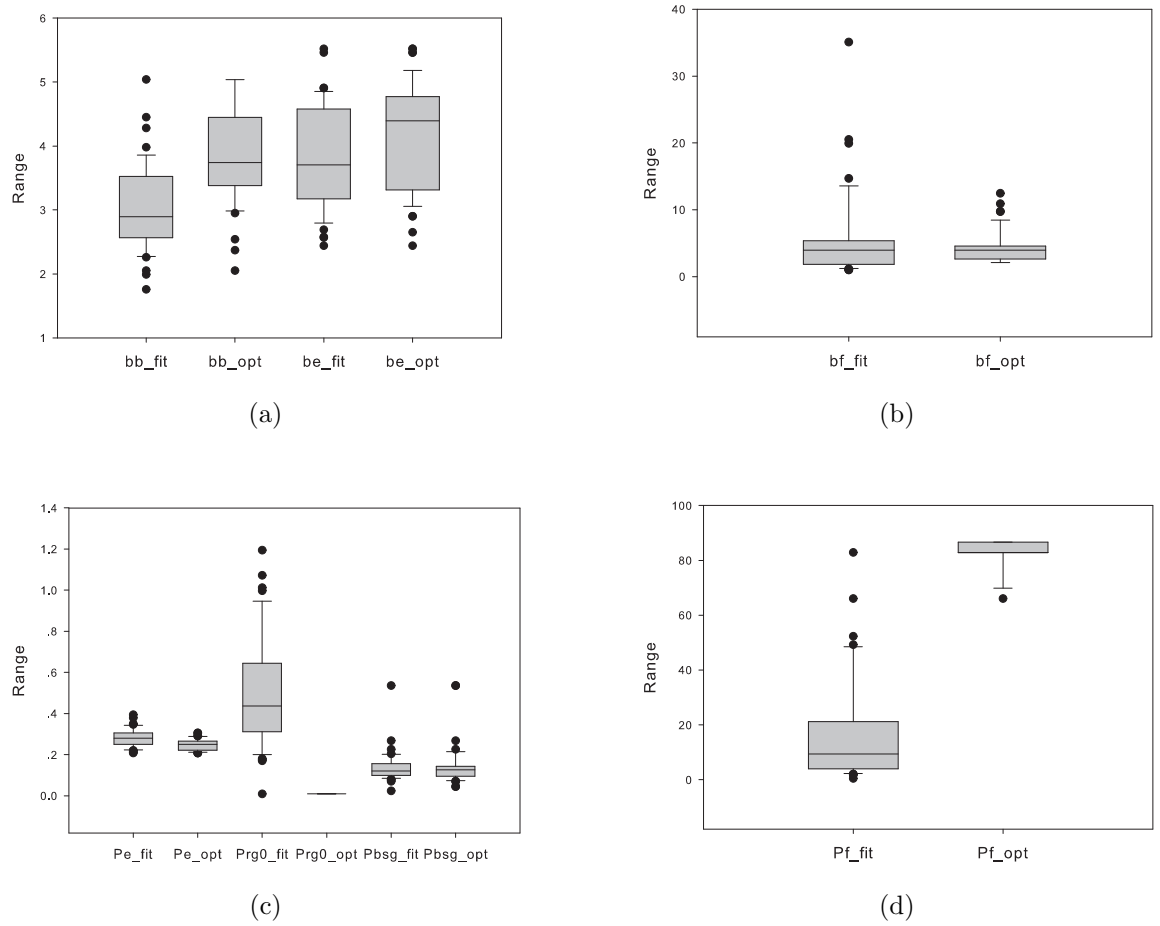


Figure 6.9: Range of the estimated and optimal values for each parameter. The suffix "_fit" represents the estimated parameter value; the suffix "_opt" represents the optimal parameter value. The prefix "bb" is the coefficient of the leaf sink variation function; "be" is the coefficient of the internode sink variation function; "bf" is the coefficient of the fruit sink variation function; "Pe" is internode sink strength; "Pf" is fruit sink strength; "Prg0" is the constant demand parameter of biomass for internode secondary growth; "Pbsg" is the constant demand parameter of biomass for leaf thickening.

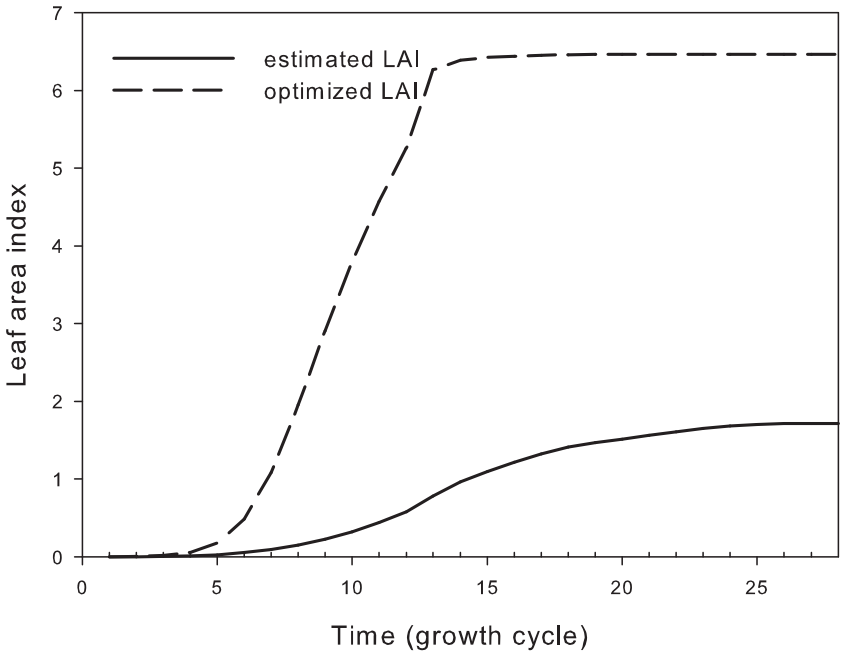


Figure 6.10: Comparison of the leaf area index for tomato IL2-5.

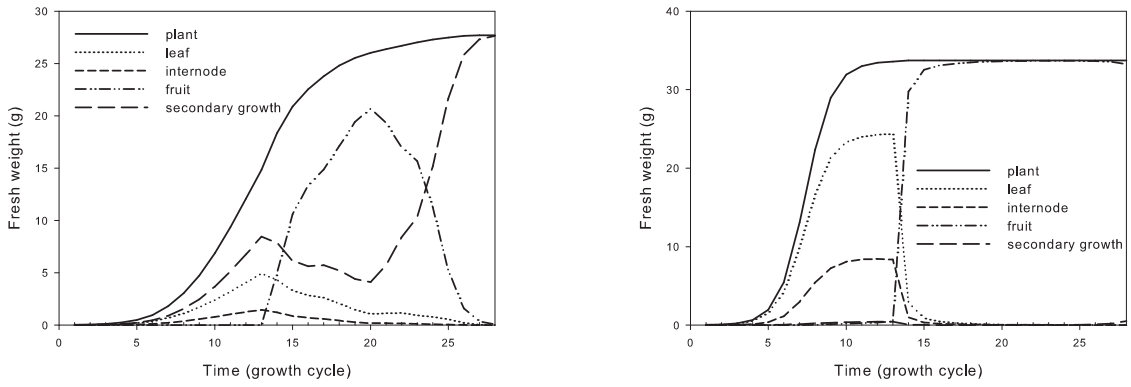


Figure 6.11: Biomass partition for tomato IL2-5 (a) with the estimated parameter values (b) with the optimal parameter values.

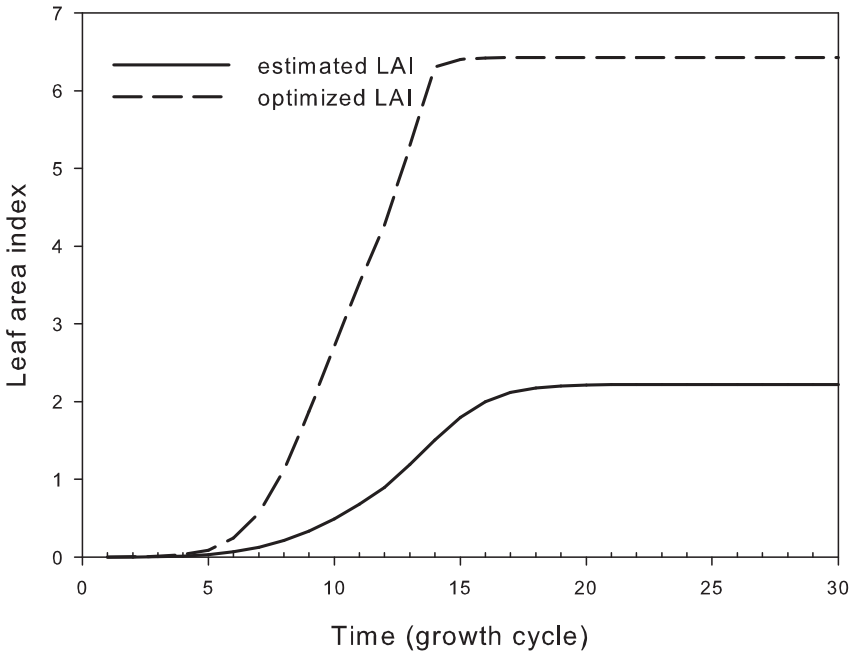


Figure 6.12: Comparison of the leaf area index for tomato IL8-1.

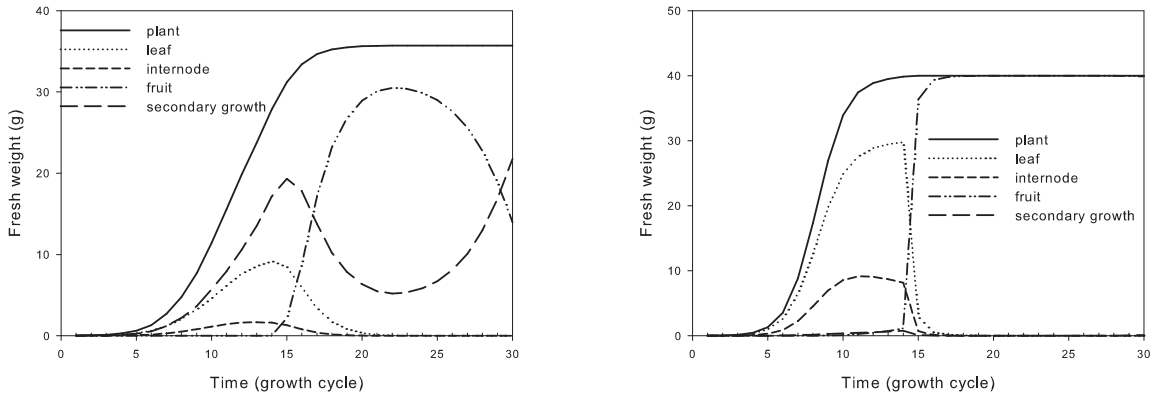


Figure 6.13: Biomass partition for tomato IL8-1 (a) with the estimated parameter values (b) with the optimal parameter values.

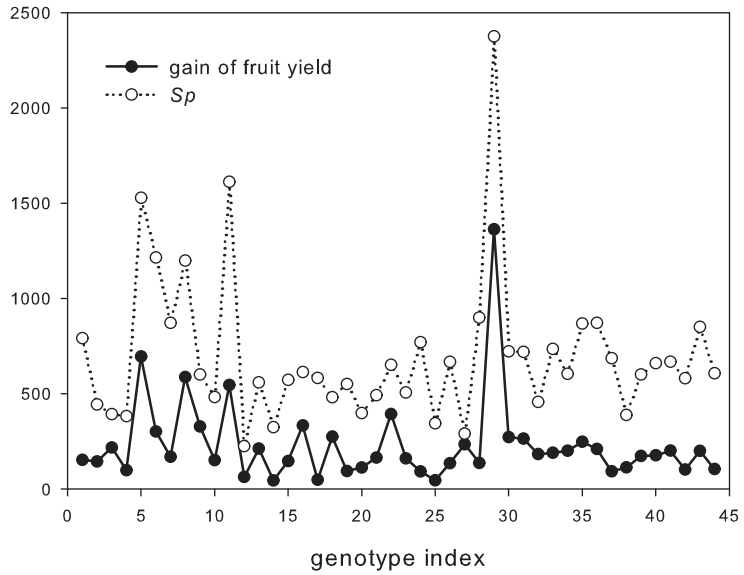


Figure 6.14: Relation between the gain of fruit yield and Sp. The index on the x-axis represents ILs, from M82, IL1-1,...,IL12-4.

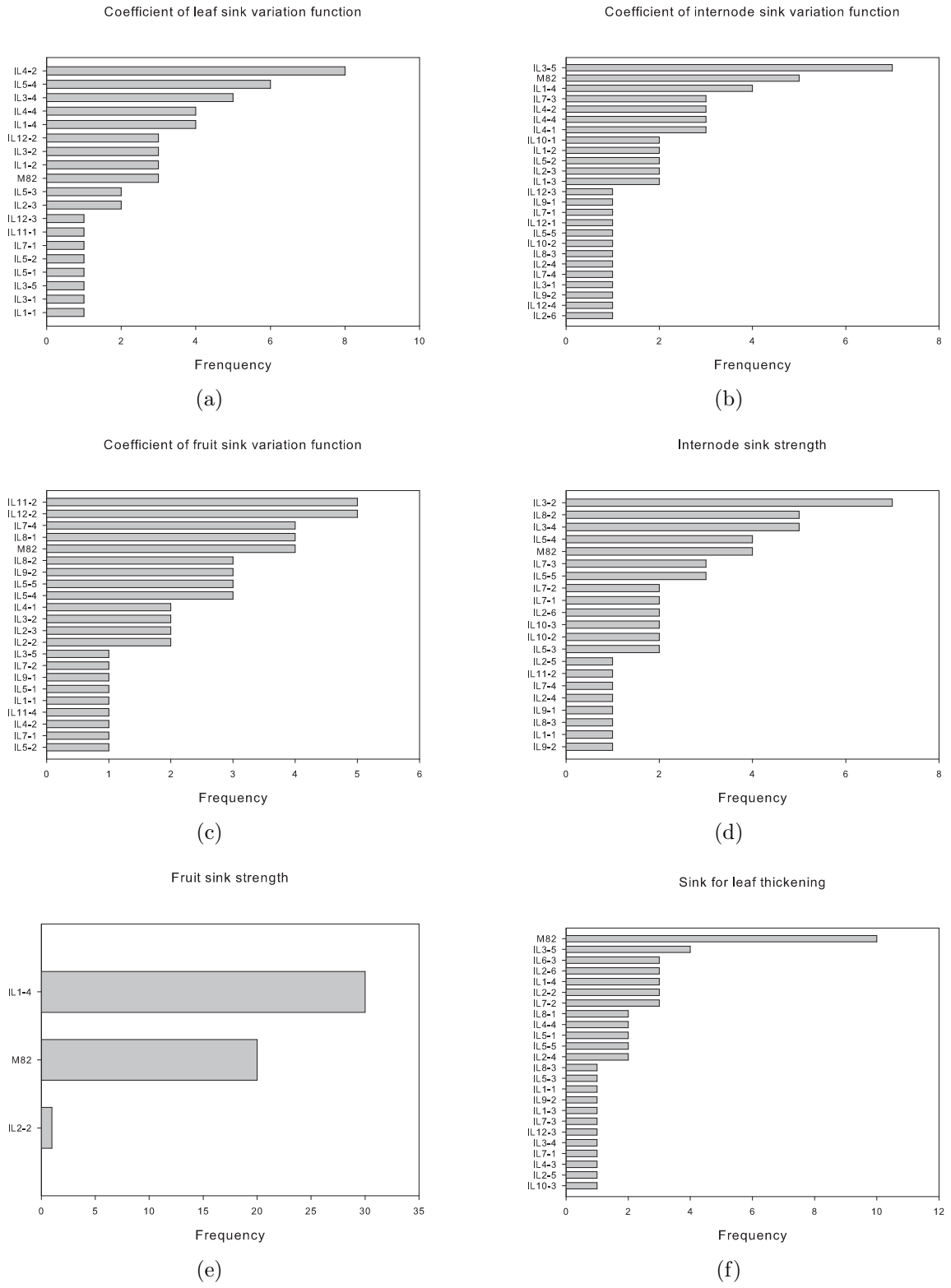


Figure 6.15: Distribution of ILs with respect to the optimal values of (a) coefficient of leaf sink variation function (b) coefficient of internode sink variation function (c) coefficient of fruit sink variation function (d) internode sink strength (e) fruit sink strength (f) constant demand parameter for the leaf thickening

Chapter 7

Optimization of fruit yield for maize

Maize (*Zea mays* L., DEA cultivar) is one of the most widely cultivated cereals all over the world. It is Commonly used as human diet in both fresh and processed forms. The cobs are sources of economical benefits. It is acknowledged that the final cob weight of maize depends on the relationship between cob sink and the availability of assimilates resulting from the plant biomass production, highly dependent on growth conditions during the early stages of grain filling, and it reflects the source-sink dynamics during the entire grain filling period (Borrás et al. [2002]). Therefore, the optimization problem of maximization of maize yield is investigated in this chapter. The optimized variables are specific parameters of the GreenLab model: the cob sink strength (P_1^f) and the two coefficients of its sink variation function (a^f, b^f). On the basis of the result of the relative stability of GreenLab parameters on maize among seasons and treatments (Ma et al. [2007], Ma et al. [2008]), these parameters are considered to be related to plant genetics here (another example based on the same hypothesis is given by Letort et al. [2008b]). In addition of these genetic-related parameters mentioned above, the impact of the cob position on the cob weight is studied. All the other parameters are maintained constant and their values are set as the estimated ones, which will be described below. Note that the model equations used in this thesis are slightly different from those presented in Guo et al. [2006] and Ma et al. [2007], which means that their results of the estimated parameter values cannot be straightforward applied.

In the usual cultivation conditions, maize is a single stem crop. The phytomer appearance is controlled by thermal time in the GreenLab model for maize. The model time step, growth cycle (GC), is thus equivalent to thermal time requirement for each phytomer appearance. The topology of maize cultivar ND108 as observed in the field is as follows: first six phytomers with short internodes appear; they are followed by 15 phytomers with longer internodes; the last one bears the male flower (tassel). Therefore, the organogenesis terminates at the end of the 21st growth cycle, but the plant is still alive until the 33rd growth cycle. Even though several phytomers may bear female flowers (cobs), Guo et al. [2006] found that the weight of the cob on the 15^{th} phytomer counted from the bottom of the stem is 95% of the total cob weight and thus chose to

consider only one cob and to gather all the potential cob weights on the 15th phytomer.

Seeds of maize cultivar ND108 were sown 0.6 m apart in north-south-oriented rows that were 0.6 m apart, at the China Agricultural University (CAU) ($39\,°50'N$, $116\,°25'E$) in 2000. The resulting planting density (28000 plants ha^{-1}) was about half of that commonly used by local farmers and was chosen to minimize competition among plants. Plants emerged on the 18^{th} May 2000. Soil, irrigation and fertilizer inputs were managed so as to avoid any mineral and water limitation, and plant disease, pest or stress symptoms. The experiments had four replications. Samples were taken destructively at 12 dates. One plant was collected per replication and sampling date. Only aboveground organs were considered as in Guo et al. [2006]. Fresh weights of blades, sheaths, internodes, cob and tassel; lengths and widths of sheaths; lengths, widths and areas of blades; and lengths and diameters of internodes, were measured and recorded individually at each sampling date. The specific leaf weight was 0.025 g/cm^2, for all leaves. The detailed information about the environmental conditions, sampling strategy, measured data and expansion duration and longevity of organs can be found in Guo et al. [2006].

7.1 Parameter estimation

A set of parameters, assumed to be species-specific, is estimated simultaneously by fitting several plants of a given species at different development stages. This procedure is called multi-fitting. For this estimation procedure, we used the dataset taken from Guo et al. [2006] as the measured data. In this chapter, the data of fresh weight of all organs measured at three stages (8^{th} growth cycle corresponding to the vegetative stage, 18^{th} growth cycle approximately corresponding to flowering stage and 33^{rd} growth cycle corresponding to physiological maturity) were used as target data. A generalized nonlinear least square method adapted from Levenberg-Marquardt's algorithm was used to estimate the parameters of GreenLab (Zhan et al. [2003]).

The estimated values of the hidden parameters of GreenLab are listed in Table 7.1. They are different from the ones in Guo et al. [2006] and Ma et al. [2007], as the sink variation function (Beta function) in this thesis is slightly improved as shown in Eq.2.7: the sink strength is defined as the maximum sink value (sink amplitude) in this thesis, while it is defined as the total sink capacity in the previous studies of Guo et al. [2006] and Ma et al. [2007]. The simulated values of organ fresh weight by GreenLab with the estimated parameter values are shown in Fig.7.1, compared with the measured data. The root mean squared error (RMSE) is 10.50 for all data of the three measurement stages and the coefficients of determination (R^2) for blade, petiole, internode, cob and tassel at the maturity (33^{rd} growth cycle) are 0.98, 0.95, 0.94, 1 and 1 respectively. The optimization work presented in the following sections is based on the estimated parameter values listed in Table 7.1.

Table 7.1: Estimated parameter values of GreenLab by multi-fitting of the data of maize cultivar ND108 measured at three different development stages simultaneously, using the generalized non-linear least square method.

Parameter	Definition	Value	Unit[1]
$E(t)$	Environmental factor at growth cycle t	1	–
μ	Light use efficiency	0.048	g/cm^2
Sp	Characteristic surface area related to plant crown projection	3600	cm^2
k	Light interception coefficient	0.68	–
P_1^b	Sink strength of blade	1 (fix)	–
a^b, b^b	Coefficients of beta function for blade	$3.59, 5.38$	–
P_1^s	Sink strength of sheath	0.6	–
a^s, b^s	Coefficients of beta function for sheath	$3.05, 3.69$	–
P_1^e	Sink strength of internode[2]	1.4	–
a^e, b^e	Coefficients of beta function for internode	$3.34, 1.65$	–
P_1^f	Sink strength of cob	806.47	–
a^f, b^f	Coefficients of beta function for cob	$8.34, 2.60$	–
P_1^{fm}	Sink strength of tassel	4.23	–
a^{fm}, b^{fm}	Coefficients of beta function for tassel	$1, 1$	–

[1] "–" means no unit is associated to the parameter.

[2] It is the sink strength of long internodes. The sink strength of short internodes is assumed to be proportional to the one of long internodes, and they share the same expansion mode. The ratio of the sink strength of long internodes to the one of short internodes is also a parameter that needs to be estimated and its value 4.67 is derived from the parameter estimation procedure with the other parameters.

7.2 Single optimization of maize cob yield

As introduced at the beginning of this chapter, the cob is economic valuable. Therefore, we investigate an optimization problem of maximization of the final weight of cob when plant age is t in this section. The formula of the corresponding objective function J is given by Eq.7.1. The variable to optimize is the cob sink value, p_1^f. Note that this variable is not only a multiplier in each item of J, but also implicitly involved through the plant demand D.

$$J = \sum_{k=1}^{t_x^f} p_1^f(k) \frac{Q(t-(t_x^f-k))}{D(t-(t_x^f-k))} \tag{7.1}$$

The parameters that we optimize are the cob sink strength and the coefficients of cob

sink variation function. The variation of the optimized parameters is limited to a reasonable range referred to (Ma et al. [2007], Ma et al. [2008]), as listed in Table 7.2. Besides the parameters listed in Table 7.2, the cob position is a factor that will be investigated. The total number of metamers of maize cultivar ND108 is 21. Hence, we assume that the cob can be initiated from the first metamer to the 21^{st} metamer, the metamers being numbered from the bottom of the stem. In reality, for maize cultivar ND108, the cob is rarely initiated on the metamers at the bottom of the main stem. However, for maize (*Zea mays* L.), it can occur that the cob is initiated on the lower metamers: in this case, the cobs are the remaining of lateral branches that existed for former maize lines (West-Eberhard [2003]). The results of optimization and simulation with respect to the cob position varying within a wide range also provide a deeper understanding of the source-sink dynamics involved in plant growth.

Table 7.2: Definitions and variation ranges of the GreenLab parameters that are optimized in the optimization problems.

Parameter[1]	Definition	Range
P_1^f	Sink strength of cob	$[0, 1500]$
a^f	Coefficient of beta function for cob	$[0, 25]$
b^f	Coefficient of beta function for cob	$[0, 25]$

[1] The parameters are unitless.

7.2.1 Expansion duration of the cob is independent of its position

In this section, the cob expansion duration is restricted to 19 growth cycles as observed in reality, regardless of the cob position. The optimization result obtained by maximizing the cob weight with respect to the cob position is illustrated in Fig.7.2(a). This figure shows that the inflexion point of the curve of the cob yield with respect to the cob position is the point with the coordinate of the cob position equal to 15. The cob weight increases monotonously when the cob position increases from the bottom to the 15^{th} internode. Then there is slight decrease when the cob position is between 16 and 21. The ratio of the cob weight when the cob is borne by the 21 internode to the one when the cob is borne by the 15 internode is 98.03%, the cob weight corresponding to the cob position 15 being the maximal cob weight among all cob weights shown in Fig.7.2(a) and the cob weight corresponding to the cob position 21 being the minimal cob weight among all cob weights when the cob position varies from the position 15 to the position 21. The shape of the curve shown in Fig.7.2(a) can be interpreted as follows: the cob weight is mainly determined by the time when the cob appears in the competition for biomass against other organs. The course of the cob development can be revealed by its sink value variation. Three randomly selected examples of the

corresponding optimal cob sink variations where the cob is at the 1^{st}, 5^{th} and 13^{th} internode on the stem are illustrated in Fig.7.2(b), respectively. Comparing the optimal sink variations, we found that the sink value is almost 0 at the beginning of the cob development to let leaf area index increase as illustrated in Fig.7.2(c) and thus more biomass was produced by photosynthesis. And then, the cob sink value increases very quickly at the last short period of its expansion duration, and nearly all the biomass is thus distributed to the cob as illustrated in Fig.7.2(d). When the cob stops getting biomass, the biomass of plant will only be distributed to leaves and internodes, or the tassel in the following growth cycles.
Comparing Fig.7.2(a) and Fig.7.2(b), we found that the cob at the 1^{st} internode developed earlier than the cob at the 5^{th} internode. However, the optimal cob weight for the plant with the cob at the 1^{st} internode is lower. It is because, on one hand, the cob at the 1^{st} internode competes for biomass against source organs (leaves) too earlier when the leaf area index is small and the plant biomass is not too much. On the other hand, the plant growth will be inhibited if the cob develops too early, due to the larger cob sink strength compared with the other organs, especially leaves which are the source organs used to produce biomass through photosynthesis. The optimal results in this section coincides with the conclusion in section 5.
The cob expansion duration is set to be 19 growth cycles as observed. To guarantee that the cob expands for 19 growth cycles before the plant terminates its growth (i.e. 33^{rd} growth cycle), the highest position of the cob is at the 15^{th} internode on the stem. If the cob position is higher than the 15^{th} internode, it will miss the potential ability to obtain biomass. The higher the cob is, the more potential ability missed. Therefore, the maximal cob weight with respect to cob position decreases where the cob position is higher than the 15^{th} internode.

7.2.2 Expansion duration of the cob depends on its position

To make sure that the cob has enough time to develop, and to avoid missing its potential ability of obtaining biomass, the cob expansion duration varies according to its position. The cob will have possibilities to obtain biomass until the plant stops growing, but the quantity of biomass it obtains at each time is controlled by its sink value. For instance, if the cob is at the 1^{st} position, its expansion duration is 33 growth cycles, where the plant terminates its growth at the 33^{th} growth cycle.
The maximal cob weights with respect to the cob position are not significantly different, as shown in Fig.7.3(a). The corresponding optimal cob sink variations are similar as shown in Fig.7.3(b) marked only with symbols, and the estimated one is represented by the curve in "–•–". Wherever the cob is and whatever its expansion duration is, we found that the optimal cob development strategies are similar. The optimal cob sink is almost zero at the beginning of cob development, and then the sink increases monotonously till the end of plant growth.

Comparing the estimated and the optimal cob sink variations with cob particularly born by the 1^{st} phytomer counted from the bottom as shown in Fig.7.3(b), we can separate the cob development process into four stages. During the first stage from the 1^{st} growth cycle to 15^{th} growth cycle, the optimal cob sink is a little larger than the estimated one (during the first stage, the estimated cob sink value is zero), but can be neglected. It has no negative effect on the source organ development. In this period, the leaf surface areas for the estimated plant and the optimal plant are the same. Consequently, there is no difference among the optimal cob weights when the cob is born by the 1^{st} phytomer to by the 15^{th} phytomer. However, it leads to a slight difference among the optimal cob weights for the plants with the cob born by the phytomer lower than the 15^{th} position and higher than the 15^{th} position, as shown in Fig.7.3(a). It is because even though the cob sink value is a little larger than the estimated one but does not affect the leaf surface area and thus affect the plant biomass, it does make the cob absorb biomass as the cob sink value is almost zero but not exactly equals to zero. The second stage is from the 16^{th} growth cycle to the 23^{rd} growth cycle. Even though the optimal cob sink is still a little larger than the estimated one, it does affect the source organ development during this stage. The cob competes for biomass against the source organs and the other organs. This competition leads to the decrease of the leaf surface area as shown in Fig.7.3(c). During the third stage from the 23^{rd} growth cycle to the 31^{st} growth cycle, the optimal cob sink value keeps increasing, but smoothly. On the contrary, for the observed plant, the cob sink increases significantly, and biomass allocation to the cob is done with detriment to leaves (i.e. less biomass is allocated to leaves). Hence, the leaf surface area begins to decrease. During the last stage of plant growth within two growth cycles, the optimal cob sink begins to increase significantly and quickly. Since the other organ sinks are negligible compared with the cob sink value, all the biomass is allocated to the cob as shown in Fig.7.3(d): the ratio of cob weight to the total weight of leaves and stem in this stage tends to infinity.

The comparison results of the cob sink variation reveal the source-sink dynamics. The increment of the cob weight is the product of the cob sink value and the ratio of the plant biomass production, which depends on the leaf surface area, to the plant demand that is the sum of all the organ sinks, as described in Eq.2.6. Even though the cob sink value is smaller than the estimated one, the leaf surface area is higher and the biomass production may thus be bigger. On the contrary, even though bigger cob sink value results in more biomass allocated to the cob instantaneously, it leads to less biomass allocated to other organs, especially leaves and less biomass production at the following cycles. Hence, to obtain maximal cob weight, the optimal trade-offs between sources and sinks should be considered. Compared with the estimated cob weight (1013 g), the optimal cob weight (1032 g) is 2% greater. The maize cultivar ND108 that we study, which results from long-term breeding programs, may already be close to optimum regarding cob yield. Hence, multi-objective optimization considering co-products is more interesting than the single optimization of maximization of cob weight for maize

cultivar ND108, since it may provide new information.

7.3 Multi-objective optimization of maize cob and vegetative compartment

Commonly, maize is used in the human diet in both fresh and processed forms; the grain and vegetative parts of maize are fed to livestock, and the components of the grain (e.g. starch) may be refined for direct consumption (Pratt [2001]). Moreover, the fact that the cob, and leaves and stem can be used as biofuel becomes of important economical interest (Baenziger et al. [2006]). Therefore, in this section, we investigate a multi-objective optimization problem of maximization of yields of both cob and the vegetative part consisting of leaves and stem for the plant with the cob borne by the 15^{th} internode that is the same as the observed one (Guo et al. [2006]). Cob weight and tassel weight are interrelated. Cob weight is controlled by pollen production, while pollen production depends on the tassel size of maize. Moreover, Uribelarrea et al. [2002] showed that if the tassel size is reduced, the cob size will be limited. Therefore, a constraint on tassel weight that should be beyond a threshold is imposed to the multi-objective optimization problem. The formula of the multi-objective optimization problem with the constraint is given by

$$\begin{aligned} & J = (J_1, J_2) \\ & \text{subject to} \quad g \geq \text{threshold} \end{aligned} \tag{7.2}$$

where J_1 is one of the objectives, which is the cob yield as expressed by Eq.7.1; J_2 is the other objective, which is the total weight of leaves and stem, as expressed by Eq.7.3; g is the tassel weight, as expressed by Eq.7.4.

$$\begin{aligned} J_2 \;=\; & \sum_{j=1}^{t} N_1^b(t-j+1) \sum_{k=1}^{\min(j,t_x^b)} p_1^b(k) \frac{Q(t-j+k)}{D(t-j+k)} \\ & + \sum_{j=1}^{t} N_1^e(t-j+1) \sum_{k=1}^{\min(j,t_x^e)} p_1^e(k) \frac{Q(t-j+k)}{D(t-j+k)} \end{aligned} \tag{7.3}$$

$$g = \sum_{k=1}^{t_x^{fm}} p_1^{fm}(k) \frac{Q(t-(t_x^{fm}-k))}{D(t-(t_x^{fm}-k))} \tag{7.4}$$

The optimal result of this multi-objective optimization problem, known as *Pareto front*, is shown in Fig.7.4. *Pareto front* is given by about 500 optimal solutions of cob sink variation. Hence, in Fig.7.5, we outlined the area covered by all the optimal solutions

of cob sink variation. Particularly, one example of the optimal cob sink variation is given, the corresponding cob weight being 503 g and total weight of leaves and stem being 2050 g.
The *Pareto front* of our multi-objective optimization problem is characteristic of source-sink dynamics and reveals the necessary balance between both objectives. Maximization of the total weight of leaves and stem leads to a zero cob sink strength. On the other hand, to maximize the cob weight, the cob sink value can not be maximal all the way, otherwise there would not be enough leaf surface area, and the reduced biomass production would decrease the final cob weight. For this reason, the left extremity of the *Pareto front* corresponds to a zero cob weight whereas the right extremity corresponds to a strictly positive weight of stem and leaves. For maize cultivar ND108 (*Zea mays* L., DEA cultivar), the tassel appears and begins to develop at the 21^{st} growth cycle, with a very quick expansion (2 growth cycles). From the optimal results shown in Fig.7.5, we found that the tassel expansion corresponds to the early stages when the cob sink begins to increase.
In Fig.7.6, the evolution of the tassel weight corresponding to the points on the *Pareto front* is illustrated. We see that for a wide range, the tassel weight does not vary since its expansion corresponds to growth cycles when the cob sink is still very low. However, we found that for the maximal cob weights (above 900 g), the tassel weight is decreasing. It corresponds to experimental observations of Westgate et al. [2003] who indicated that there is a potential gain of cob yield by decreasing the tassel weight.
The cob weight simulated by GreenLab with the estimated parameter values is 1013 g, the corresponding total weight of leaves and stem is 927 g and the tassel weight is 29 g. With the optimal parameter values, the maximal cob weight among the *Pareto front* in Fig.7.4 is 1032 g, the corresponding total weight of leaves and stem is 959 g and the tassel weight is 29 g. Compared the Harvest Index (HI), which is defined by the ratio of the cob weight to the weight of plant, of the estimated plant with HI of the optimal plant, HI of the optimal plant is surprisingly a little smaller than the estimated one, even though both the cob weight and the total weight of leaves and stem are higher than the estimated one. It revealed the trade-offs between sources and sinks. Post-expansion and fast growing rate as shown in Fig.7.3(b) and Fig.7.5 will enhance not only the cob weight but also the weight of leaves and stem. This optimal cob development strategy is in agreement with Weiner [1988] and Vega et al. [2000]: there is a threshold size for plants to produce flowers and fruits, the plant will grow as much as it can until its biomass reaches a threshold, and then the biomass may be distributed to fruits and flowers.

7.4 Conclusion

In our test case, the ideotype of maize can be deduced from the optimal results. It provides a reference to improve breeding strategies. From a physiological point of view,

the cob begins to absorb biomass at about the 20^{th} growth cycle when the leaf area saturates. And then, it should absorb biomass smoothly or significantly, depending on the breeding objective. If the objective is to have a maximal cob weight, the cob should have a bigger reproductive capacity, and the cob should grow with post-expansion (i.e. a long delay for expansion) and fast growth rate (i.e. expand within a short period). From an architectural point of view, the leaf size is reduced during the last vegetative and reproductive stages of growth. The harvest index is above 50%. It is coincident with the ideotype of maize proposed by Mock and Pearce [1975] by analysing the research results of other people with experiment based approaches. The *Pareto front* of the multi-objective optimization problem presents all the different optimal strategies, and the decision-maker could choose his optimal strategy according to market prices or the application purposes for example.

The maize that we optimized is assumed to have only one cob like the experimental data used for parameter estimation. However, the methodology that we used in this work is not restricted by the number of cobs on the stem. For the objective of the optimization problem, which is the maximization of cob weight, the crucial factor is not the number of cobs, but the optimal trade-offs between sources and sinks. However, in order to have more realistic optimal values, more constraints should be concerned. Since the cob growth requires pollen from the tassel and since there exists a strong interaction between cob and tassel (Borrás et al. [2002]; Uribelarrea et al. [2002]), we integrate the tassel weight into the multi-objective optimization problem as a constraint. A threshold is set for the tassel weight (not less than 10 g referred from experimental data). However, for all optimal solutions, this constraint is not active (tassel weights strictly above 10 g). One reason is that so far we do not know the relationship between cob and tassel quantitatively. Hence, it is difficult to set the threshold value. Another reason is that tassel sink variation is fixed and it does not change according to the cob sink variation, in this work. The understanding of the interaction between cob and tassel should be improved in our future works. Finally, in cobs, only kernels give the food for human beings or for livestock. The number of kernels is a critical factor that affects the final kernel weight (Borrás and Otegui [2001]). Therefore, taking into account the number of kernels per cob could be an interesting complement to this study. So far, we do not have the information about the proportion of the kernel weight to the cob weight, which raises the difficulty to estimate the corresponding model parameters. In this chapter, we have illustrated how the optimization of the parameters of plant growth models could be used as the first step to design ideotypes for genetic selection. The GreenLab model was chosen for the simplicity of its parametrization. Moreover, it describes plant growth, both from ecophysiological and architectural points of view, at the individual organ scale. Breeders can get information about physiological characteristics in determining yield from the optimal results.

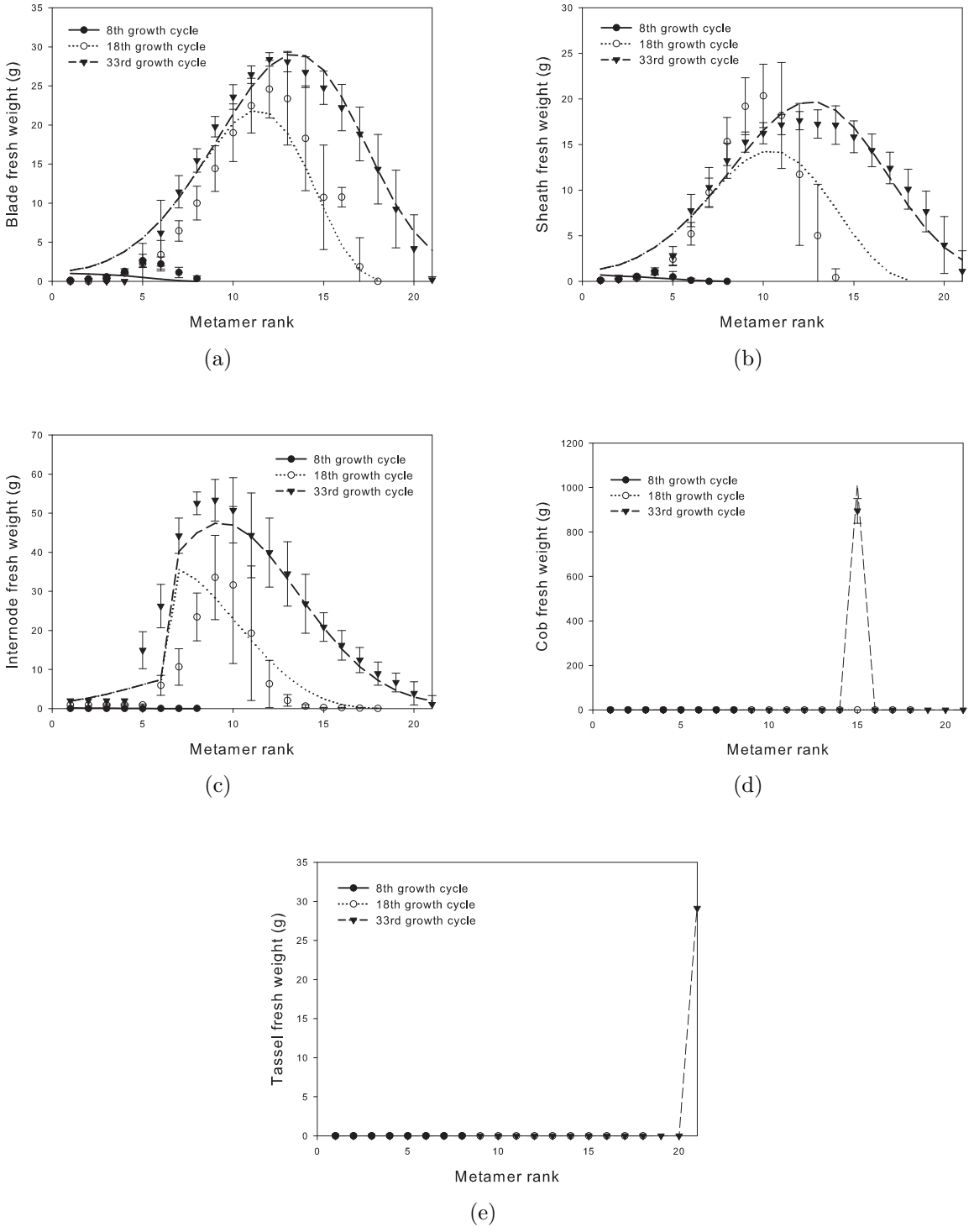


Figure 7.1: Simulation results of biomass partition to (a) blade (b) sheath (c) internode (d) cob (e) tassel, with estimated parameter values by generalized non-linear least square method. "•" represents measured data at the 8^{th} growth cycle corresponding to vegetative stage; "○" measured data at the 18^{th} growth cycle corresponding to flowering stage; the symbol of black triangle-down represents measured data at the 33^{rd} growth cycle corresponding to physiological maturity; "—" simulation result at the 8^{th} growth cycle; "· · ·" simulation result at the 18^{th} growth cycle; "– –" simulation result at the 33^{rd} growth cycle.

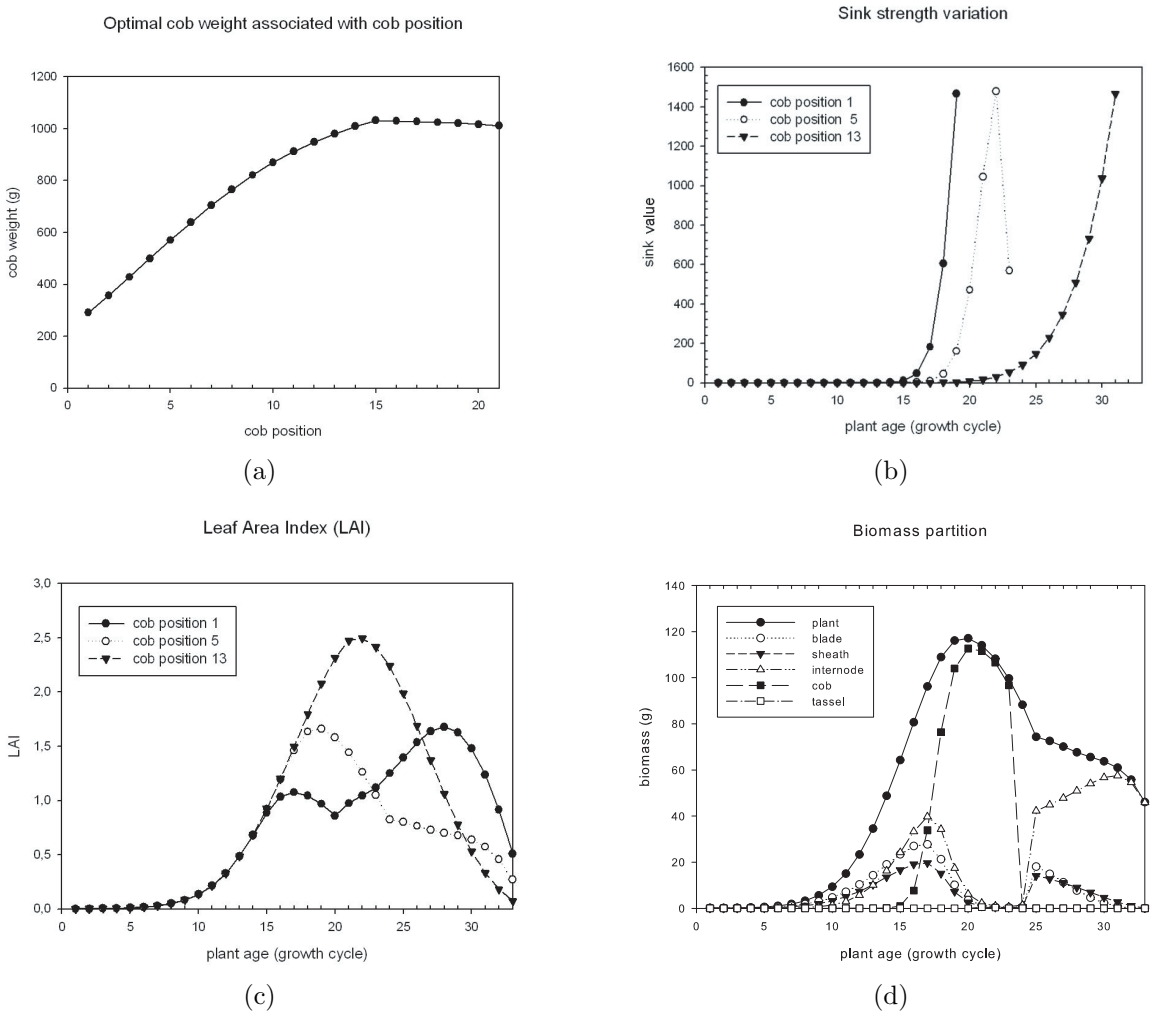


Figure 7.2: (a) Maximal cob weight associated with the cob position (b) Three randomly selected examples of the optimal cob sink variations where the cob is at the 1^{st}, 5^{th} and 13^{th} internode on the stem (c) Leaf Area Index (LAI) of maize with the cob at the 1^{st}, 5^{th} and 13^{th} internode on the stem, with the optimal cob sink variations as shown in (b) (d) Biomass partition to each kind of organs associated with the optimal sink variation as shown in (b), where the cob is at the 5^{th} internode on the stem.

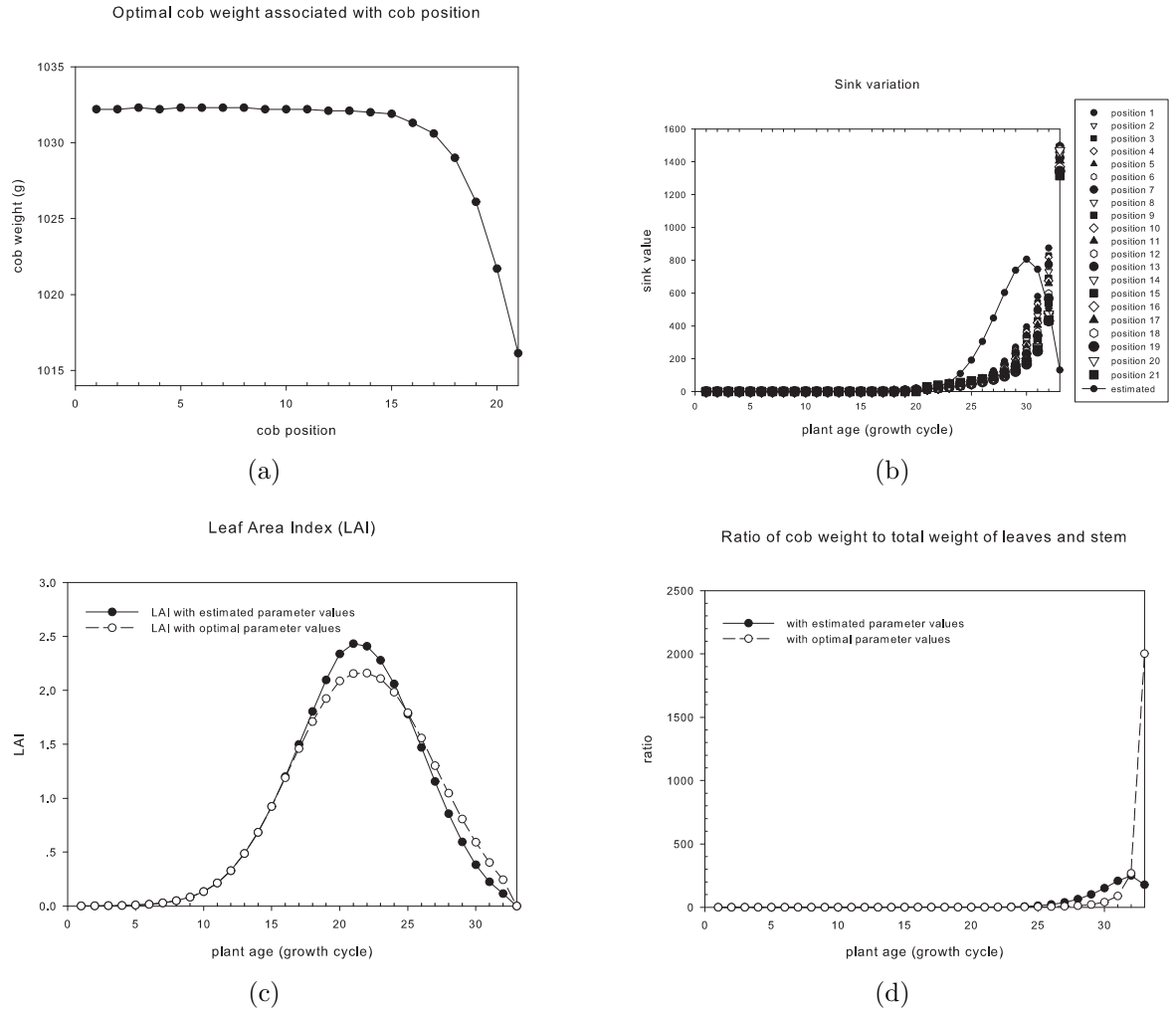


Figure 7.3: (a) Maximal cob weight with respect to the cob position (b) Cob sink variations: the curve marked only with symbol represents the corresponding optimal cob sink variation; the curve marked by "–●–" represents the estimated cob sink variation. (c) Simulation result of leaf area index for maize with estimated parameter values marked by "–●–" and with the optimal one marked by "– – ∘ – –" where the cob is borne by the 15^{th} phytomer counted from the bottom. (d) Comparison of the ratio of cob weight to vegetative compartment weight during the plant growth. "–●–" represents the result with estimated parameter values and "– – ∘ – –"represents the result with the optimal parameter values where the cob is on the 15^{th} phytomer.

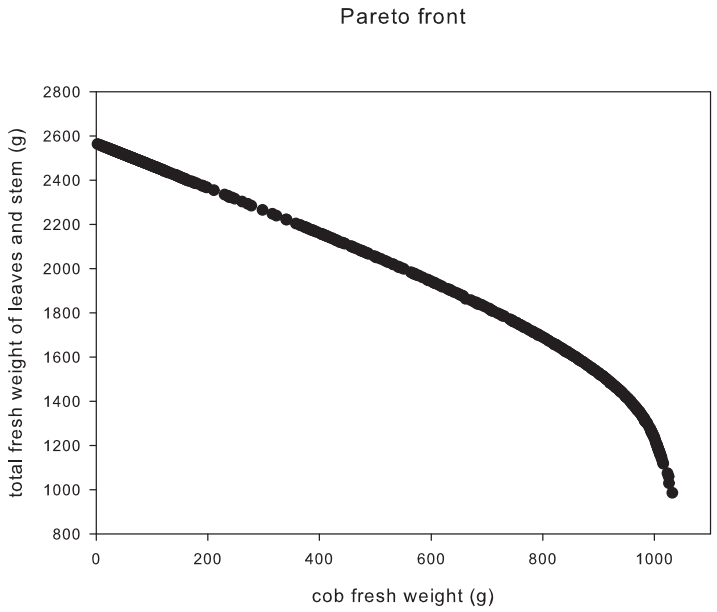


Figure 7.4: *Pareto front* of the multi-objective optimization problem.

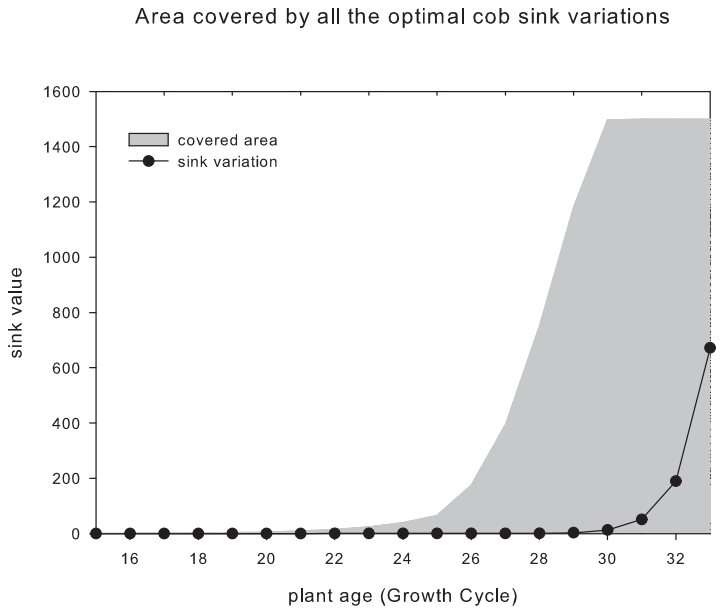


Figure 7.5: Area covered by all the optimal cob sink variations associated with the *Pareto front*. One example of the optimal cob sink variations is given. "–●–" represents the optimal cob sink variation for maize where the cob weight is 503 g and the total weight of leaves and stem is 2050 g.

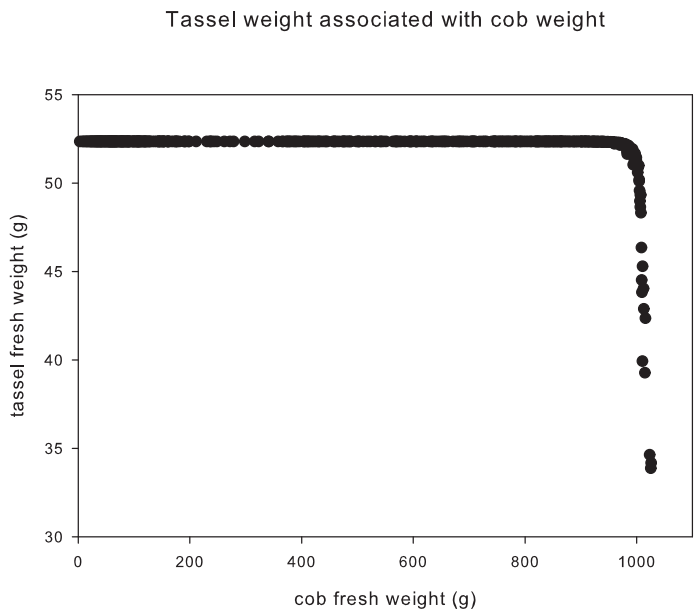


Figure 7.6: Tassel weight with respect to cob weight, associated with the *Pareto front*.

Chapter 8

Optimization of wood yield

Nowadays, the economical importance of wood production is enhanced by an increasing demand from South-East Asia (Zhang et al. [2006]) by increasing custom duties on wood export from Russia (Russia to go ... 2007) and by extending fields of applications: paper, pulp, joinery, furniture, biofuel, etc (Nepveu [1993]). In terms of environmental protection and resource utilization, wood as a renewable source attracts more and more attention (Wang et al. [2004]). For most application fields, not only wood production but also wood quality is important. These two criteria are of primordial importance to determine the economical value of logs. Therefore, the objective of the optimization problem in this chapter is to maximize wood production and to have better wood quality. Two GreenLab parameters are considered in order to maximize wood production: the sink strength for cambial growth and a coefficient that determines the way the biomass assigned to cambial growth is allocated to each metamer, through optimization and simulation respectively. The optimization procedure is based on a heuristic optimization algorithm called Particle Swarm Optimization (PSO). First, wood production is maximized without considering the effect of wood distribution on tree mechanical stability. In a further step, the mechanical stability of trees submitted to their self weight is taken into account based on simplified mechanical assumptions described in section 2.5.

Tree response to mechanical forces, e.g. self weight and wind loads, depends on tree structure and wood mechanical properties, which can be both modified by insects and diseases or silvicultural practices. In addition to external loads, trees can develop internal biomechanical stresses due to wood cells maturation (Fournier et al. [2006]). This active response aims at controlling the shape of tree axes and their position in space. In the particular case of trunks, this active mechanism usually corresponds to a negative gravitropism that allows trees to remain vertical. Even though these biomechanical phenomena have been widely studied in the past, the underlying biological mechanisms, e.g. mechanoperception and signal transduction, are still poorly known (Moulia et al. [2006]). Modelling and simulating tree biomechanics is a helpful approach that can be used to support or test biological hypotheses (Fourcaud et al. [2003]). Nevertheless

such biomechanical models are not classical as the mechanical solicitations apply on a growing structure (Ancelin et al. [2004b], Fourcaud and Lac [2003], Fourcaud et al. [2003]). However, the goal of this study is not to consider these complex biomechanical processes, but to introduce a simple mechanical criterion of stem stability in the functional-structural plant growth model GreenLab in order to assess the influence of model parameters on wood production and wood quality.

Wood quality can be defined with regard to several criteria including for instance ring and stem profiles, knot distribution and sizes, wood mechanical and chemical properties, heterogeneity of these properties within the stem and rings, stem straightness and/or cross section shapes. So far, most of the existing wood quality models include one or more of these properties (Nepveu [1993], Denne et al. [1994], Houllier et al. [1995], Deleuze and Houllier [1997], Woodcock and Shier [2003], Lasserre et al. [2005]). Jayawickrama [2001] pointed out that among these wood properties, wood stiffness and stability are of major importance in fast growth radiate pines. Moreover, Lasserre et al. [2005] found that there is a negative relationship between stand density and wood stiffness (an increase in stand density being associated to a decrease in stem diameter at breast height (DBH)). Therefore, in this thesis, wood quality is evaluated according to tree mechanical stability, which involved the coupling influences between tree structure, load distribution and wood mechanical properties (Ancelin et al. [2004b]). For this purpose, a mechanical stability criterion is used, which is based on stem buckling equations. We assume that this criterion summarizes both wood mechanical properties and the mechanical state of the tree, excluding the effects due to biomechanical responses as defined above, i.e. formation of reaction wood and induction of growth stresses. The final displacement angle of the top stem is considered as an indicator of tree stability.

This study aims at proposing a methodology for the investigation of wood quantity and wood quality. It is not applied on a specific tree species. However, the GreenLab parameters used in this study (see Table 8.1) are chosen based on the previous published research works of GreenLab carried out on trees (Guo et al. [2008], Letort et al. [2008a]). The constitutive material of a tree stem is supposed to be isotropic. It was observed on several tree species that the structural Young's modulus is only slightly different as well along the stem (Niklas [1997a], Niklas [1997b]) as within the stand (Brüchert et al. [2000], Alméras et al. [2004]). Therefore the structural Young's modulus is set to be constant in our study. Moreover, the first method "Number of leaves" is chosen to calculated the global allocation of biomass to cambial growth, and to begin from the simple one, the parameter P_1^{rg} is set to be 0 (see section 2.3.2 for detail information of secondary growth in GreenLab). Moreover, the sink value of each metamer of each physiological age p_p^{rg}, which determines the amount of biomass of cambial growth allocated to each metamer, is set to be identical, which is 1. The organogenesis mechanism of the tree is predefined as shown in Fig.2.1.

Table 8.1: Parameter values of the GreenLab model for tree. The definitions of the parameters are referred to chapter 2.

Parameter	Value	Unit
$E(t)$	1	_[1]
μ	0.1	g/cm^2
Sp	$1E+6$ (isolated)	cm^2
	$1E+5$ (high density)	
k	1	–
slw	0.04	g/cm^2
Q_{seed}	1	g
ta	1	growth cycle
P_m	5	–
P_p^b	1,0.8, 0.6 0.4, 0.2 from PA 1 to PA P_m	–
P_p^e	0.4,0.32, 0.2 0.1, 0.08 from PA 1 to PA P_m	–
E_y	10	GPa

[1] "_" means no unit is associated to the parameter.

8.1 Optimization of wood quantity

The effect of the sink strength for cambial growth (P_0^{rg}) and of the coefficient (λ) on stem wood production is investigated in this section. The two parameters (P_0^{rg}, λ) are optimized by the Particle Swarm Optimization algorithm in order to maximize stem wood production. The stem wood quality is not considered in this section. As described in section 2.3.2, P_0^{rg} and λ have distinct effects: P_0^{rg} determines the total biomass allocated to wood formation at the whole-plant level, while λ drives the way this total amount of biomass is partitioned in the tree architecture. The effects of these two parameters on tree growth being distinct, they were considered independently in this section.

8.1.1 Optimization formula

The formula of the optimization problem of maximization of wood production (trunk weight) J is given by Eq.8.1. The optimization variable is the sink strength for cambial growth (P_0^{rg}).

$$
\begin{aligned}
J &= Q_{trunk} = Q_{pith} + Q_r \\
&= \sum_{j=0}^{t} N_1^e(j) \cdot q_1^e(j,1) + \sum_{j=0}^{t} \sum_{h=1}^{j} q_1^{rg}(j,h) \\
&= \sum_{j=0}^{t} N_1^e(j) \cdot P_1^e \frac{Q(j)}{D(j)} \qquad \text{where} \\
&\quad + \sum_{j=0}^{t} \left(P_0^{rg} \sum_{p=1}^{P_m} N_p^b(j-1) \frac{Q(j)}{D(j)} \right) \sum_{h=1}^{j} \left[\left(\frac{1-\lambda}{D_{sg1}(j)} + \frac{\lambda \cdot N^b(j-h)}{D_{sg2}(j)} \right) l_1(j-h) \right]
\end{aligned} \tag{8.1}
$$

Q_{trunk} represents the trunk weight, Q_{pith} represents the total pith weight of the trunk and Q_r represents the total weight of the rings surrounding the pith on the trunk; the physiological age of the metamers on the trunk is 1. For trees in temperate zones, organs finish elongation within one year that corresponds to one growth cycle in GreenLab. Hence, $q_1^e(n,j) = q_1^e(n-j+1,1)$.

Note that P_0^{rg} is not only a multiplier of the second item of Eq.8.1 explicitly, but also implicitly in the equation for plant total demand D. As plant biomass increment Q is a function of D, P_0^{rg} appears also implicitly in the equation of Q as expressed by Eq.2.23. Hence, it is impossible to get the differentiation information of the objective function of the maximization problem as shown in Eq.8.1.

8.1.2 Influence of the sink strength for cambial growth (P_0^{rg}) on wood production

Simulated trunk weight and tree height with regard to P_0^{rg} are shown in Fig.8.1. Tree height decreases monotonously with increasing in P_0^{rg}. This result is straightforward if we consider the GreenLab equations describing biomass allocation in the tree. Tree height is indeed deduced from the summation of trunk metamer lengths. Metamer length depends on pith weight and is therefore a function of the amount of biomass allocated to primary growth. Due to mechanisms of sink competition, this quantity is reduced when the demand of the biomass compartment for cambial growth increases with increasing in P_0^{rg}. According to Eq.2.3, Eq.2.5, Eq.2.9, and Eq.2.12, increasing in P_0^{rg} leads to increasing plant demand D, which leads to the decrease of the amount of biomass allocated to leaves and piths. Hence increase in P_0^{rg} results in lower pith weights which lead to shorter metamers.

Trunk weight is determined by summing the pith weights of all its constitutive metamers and the total weight of rings surrounding these piths. Increasing P_0^{rg} has two contradictory effects: a negative effect on pith weights and a positive effect on ring weights. With increasing in P_0^{rg}, a larger amount of the tree total biomass is allocated to cambial

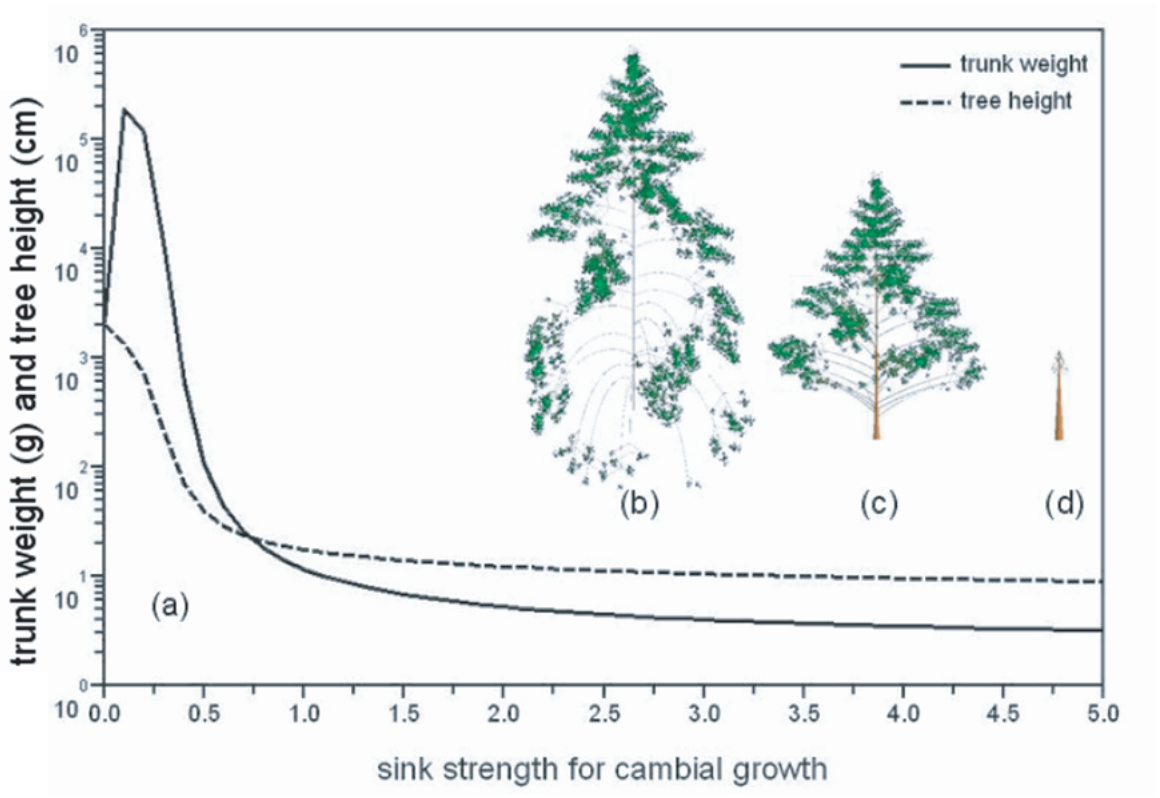


Figure 8.1: (a) Variations of trunk weight and tree height with regard to P_0^{rg} (y-axis is in logarithm scale.). (b), (c) and (d) 3D image of tree: in (b) $P_0^{rg} = 0$, trunk weight is 2.03 kg, height is 19.79 m; in (c) $P_0^{rg} = 0.1$, trunk weight is 187.00 kg, height is 13.54 m; in (d) $P_0^{rg} = 0.5$, trunk weight is 0.11 kg, height is 0.39 m.

growth at the expense of primary growth and in particular of leaf growth. Hence, leaf area index (LAI) decreases, which reduces in turn the plant photosynthetic potential. This means that an optimal P_0^{rg} value can be found that maximizes the trunk weight (Fig.8.1(a)). Under the condition that the other parameters of the GreenLab model are fixed, the PSO algorithm with 20 particles in the swarm provides after 50 iterations the optimal value of P_0^{rg}, which is 0.1, that maximizes the trunk weight. Fig.8.1(b-d) represent the 3D shapes of three simulated trees with different values of the sink strength for cambial growth.

Fig.8.2 shows that when sink strength for cambial growth takes a value exceeding 50, 95% of the biomass produced by the tree is allocated to cambial growth and piths. As a result, not enough biomass remains for leaves, thus inducing a slowdown of tree growth. This result coincides with the previous results of (Mäkelä [1986]). Therefore, far from generating higher yield, increasing tree cambial sink strength (P_0^{rg}) has a negative effect on the tree growth rate. It leads to shorter tree height, branches and sizes of leaves

which can reduce the tree performances in situations of high competition with neighbor trees, due to less light amount intercepted by the leaves.

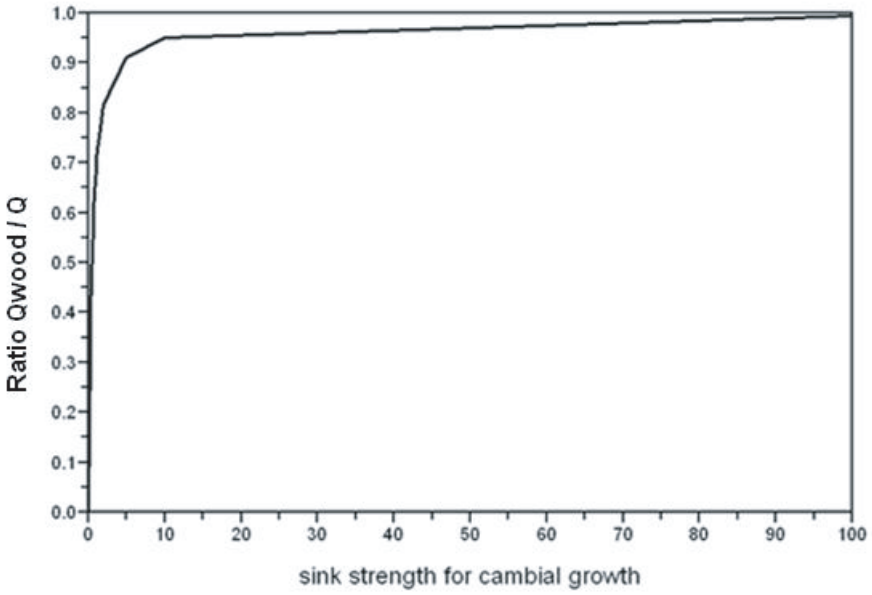
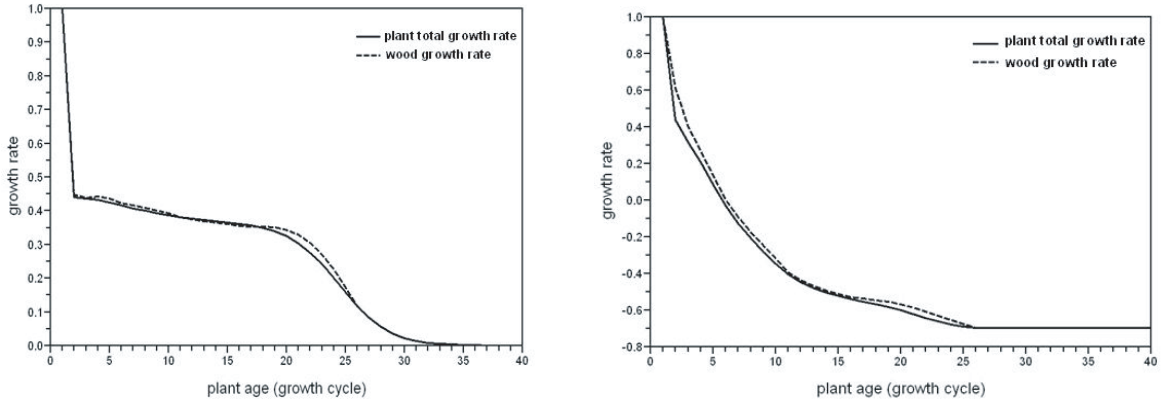


Figure 8.2: Ratio of biomass for cambial growth and for piths Q_{wood} to tree biomass Q with respect to P_0^{rg}.

Fig.8.3 enables to compare relative growth rates of wood (pith and rings) to relative tree growth rate. Relative growth rate is defined as the ratio $(V(t)-V(t-1))/V(t-1)$, $V(t)$ being the biomass allocated to wood or the tree biomass production at growth cycle n. As shown in Eq.2.3, Eq.2.23, Eq.2.9, and Eq.2.12, the biomass production of plant Q and the total biomass for cambial growth are both related to the product between numbers of organs and the ratio of biomass production Q to total demand of plant D. In addition, the sink strength for cambial growth P_0^{rg} is set to be constant. Therefore, the relative growth rates of wood and of photosynthetic potential are similar. At the first growth cycle, as plant biomass production and biomass allocated to wood increase from 0, their growth rates are both 100%. After that, the growth rates decrease due to the competition of cambial growth against the growth of leaves for photosynthesis. For the particular tree with the maximal trunk weight and with the optimal balance between cambial growth and photosynthetic potential, the relative growth rates for wood growth and tree growth are zero at the end of plant growth, as shown in Fig.8.3(a). However, for larger P_0^{rg}, relative growth rates tend to be negative as shown in 8.3(b), and it implies that trees are under suppression.



(a) (b)

Figure 8.3: Comparison of relative growth rates between wood growth (pith and rings) and tree growth. (a) P_0^{rg} is set to the optimal value 0.1 (b) $P_0^{rg} = 1$.

8.1.3 Effect of λ on wood production

The partition coefficient λ in Eq.2.17 determines the distribution of total biomass assigned to cambial growth to each metamer. It does not change the tree height which is controlled by primary growth. Eq.2.17 implies that λ should be equal to 1 in order to allocate the maximal proportion of wood to the stem. In that case, the proportion of biomass allocated to the stem is higher than that allocated to branches. Fig.8.4 shows the variation of the trunk weight as λ increases through simulation and gives two examples to illustrate the impact of λ, all the other parameters being fixed and P_0^{rg} being set to be the optimal value. As the simulation results showed that the trunk weight increases monotonously with increasing in λ, the use of the optimization algorithm to optimize trunk weight with respect to λ is not required.

8.2 Optimization of wood quality with biomechanical constraint

The effect of P_0^{rg} and λ on stem wood quality is studied in this section. The mechanical criterion of stem stability is represented by the final rotation angle at the stem tip. It depends on tree weight, tree height and diameters, and stem conicity. The variation of this final rotation angle versus stem diameter at breast height (DBH) with increasing in λ is shown in Fig.8.5(a), where P_0^{rg} is set to be the optimal value 0.1. DBH decreases with decreasing in λ, due to the dominant effect of mode D_{sg1}. Meanwhile, the stem tip tends to bend more.

Fig. 8.5(b) shows the variations of stem diameter versus the final rotation angle at

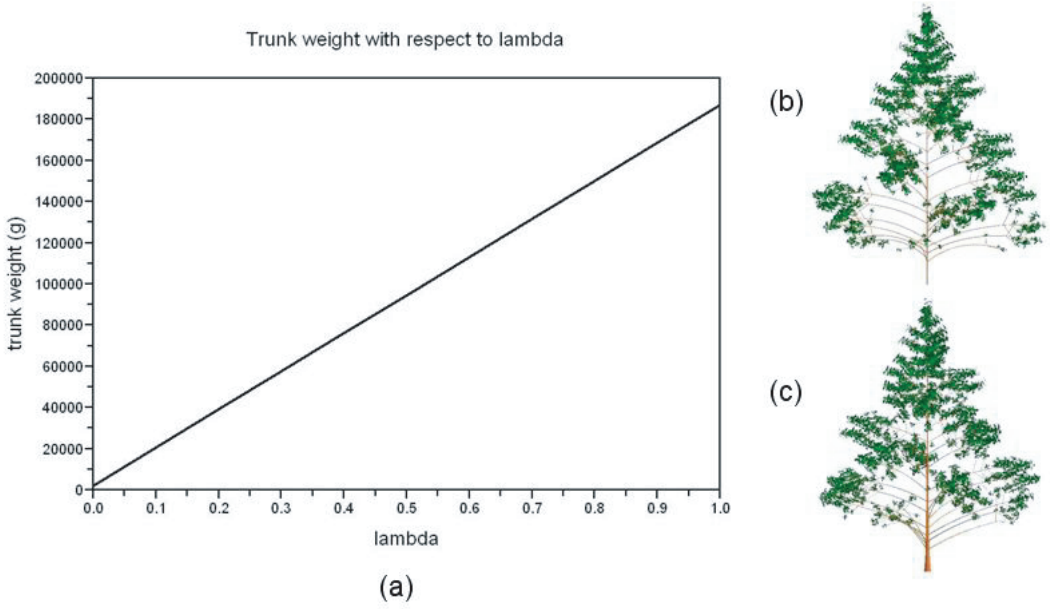
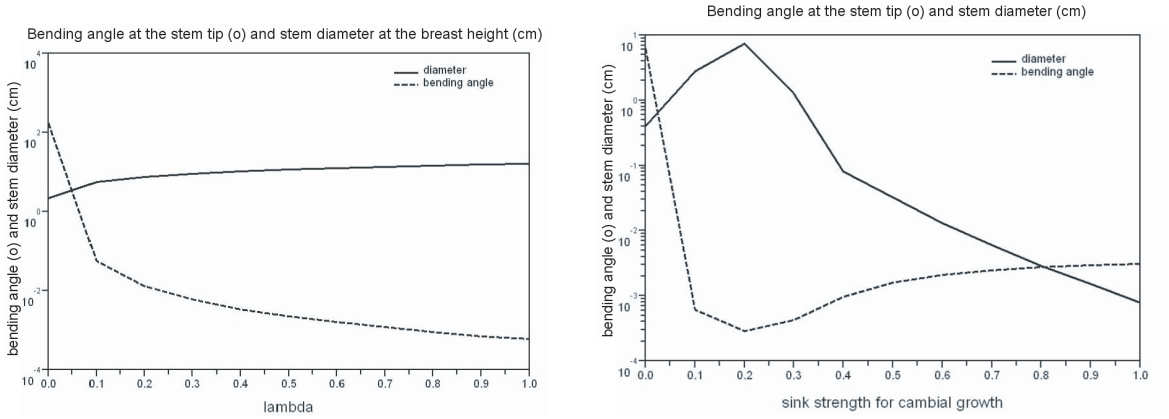


Figure 8.4: (a) Variation of the trunk weight with respect to λ. (b) and (c) 3D image of tree with the optimal value of sink strength for cambial growth: $P_0^{rg} = 0.1$: in (b) $\lambda = 0$, trunk weight is 1.63 kg, tree height is 13.54 m; in (c) $\lambda = 1$, trunk weight is 187.00 kg, tree height is 13.54 m.

the stem tip with increasing in P_0^{rg}, where λ is set to be 1. As shown in Fig.8.1(a), when P_0^{rg} is larger than 0.3, tree height is less than 1.30 m. In this particular case, the diameter at the stem tip is considered instead of DBH. Contrary to λ, P_0^{rg} influences the photosynthesis process and tree height as illustrated by Eq.2.23 and Fig.8.1. Consequently the load (self-weight) applied to the stem as well as the tree height vary with increasing in P_0^{rg}. Nevertheless, Fig.8.5(a) and Fig.8.5(b) reveal that the stem DBH and the final rotation angle at the stem tip always have opposite slopes, i.e. increasing bending angle always coincides with decreasing DBH and reciprocally. Two examples are given to illustrate the impact of the coefficient λ on stem stability as shown in Fig.8.6.

8.2.1 Impact of stand density on tree stability

So far, we investigated the growth behavior of an isolated tree with respect to the GreenLab parameters. In this section, the effect of the GreenLab parameters on the growth behavior of a tree in high stand density will be investigated. Fig.8.7 shows differences in old tree diameters considering different densities. In the GreenLab model, density is related to total ground projection area available of the crown for plant *Sp*.



(a) (b)

Figure 8.5: Relationship between stem diameter and bending angle at the stem tip with respect to (a) lambda (λ) (b) sink strength for cambial growth (P_0^{rg}). (y-axis is in logarithm scale.)

From Fig.8.7, we see that with respect to λ, the differences between diameters at the stem base and at the stem tip for an old tree in high stand density (approximately 1000 plants/ha in Fig.8.9) are all less than the ones for an old isolated tree. It implies that as stand density increases, the stem profile tends to be more cylindrical. Furthermore, for any density value, DBH monotonously increases with respect to λ, while the rotation angle at the stem tip decreases, as shown in Fig.8.5(a) and Fig.8.8. This result is in agreement with the beam theory in mechanics where the beam stiffness is proportional to beam diameter to the fourth power. Two 3D images of an old isolated tree and an old tree growing in high density are shown in Fig.8.9. These two trees share the same parameter values except one: the total ground projection area available of the crown for plant which is related to stand density.

8.3 Conclusion

Contrary to the common idea that increasing sink strength for cambial growth would increase the final wood production, we showed that there is an optimal value of the sink strength. This sink strength, P_0^{rg}, determines the quantity of total biomass allocated to the rings of trunk and of branches surrounding the piths. It controls the global allocation of the tree biomass into cambial growth. The level of competition of cambial growth against the other organs is reflected by the ratio of biomass production to biomass demand as shown in Eq.2.23, Eq.2.9, Eq.2.12 and Eq.2.3 (see also Mathieu et al. [2009]). In terms of source-sink dynamics, an increase in P_0^{rg} has two contradictory effects. On one hand, it induces a slowdown of the plant photosynthetic potential. It

(a) (b)

Figure 8.6: 3D image of bending stems with respect to λ, where $P_0^{rg} = 0.1$, tree weight is 1120 kg. (a) $\lambda = 0$, trunk weight is 1.63 kg, bending angle at the stem tip is 176.2 °; (b) $\lambda = 1$, trunk weight is 187 kg, bending angle at the stem tip is 0 °.

results in less biomass allocated to leaves, and consequently in a decrease in leaf area index. Meanwhile, it reduces biomass allocated to piths (pith weight). On the other hand, it increases the total weight of rings surrounding piths. Therefore, to maximize wood production (pith and rings), a balance had to be found between cambial growth and primary growth.

In the literature, for a branching point on the trunk, one part of the assimilates produced by a leaf goes downward towards the root system and the other part goes upward to the crown. However, the Pressler law described by Eq.2.16 only reproduces the downward propagation of assimilates in the plant hydraulic system. In our model, we introduced the coefficient λ to drive the proportion of upward to downward propagation. The value of λ can be estimated from real data for a given tree species, by fitting the stem profile (Guo et al. [2008], Letort et al. [2008a]). According to our simulation results, in order to maximize the trunk weight, the optimal value of λ is 1, which leads to minimizing the load of branches. Hence, this preferential allocation of biomass to the trunk penalizes the cambial growth of branches and thus reduces their hydraulic conductivity. The case where λ is equal to 0 is favourable to water conduction in

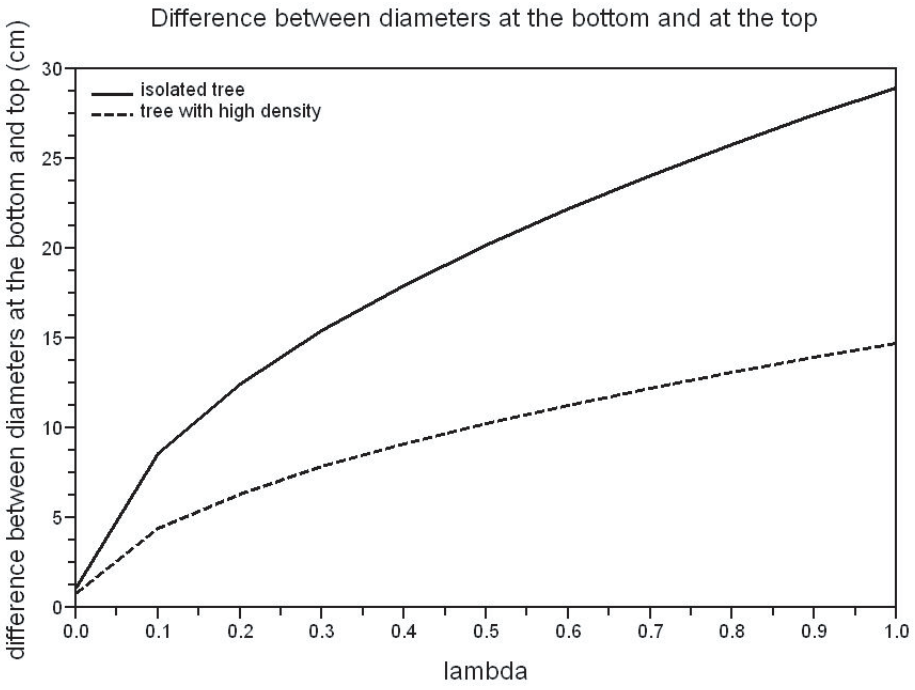


Figure 8.7: Difference between diameters at the bottom and at the top of the stem with respect to λ. The maximal total ground projection area available for an old tree in high density is 10 m^2 (approximately 1000 plants/ha).

branches, whereas it has negative effect on stem stability; and vice versa. Thus, the value of this parameter λ has also indirect effect on water conductivity, which should be considered as additional constraints when applying the methods presented in this chapter to real case studies. This can be done through multi-objective optimization, i.e. maximization of stem stability and water conductivity of branches simultaneously with respect to λ. Considering the contradictory effects of λ on these two objectives can be an interesting extension of this work. However, how to evaluate the water conductivity of branches is not clear so far.

The analysis of the simulated results based on GreenLab in this chapter reveals emergent properties of GreenLab that are in accordance with the literature. No matter considering variations of λ or P_0^{rg}, as DBH increases, the stem bends less and less. The same relationship between DBH and stem stiffness is obtained for trees both in isolated and dense states, as illustrated in Fig.8.5(a) and Fig.8.8, respectively. With the simplified mechanical assumptions, because the bending angle of a stem resulting from a load is negatively proportional to the second moment of area which is biquadratic of diameter, the stiffness of a stem is mainly influenced by stem dimensions. Therefore, a bigger diameter coincides with a smaller bending angle. This phenomenon is observed for most kinds of trees (Wiemann and Williamson [1989a], Wiemann and Williamson

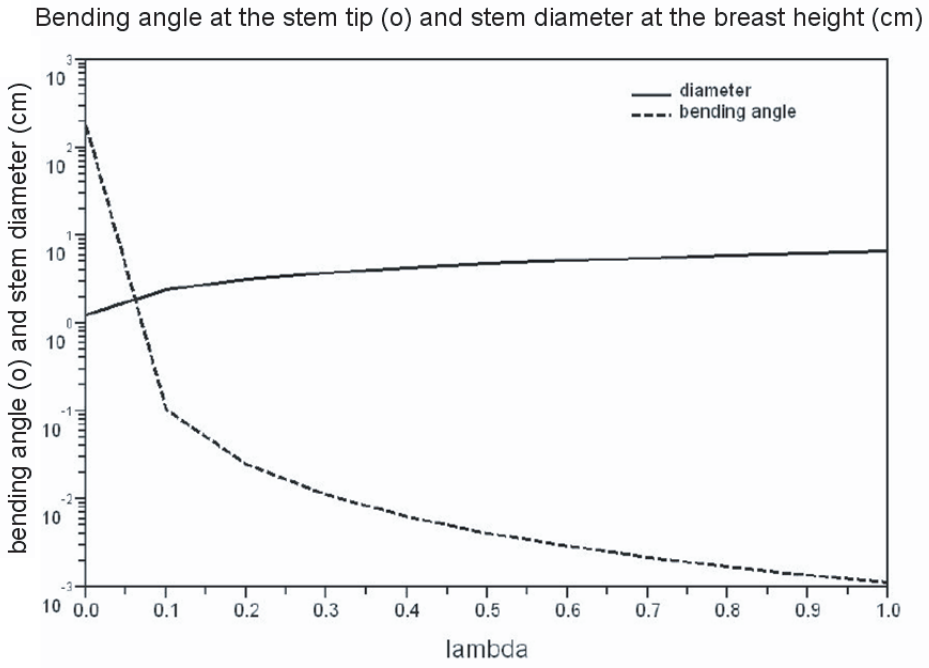


Figure 8.8: Simulation results of bending angle at the stem tip and stem diameter at breast height for old trees growing in high stand density (y-axis is in logarithm scale). The maximal total ground projection area available for an old tree in high density is 10 m^2 (approximately 1000 plants/ha).

[1989b], Woodcock and Shier [2003], Ancelin et al. [2004a]). The results coincide with the hypothesis in (Wiemann and Williamson [1989a] and Wiemann and Williamson [1989b]): radial increases are related to mechanical support; stiffness increases as diameter increases. The increase in stiffness conferred with an increase in diameter was observed, no matter whether a tree is isolated or it is in high density. Compared to isolated trees, the stem profile of trees grown in high density tends to be cylinder (Fig.8.7). It is in agreement with common observations in forestry (Assmann [1970]).

The optimization results that we presented in this chapter and chapter 7 open interesting perspectives for applications in breeding programs. In the present work in this chapter and chapter 7, it was assumed that the model parameters were genotypic, i.e. they only depend on the plant genes and are independent of environmental influences. A preliminary study to validate this assumption on real tree species is presented in Letort et al. [2008a] where two beech trees (*Fagus sylvatica*) growing in different local environments were fitted with the same set of genotypic parameter values. In addition, as described in chapter 7, the results found by Ma et al. [2007] and Ma et al. [2008] based on maize revealed the relative stability of the GreenLab parameters in different environmental conditions and different planting densities. Recently, efficient methods

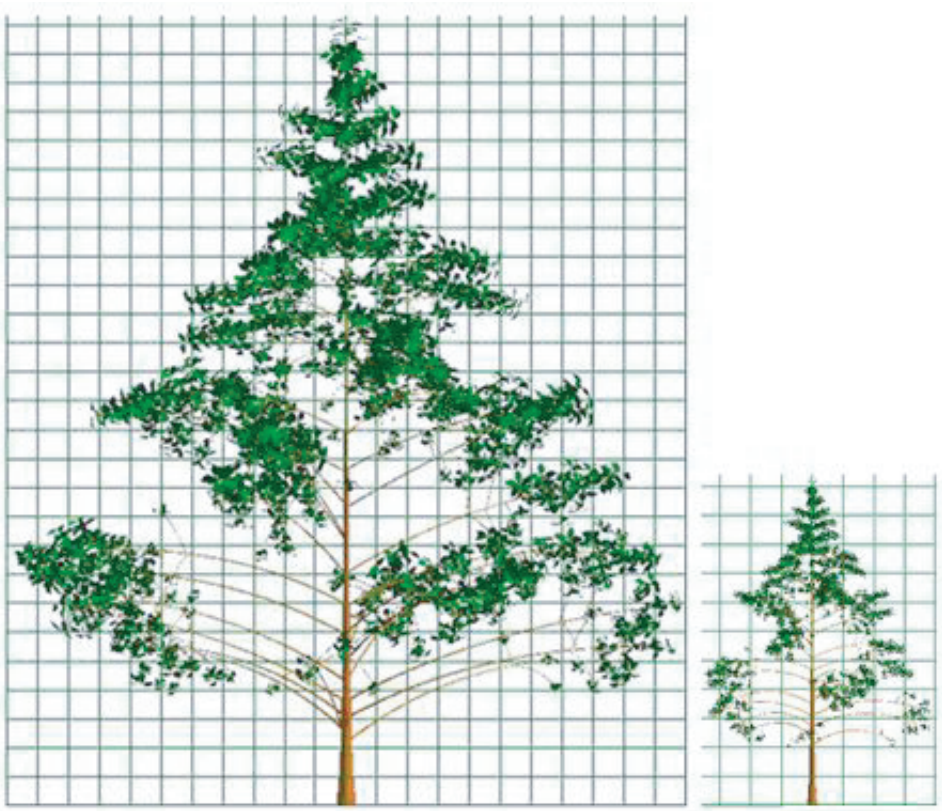


Figure 8.9: Comparison of 3D images of an isolated tree on the left and a tree growing in high density (approximately 1000 plants/ha) on the right, where the other parameter values are the same, especially P_0^{rg} is 0.1 and λ is 1. The heights of an isolated tree and a tree growing in high density are 13.54 m and 5.60 m, respectively; DBHs are 16.18 cm and 6.56 cm, respectively.

have been developed in the field of quantitative genetics to identify particular loci on the plant chromosomes that contribute to phenotypic traits (de Vienne [1998]). These methods rely on establishing statistical correlations between quantitative traits that can be measured on plants (e.g. yield, duration, height) and the values of a set of particular genes (markers). Several authors have emphasized the potential benefits of coupling these methods to models of plant growth in order to unravel the complex genotype $\times$ environment interactions and thus to improve genetic selection (Hammer et al. [2002], Tardieu [2003], Yin et al. [2004], Hammer et al. [2006]). Models such as GreenLab can play the role of intermediate link between genotype and phenotype. It requires identifying the genotype loci that have influence on the parameters of GreenLab instead of the classical breeding traits. A simulation study of these methods applied to the case of GreenLab can be found in Letort et al. [2008b]: a simple genetic model was built to determine the model parameters as functions of the genotype of the plant considered.

In this context, defining a set of optimized parameter values under given objectives is equivalent to defining ideotypes. Ideotype is the set of desirable traits that a plant should present to enhance yield or any other objective trait under specified conditions (Donald [1968]). To benefit from the recent advances in plant growth modelling, this set of traits should include the values of model parameters (Dingkuhn et al. [2007], Letort et al. [2008c]). It implies that modellers should be able to provide a set of optimal parameter values in response to specified objectives, which is precisely what was done in this chapter and chapter 7.

Chapter 9

Optimal control on leaf harvest

Pruning is a common and necessary action when people cultivate crops or trees, in order to make plants grow well and to reach their own aims. For example, for tomato, branches are pruned to reduce their competition for biomass against fruits; in the horticulture point of view, trees have to be pruned to be a certain shape. Once pruning is done during the growth of plants, plant growth behavior does have been influenced, as the architecture of plants is changed due to pruning. Considering the cases mentioned above, tomato yield and wood quality are variable.
To make plant growth well or to make plant produce more biomass, the amount of functioning leaves are crucial. However, the practices described above, i.e. pruning branches, will result in the reduction of the number of leaves. Particularly, in the domain of agriculture, one of the most special cases is to harvest leaves, as leaves are economical valuable, tea plant for instance. However, harvesting leaves will affect the growth behavior of a plant and thus affect leaf yield. Hence, in this section, we optimize the pruning strategy in order to obtain maximal yield of leaf harvest.

9.1 Problem statement

The objective of the optimization problem in this section is to maximize yield of leaf harvest within certain range of ages. The control variable is pruning strategy.
We abstract it and generalize it as an optimal control system. In this section, we take organs of physiological age 1 as example to introduce the optimal control system and the optimization procedure.
State variable $\mathbf{X}(\mathbf{t})$ is

$$\begin{cases} \mathbf{X}(t) = \begin{bmatrix} x_1(t) & x_2(t) & \dots & x_{ta}(t) \end{bmatrix}^T, & 1 \le t \le T_N \\ \mathbf{X}(0) = 0, \end{cases}$$

where $x_i(t)$ is the total mass of leaves of age i at growth cycle t; T_N is the final age of the plant; ta is the functioning time of leaves. When $t = 0$, there is no leaves, hence $x_i(0) \equiv 0, \forall\ i \in [1, ta]$.

Control variable $\mathbf{U(t)}$ is

$$\begin{cases} \mathbf{U}(t) = \begin{bmatrix} u_1(t) & u_2(t) & \dots & u_{ta}(t) \end{bmatrix}^T, & 1 \le t \le T_N \\ \mathbf{U}(0) = 0, \end{cases}$$

where $u_i(t)$ is the proportion of number of harvested leaves of age i at growth cycle t among total number of leaves of age i at growth cycle t, $0 \le u_i(t) \le 1$. Particularly, if $u_i(t)$ is equal to 1, all the leaves of age i at growth cycle t are harvested. Conversely, if $u_i(t)$ is equal to 0, we don't harvest any leaf of age i at growth cycle t.
According to the GreenLab principles, the state variable $x_i(t)$ is calculated by Eq.9.1.

$$x_i(t+1) = Nb(i,t+1)\frac{x_{i-1}(t)}{Nb(i-1,t)} + Nb(i,t+1)p_1^b(i)\frac{Q(t)}{D(t)}, \quad \forall\, t \in [0, T_N] \tag{9.1}$$

where

$$D(t) = \sum_{j=1}^{\min(t,\ t_x^b,\ ta)} Nb(j,t)p_1^b(j) + \sum_{j=1}^{\min(t,\ t_x^o,\ ta)} \left(\sum_{o=s,e,f,fm} N_1^o(t+1-j)p_1^o(j) \right), \ \forall\, t \in [0, T_N]$$

$$Q(t) = \begin{cases} E(t)\mu Sp \left(1 - exp\left(-\frac{k}{Sp \cdot slw} \sum_{i=1}^{min(t,\ ta)} x_i(t) \right) \right), & \forall t \in [1, T_N] \\ Q_{seed}, & t = 0 \end{cases}$$

$$Nb(i,t) = N_1^b(t-i+1)\prod_{k=1}^{i-1}(1-u_{i-k}(t-k)), \forall t \in [1, T_N+1]\,.$$

where $Nb(i,t)$ is the number of leaves of age i at cycle t.
Eq.(9.1) is explained by an example and is illustrated in Fig.9.1. The first item on the right side of Eq.(9.1) is the total biomass of leaves left after the harvesting operation at the end of the last growth cycle. It is marked in circle without cross on the left image of Fig.9.1, the ones marked in circle with cross being harvested. The last item on the right side of Eq.(9.1) is the biomass obtained at the current growth cycle. It is filled in orange on the right image.
The state equation as expressed by Eq.9.1 can be written in vector notation as expressed by Eq.9.2.

$$\mathbf{X}(t+1) = \mathbf{F}^t\left(\mathbf{X}(t), \mathbf{U}(t)\right), 0 \le t \le T_N \tag{9.2}$$

Therefore, the yield of leaf harvest at cycle t is

$$y(t) = \mathbf{U(t)}^T\mathbf{X(t)} \tag{9.3}$$

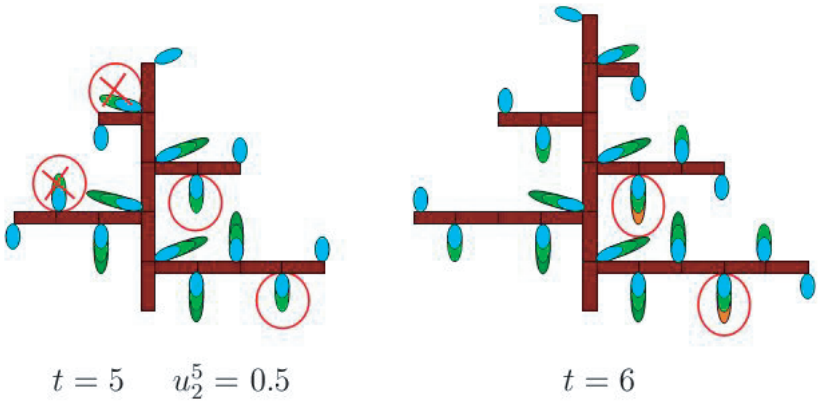


Figure 9.1: Illustration of state equation.

The total yield of leaf harvest at the final age of plant is thus given by

$$Y = \sum_{t=1}^{T_N} y(t) = \sum_{t=1}^{T_N} \mathbf{U(t)}^T \mathbf{X(t)} = \sum_{t=1}^{T_N} \mathbf{g^t}\left(\mathbf{X(t)}, \mathbf{U(t)}\right) \tag{9.4}$$

where T_N is the final plant age.

9.2 Gradient of the objective function

Because gradient based method is always used to solve minimization problems, we transform our maximization problem to minimization one.

$$J' = \sum_{t=1}^{T_N} -\mathbf{g^t}\left(\mathbf{X(t)}, \mathbf{U(t)}\right) = \sum_{t=1}^{T_N} \mathbf{G}^t\left(\mathbf{X}(t), \mathbf{U}(t)\right) \tag{9.5}$$

Now, we can write the formula of the optimal control problem as Eq.(9.6).

$$\begin{cases} J' = \sum_{t=1}^{T_N} \mathbf{G}^t\left(\mathbf{X}(t), \mathbf{U}(t)\right) \\ \mathbf{X}(t+1) = \mathbf{F}^t\left(\mathbf{X}(t), \mathbf{U}(t)\right), \qquad 0 \le t \le T_N. \end{cases} \tag{9.6}$$

By using the method described in section 4.3.2, the gradient of the objective function with respect to $\mathbf{U}$ as expressed by Eq.9.5 is given by

$$\nabla_{\mathbf{U}} \mathbf{J'}\left(\mathbf{U}\right) = \mathbf{G}^t_{\mathbf{U}}\left(\mathbf{X}(t), \mathbf{U}(t)\right) - \boldsymbol{\lambda}^{t+1^T} \mathbf{F}^t_{\mathbf{U}}\left(\mathbf{X}(t), \mathbf{U}(t)\right), \ \forall\, t \in [1, T_N - 1], \tag{9.7}$$

And the adjoint equation of the optimal control problem, with the final condition $\lambda^{\mathbf{T_N}} = \mathbf{0}$, is given by

$$\boldsymbol{\lambda}^t = \mathbf{G}^t_{\mathbf{X}}\left(\mathbf{X}(t), \mathbf{U}(t)\right) + \boldsymbol{\lambda}^{t+1^T}\mathbf{F}^t_{\mathbf{X}}\left(\mathbf{X}(t), \mathbf{U}(t)\right) = 0,\ \forall\ t \in [2, T_N - 1]\,, \tag{9.8}$$

9.3 Description of the controlled plant

The plant we optimized has only one order branch, and the maximum number of growth units on the branches is 20. The plant age is 70 growth cycles. The functioning time and the expansion time of organs are 12 growth cycles and 3 growth cycles, respectively. The leaves on the stem and on the branches have the same sink strength, but the internodes of the stem and the ones of the branches are different. Besides, the internodes have rings round piths. Mode one *Number of leaves* is chosen for the calculation of the internode secondary growth. The GreenLab parameters for the control plant are given in Table 9.1.

Table 9.1: Parameter values of the GreenLab model for the control plant for leaf harvest. The definitions of the parameters are referred to chapter 2.

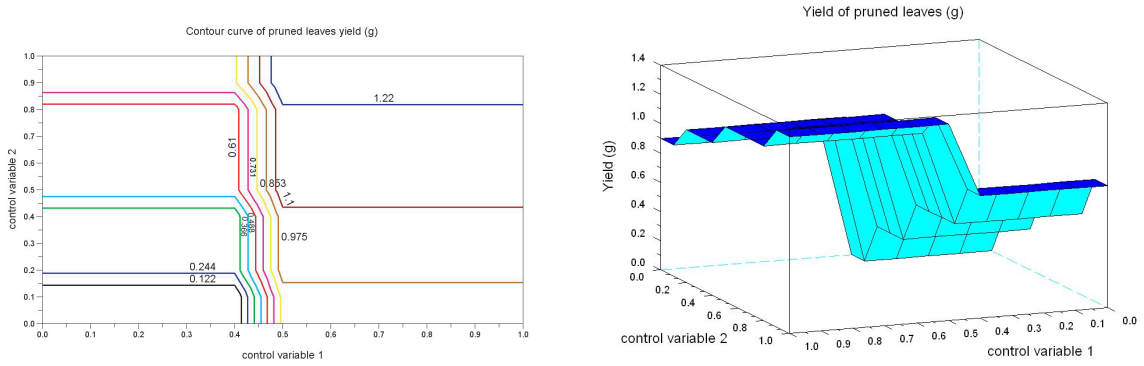
Parameter	Value	Unit
$E(t)$	1	–[1]
μ	0.05	g/cm^2
Sp	10000	cm^2
k	1	–
slw	0.04	g/cm^2
Q_{seed}	1	g
ta	12	growth cycle
t_x^o	3	growth cycle
P_m	2	–
P_p^b	1,1 from PA 1 to PA P_m	–
P_p^e	0.8,0.64 from PA 1 to PA P_m	–
a^b, b^b	1, 1	–
a^e, b^e	2, 0.8	–

[1] "–" means the parameter is unitless.

9.4 Convexity analysis of the optimal control problem

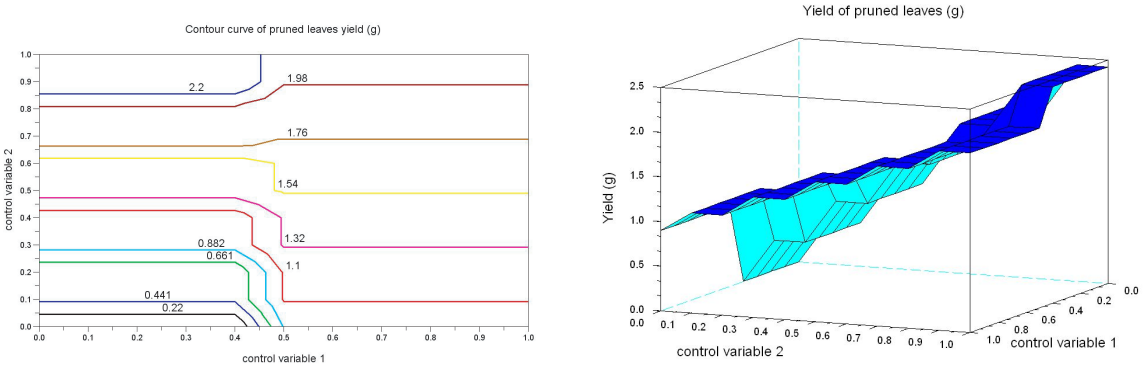
Before we search the optimal control variable values using the gradient information as expressed by Eq.9.7 and Eq.9.8, we first study the convexity of the optimal con-

trol problem. It helps us analyze the problem complexity and choose an appropriate algorithm.
Suppose that at each growth cycle, only the leaves of age t_l will be harvested. For this case, the number of control variables to be optimized is $T_N - t_l + 1$. Extended to general cases, the number of control variables to be optimized is much more than $T_N - t_l + 1$. The optimal control problem studied in this chapter is thus much dimensional.
As the optimal control problem investigated in this chapter is much dimensional, it is difficult to analyze the geometrical properties of the general cases of the problem. Here, we consider only two control variables, while the others are set to be zero.



(a) (b)

Figure 9.2: (a) Contour curve (b) 3D representation of the yield of harvested leaves of age 1. The leaves are harvested at growth cycle 1 and at growth cycle 3.



(a) (b)

Figure 9.3: (a) Contour curve (b) 3D representation of the yield of harvested leaves of age 1. The leaves are harvested at growth cycle 1 and at growth cycle 5.

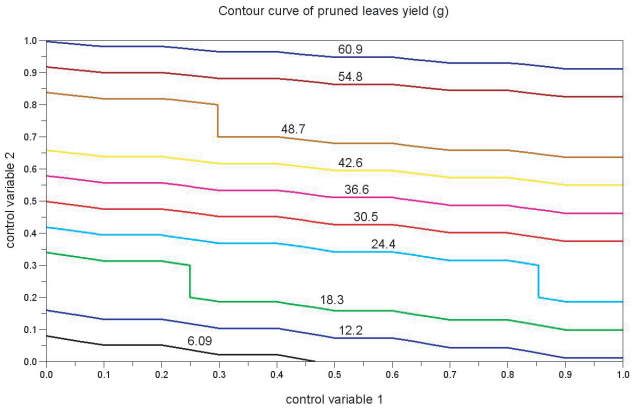


Figure 9.4: Contour curve of the yield of harvested leaves of age 3. The leaves are harvested at growth cycle 7 and at growth cycle 10.

Three examples are given and the simulation results of the yield of leaf harvest with respect to control variable values are shown in Fig.9.2, Fig.9.3 and Fig.9.4 respectively. From the figures, we found that the problem is not convex. Although each of them is monotonous, their tendencies are different. It may raise difficulties to optimization algorithms, regardless of gradient-based methods or other kinds of methods, as the problem is multimodal and complex.

9.5 Results

9.5.1 Single objective optimization problem

The objective of the optimal control problem investigated in this subsection is to find the optimal pruning strategy in order to obtain maximal yield of leaf harvest. The objective function J of the problem in this section is as expressed by Eq.9.9, which is subject to the boundary condition of the control variables.

$$\begin{aligned} J &= \sum_{t=1}^{T_N} g^t\left(\mathbf{X}(t), \mathbf{U}(t)\right) \\ s.t. \quad & 0 \leq \mathbf{U}(t) \leq 1, \qquad \forall\ t \in [0, T_N] \\ & u_i(t) = 0, \qquad \text{if } i \neq t_l, \end{aligned} \tag{9.9}$$

where t_l is the age of leaves to be harvested.
The functioning time of blades is 12 GCs, the age of leaves that can be harvested is thus from 1 GC to 12 GC. The optimal yield of leaf harvest with respect to the age of harvested leaves is shown in Fig.9.5 and listed in Table 9.2. The optimal results found by the gradient based method and PSO are similar. However, for some cases, the optimal

solution found by the gradient based method is better than the one found by PSO, while for other cases, the optimal solution found by PSO is better than the one found by the gradient based method. The problem is much dimensional and multimodal, which is a challenge for both optimization algorithms.

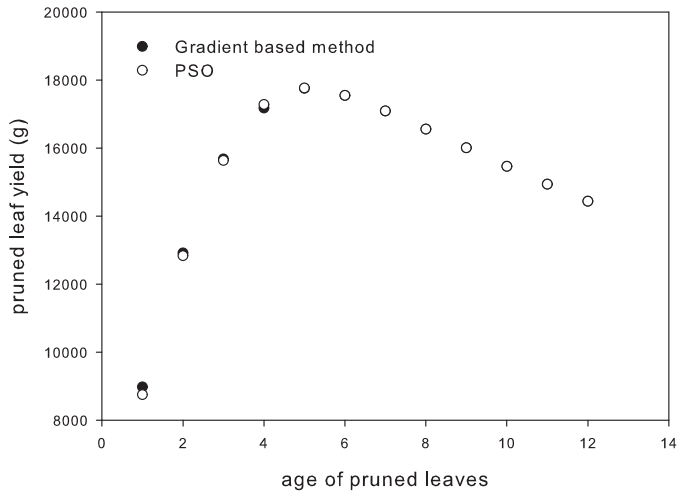


Figure 9.5: Optimal yield of leaf harvest with respect to the age of harvested leaves.

One specific case is chosen to analyze the optimal strategy. The optimal pruning strategy by gradient based method when we harvest leaves of age 9 is shown in the figure at the top of Fig.9.6. The corresponding optimal yield of leaf harvest is 16010.8g. The control variable value $u_9(t)$ is 0, if plant age t is less than 9. It is because there is no leaves of age 9 during that period. Considered the fact that the number of leaves that be harvested should be integer, the real value of the number of harvested leaves derived by the program is rounded to its nearest integer value. If .5 is happened, we chose the integer in the upper bound. At growth cycle 9, there is only one leaf of age 9, which is initiated at the beginning of the plant growth. Even though the control variable value is 0.5 at growth cycle 9, according to the rule of rounding value, the leaf of age 9 is harvested. The leaves of age 9 at growth cycle 10 are harvested in the similar way. The number of leaves alive after harvest at each growth cycle with respect to the corresponding control variable is shown in the figure at the bottom of Fig.9.6.
In the control plant, the expansion duration of leaves is 3 growth cycles. Hence, the leaves whose age is older than 3 are source organs, which do not compete for biomass against other organs. However, the control plant has secondary growth. The biomass demand for the internode secondary growth is proportional to the total number of leaves alive. Even though the leaves themselves whose age is older than 3 do not compete for biomass, they enhance the competition capacity of the internode secondary growth for biomass against other organs. The reduction of the number of leaves of age 9 due to

Table 9.2: Comparison of the optimal yield of leaf harvest with respect to the age of harvested leaves by the gradient based method and the Particle Swarm Optimization.

Harvested leaf age	Algorithm	
	Gradient based	PSO
1	8979.7	8740.5
2	12920.5	12838.1
3	15682.2	15639.3
4	17179.2	17286.9
5	17760.1	17770.1
6	17549.1	17552.0
7	17092.8	17096.1
8	16561.6	16561.6
9	16010.8	16010.8
10	15468.1	15468.1
11	14943.2	14934.2
12	14439.7	14439.7

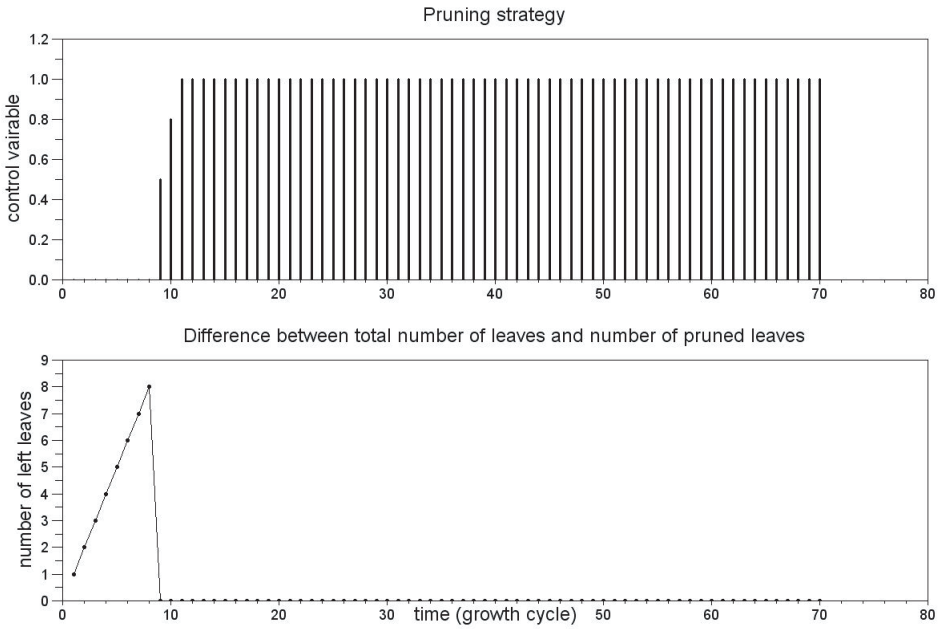


Figure 9.6: Optimal pruning strategy for harvested leaves of age 9 and the number of leaves alive of age 9 after harvest at each growth cycle.

harvest will decrease the competition of the internode secondary growth. Moreover, it does not affect the plant biomass production, which is determined by the leaf surface area (LAI). Even though the LAI of the plant with harvested leaves decreases, as shown

in Fig.9.7, it already saturates when the difference from the LAI of the plant without harvest occurs. Hence, under such condition, to have maximal leaf yield, the optimal pruning strategy is to harvest all the available leaves of age 9.

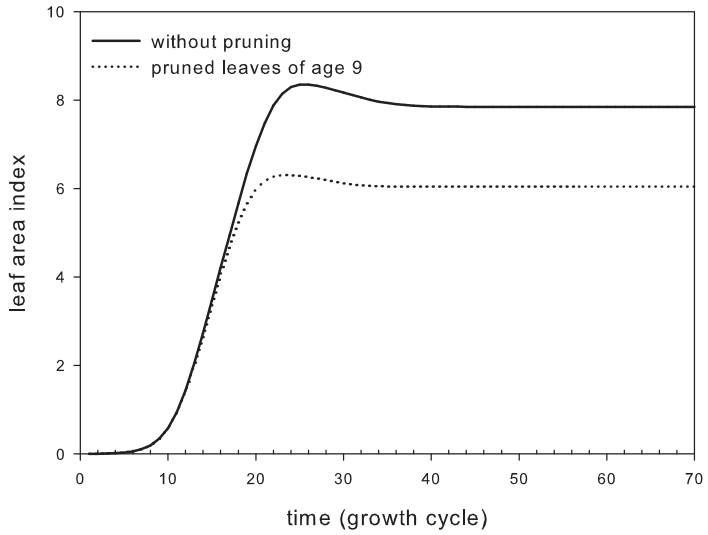


Figure 9.7: Comparison of the leaf area index.

The cases where the age of harvested leaves is less than 3 growth cycle are more complicated. For these cases, the leaves to be harvested are both source and sink organs, which have not finished expansion. The optimal result of the case where the age of harvested leaves is 1 is shown in Fig.9.8, and the optimal yield of harvested leaves is 8879.76 g.
A more complicated case is optimized, where the leaves of age 1 and age 2 will be harvested. The optimal pruning strategy is as shown in Fig.9.9. The corresponding optimal yield of harvested leaves is 12940 g. If the leaves of age 1 are harvested, there is no leaves of age 2 in the following time. Moreover, the biomass of leaves of age 1 is less than the biomass of leaves of age 2, as the older leaves obtain more biomass due to their expansion duration. Hence, for the optimal strategy to obtain maximal yield of harvested leaves, the number of the leaves of age 1 that are harvested are chosen to be as less as possible.

9.5.2 Multi-objective optimization problem

Considering the application of leaf harvest in reality, besides the benefit from the yield of leaf harvest, the final benefit should minus the cost from the harvest operation. Hence, in this section, we formulate a multi-objective optimization problem of pruning strategy.

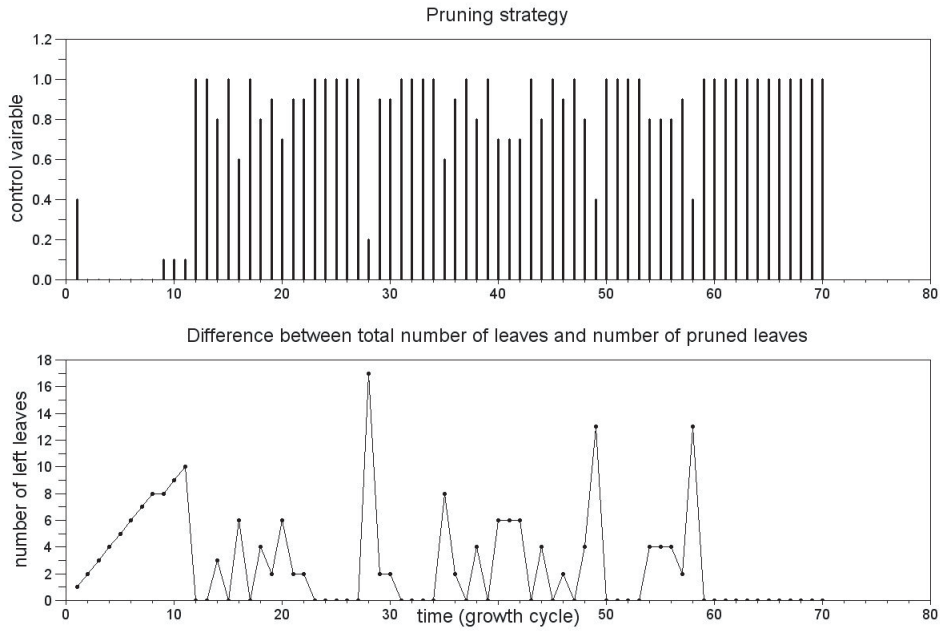


Figure 9.8: Optimal pruning strategy for harvested leaves of age 1 and the number of leaves alive of age 1 after harvest at each growth cycle.

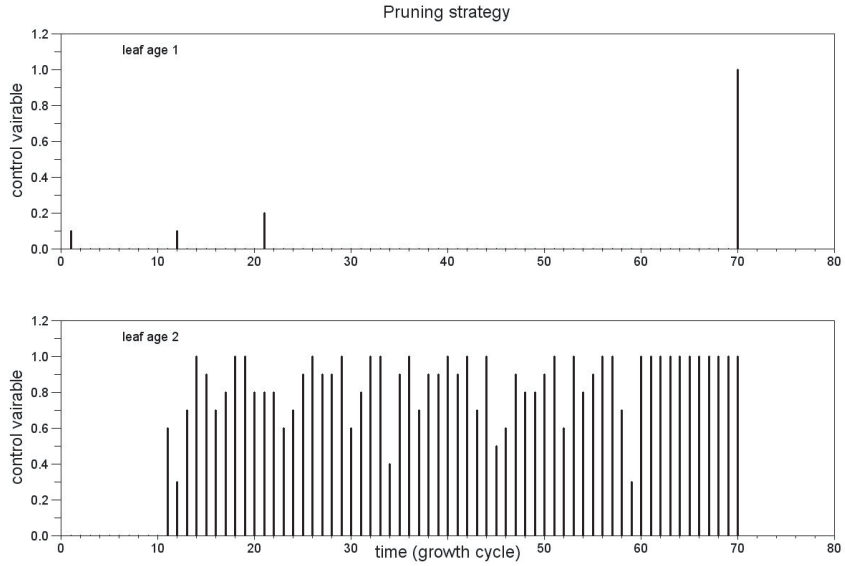
The objectives are maximization of the yield of leaf harvest and minimization of the frequency of harvest operation simultaneously. The control variables are restricted to bool values, i.e. the available leaves are all harvested or all not harvested. The formula for the multi-objective optimization problem is given by Eq.9.10.

$$\begin{aligned} Z &= (J, -C) \\ \text{subject to } \mathbf{U}(t) &= 0 \text{ or } 1 \\ u_i(t) &= 0, \text{ if } i \notin [t_l,\ t_r] \\ u_i(t) &\in \mathbf{U}(\mathbf{t}) \end{aligned}$$

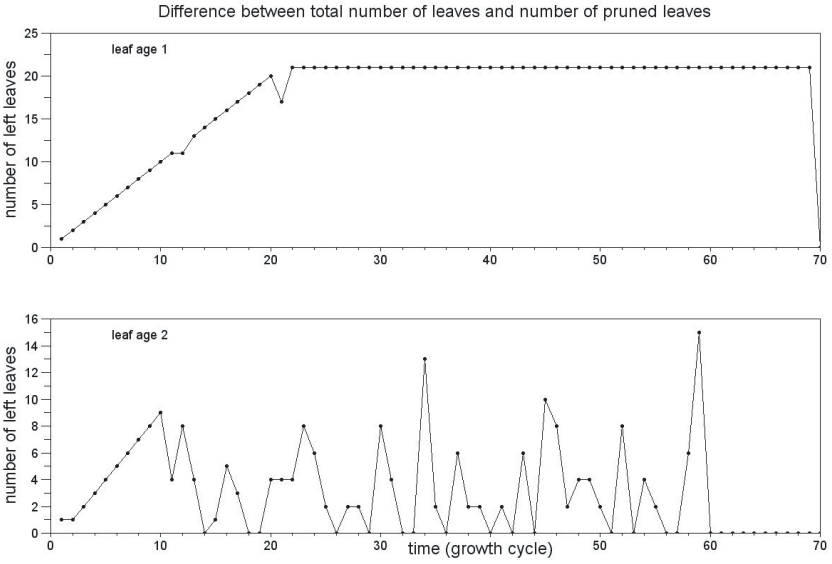
where

$$\begin{aligned} J &= \sum_{t=1}^{T_N} \mathbf{U}(\mathbf{t})^T \mathbf{X}(\mathbf{t}) \\ C &= \text{frequency of harvest} \\ &= \sum_{t=1}^{T_N} bool\left(max\left(u_i(t),\ i = 1, 2, \ldots, ta\right) > 0\right) \end{aligned} \tag{9.10}$$

where $[t_l, t_r]$ is the age range of leaves to be harvested.



(a)



(b)

Figure 9.9: (a) Optimal strategy of leaf harvest, whose age is 1 (at the top) and 2 (at the bottom). (b) Number of leaves alive of age 1 (at the top) and age 2 (at the bottom) after harvest at each growth cycle.

The optimal solution of the multi-objective optimization problem as expressed by Eq.9.10, *Pareto front*, is shown in Fig.9.10. Considered one of the objectives of the multi-objective optimization problem, which is minimization of the frequency of harvest operation, it is obvious that the minimum cost should be 0, no harvest operation

being applied. And the corresponding yield of leaf harvest is definitely 0. Hence, the left extreme of the *Pareto front* corresponds to the origin of coordinates. The right extreme of the *Pareto front* corresponds to the problem of maximization of the yield of leaf harvest, no matter what the cost is, which is the single optimization problem investigated in the previous section. The optimal strategy corresponding to the *Pareto front* when the frequency of the harvest operation is 10 is shown in Fig.9.11. The corresponding optimal yield of leaf harvest is 1255.03 g. From Fig.9.11, we found that the harvest operation is applied late, in order to harvest the leaves with large biomass, and the yield of leaf harvest could thus be much.

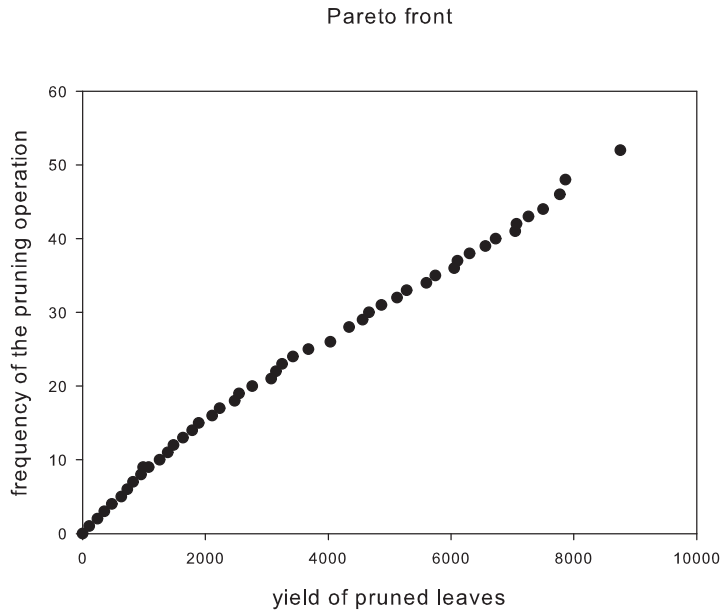


Figure 9.10: Optimal solution of the multi-objective optimization problem, the age of harvested leaves is 1.

9.6 Conclusion

This chapter illustrates the potential application of GreenLab for decision support in agriculture or forestry. The problem of maximization of the yield of leaf harvest with respect to pruning strategy is formulated. The classical optimal control algorithm is applied to GreenLab, and is used to solve the optimal control problem. Compared with the optimal results by the Particle Swarm Optimization algorithm, the classical optimal control algorithm by using variational approach and Lagrange theory is validated.
The results of the optimal pruning strategy revealed the optimal trade-off between source-sink dynamics during the plant growth, which is consistent with the results introduced in the previous chapters.

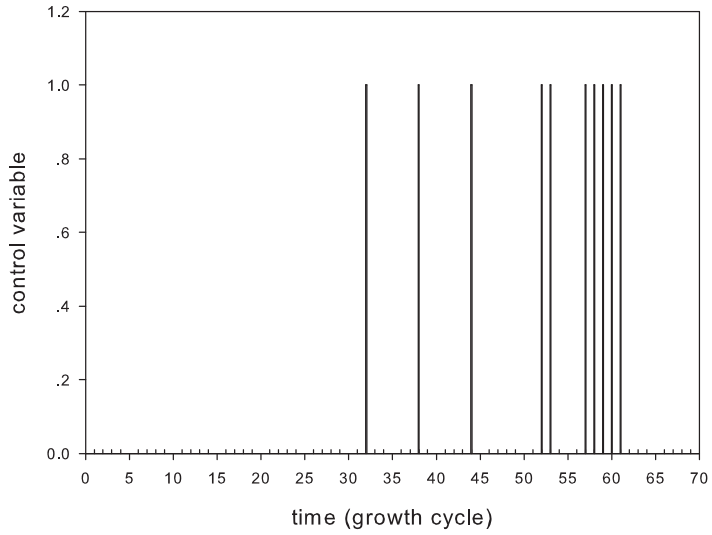


Figure 9.11: Optimal pruning strategy corresponding to the *Pareto front* shown in Fig.9.10, when the harvest operation is applied ten times, the age of harvested leaves being 1.

Moreover, the multi-objective optimization problem of maximization of the yield of leaf harvest and minimization of the frequency of the harvest operation is investigated. The problem is solved by the Particle Swarm Optimization algorithm. The optimal solutions provide the decisions to decision-makers according to their requirements. It makes GreenLab possible to be used for decision support.

Chapter 10

Analysis of tri-trophic ecosystem model and optimization of pest management

A tri-trophic ecosystem model is proposed in Chapter 3 to describe the interactions among plant (e.g. oil palm), insect pests and auxiliary insects. The levels of interaction of coexisting submodels characterized tri-trophic ecosystems as complex systems (Piqueira et al. [2009]), which induce a very heavy parametrization and thus difficulties to determine parameters accurately. Such tri-trophic ecosystem models inevitably contain uncertainties (Holland et al. [2009],Larocque et al. [2009],Peters et al. [2009]). One way for analyzing uncertainties of model parameters to model output variables is sensitivity analysis. It can distinguish two types of model parameters (Cariboni et al. [2007]): (1) the most influential parameters to model outputs, which need more accurate measures to reduce uncertainties related to the determination of the corresponding parameters and careful observations to enhance model reliability degree, and (2) the least influential parameters, which can be signed to any value in their feasible range without resulting in significant variations of model outputs.

Spraying pesticides is one of the pest management techniques. However, inappropriate application of pesticides may decimate natural enemies of insect pests, which may result in great increase of insect pest population to dramatic levels. This phenomenon has been observed widely (Trichilo and Wilson [1993]). Hence, in order to protect plants, it is crucial to rationalize applications of insect pest management techniques.

In this chapter, taking oil palm (*Elaeis guineensis*) as example of plant, we investigate three issues based on the tri-trophic ecosystem model. First, as we suppose that the parameters of GreenLab can be estimated by using the experimental data, the sensitivity of model outputs to the parameters in the insect population dynamics model is investigated using a global sensitivity analysis method. Second, we investigate the problem of parameter identification by using experimental data. Third, to enhance the efficiency, insect pest management techniques are optimized by the Particle Swarm Optimization

algorithm.

10.1 Setting of the GreenLab parameters and the time unit in the ecosystem model

10.1.1 Setting of the GreenLab parameters

Organogenesis for oil palm

To estimate the hidden parameters of GreenLab, organogenesis should be first set with the help of observation data of oil palm. Oil palm (*Elaeis guineensis*) involved in the ecosystem we studied is a single stem plant without branches. Flowers are on the stem. There is no secondary growth of internodes. For this species of oil palm, on average, one palm leaf is produced every 15 days, which is equal to one growth cycle in GreenLab. A palm leaf consists of a blade and a petiole. Leaves and internodes expand within one growth cycle. On average, the first bunch of fruits appears at the stem top between two and five years after planting and plant is around one year old at plantation time. The regularity of fruit appearance is difficult to induce due to variable environmental conditions. We suppose that fruits appear continuously in GreenLab. Hence, internodes older than two years bear a palm leaf and a bunch of fruits that have about one year delay of expansion.

Estimation of the hidden parameters of GreenLab

For established oil palm, the average weights of a blade, a petiole, an internode and a bunch of fruits are about 6.3 kg, 2.7 kg, 3.4 kg and 17 kg respectively. The leaf length is about 9 m, and the width is about 2 m. The number of leaves alive is about 40. These data are used to adjust the hidden parameter values of GreenLab. The adjusted hidden parameter values of GreenLab are listed in Table 10.1. With the adjusted parameter values listed in Table 10.1, the simulated results of the weight of a blade, a petiole, an internode and a bunch of fruits by GreenLab without considering pests and auxiliaries are 6.35 kg, 2.54 kg, 3.18 kg and 17.73 kg, which are close to the observation values.

10.1.2 Determination of the time unit in the ecosystem model

The main pest species of oil palm that we study is *Coelaenomenodera lameensis* Berti and Mariau. The hatching period is about 20 days. The longevity of juveniles is about 70 days (60 days for larvae and 10 days for nymphs). The longevity of adults is about three months. Each juvenile eats 4 cm^2 of leaf surface area during its whole life (Mariau and Morin [1972]), and each adult eats 1 cm^2 approximately. Even though the viability rate and fecundity are dependent on environmental conditions (Mariau and

Morin [1972], Mariau and Lecoustre [2004]), we set them constant values in this primary study. There are two kinds of auxiliaries: egg auxiliary (e.g. *Oligosita longiclavata*, *Achrysocharis leptocerus*) and larva auxiliary (e.g. *Closterocerus africanus*,*Sympiesis abruiana* Waterst, *Pediobius setigerus* Kerrich, *Cotterellia podagrica* Waterst). The hatching period is about 5 days. The longevity of juveniles is about 25 days (10 days for larvae and 15 days for nymphs). The longevity of adults is about one month. As the shortest time span among the plant organogenesis, pest and auxiliary development stages is 5 days, it is as time unit for the temporal discretization of the present tri-trophic ecosystem model. In addition, the oil palm of 7 years old, which is equivalent to 500 time steps, is studied in this chapter. Pests begin to attack oil palm when the age of oil palm is about three years and a half.
The parameters of the tri-trophic ecosystem model related to insects (pests and auxiliaries) are listed in Table 10.2, as well as the parameter space for sensitivity analysis and optimization. In addition, the types of the parameters are indicated. The following results of sensitivity analysis, parameter identification and optimization are based on these parameter values listed in Table 10.2 and Table 10.1. As eggs and juveniles are unable to fly and stay on the same leaves during their whole lives, and to reduce the number of parameters in the tri-trophic ecosystem model a little, we suppose that the viability rates for eggs and juveniles are identical in this chapter.

Table 10.1: Estimated parameter values of GreenLab for oil palm (*Elaeis guineensis*)

Parameter	Definition	Value
$E(t)$	environmental factor at time t (Eq.(2.23))	1
μ	light use efficiency (Eq.(2.23))	0.01
Sp	characteristic surface area related to plant crown projection (Eq.(2.23))	$1E+6$ cm^2
k	light extinction coefficient (Eq.(2.23))	1
slw	specific leaf weight (Eq.(2.23))	0.05 g/cm^2
ta	leaf functioning duration (Eq.(2.23))	120 TU[1]
$t_x^{b,s,e}$	expansion duration for blade (b), petiole (s) and internode (e) (Eq.(2.5))	3 TU
t_x^f	expansion duration for fruit (f) (Eq.(2.5))	42 TU
P_1^b	blade sink strength (Eq.(2.5))	1
P_1^s	petiole sink strength (Eq.(2.5))	0.4
P_1^e	internode sink strength (Eq.(2.5))	0.5
P_1^f	fruit sink strength (Eq.(2.5))	1.2
a^f, b^f	coefficients of fruit sink variation function (Eq.(2.7))	$1.2, 15$

[1] TU represents time unit.

Table 10.2: Parameter list of the tri-trophic ecosystem model

Parameter	Definition	RS[1]	RO[2]	Type[3]
C_{eg}	viability rate of pest eggs	[0.3,1]	[0,1] con	
C_{ju}	viability rate of pest juveniles	[0.3,1]	[0,1]	con
c	maximal viability rate of pest adults	[0.3,1]	[0,1]	con
λ	coefficient that determines the viability rate of pest adults	[0,1]	[0,1]	con
$\|F0(T_b)\|$	pest initial population size[4,5]	[1,10000]	[0,100000]	dis
ai, bi	coefficients of the beta function determining pest age distribution	[0,25]	[0,100]	con
M	pest fecundity	[1,500]	[0,100000]	dis
ar, br	coefficients of the beta function determining pest reproduction rate	[0,25]	[0,100]	con
ω	coefficient controling the mean of the Poisson distribution for pests	[0.01,10]	[0,10]	con
fly	coefficient that controls the strategy of pest adults to eat leaves[6]	[0,1]	[0,1]	dis
CP_{eg}	viability rate of auxiliary eggs	[0.3,1]	[0,1]	con
CP_{ju}	viability rate of auxiliary juveniles	[0.3,1]	[0,1]	con
cp	maximal viability rate of auxiliary adults	[0.3,1]	[0,1]	con
$P\lambda$	coefficient that determines the viability rate of auxiliary adults	[0,1]	[0,1]	con
$\|PF0(T_{bp})\|$	auxiliary initial population size[7]	[1,10000]	[0,100000]	dis
$PEsurP$	proportion of egg auxiliaries in the initial population	[0,1]	[0,1]	con
$PEai, PEbi$	coefficients of the beta function determining the age distribution of the initial population of egg auxiliary	[0,100]	[0,25]	con
$PWai, PWbi$	coefficients of the beta function determining the age distribution of the initial population of juvenile auxiliary	[0,25]	[0,100]	con
PM	auxiliary fecundity	[1,500]	[0,100000]	dis
Par, Pbr	coefficients of the beta function determining auxiliary reproduction rate	[0,25]	[0,100]	con
$P\omega$	coefficient controling the mean of the Poisson distribution for auxiliaries	[0.01,10]	[0,10]	con

[1] "RS" represents the parameter space for sensitivity analysis;
[2] "RO" represents the parameter space for optimization;
[3] "con" and "dis" represent that parameter values are continuous and integers respectively.
[4] "T_b" represents pest coming time.
[5] "$|y|$" denotes the summation of all the elements of a n-dimensional vector y.
[6] fly is a binary variable; if it is 1, adults eat any palm leaf of oil palm; if it is 0, adults eat the same palm leaves as when they were juveniles.
[7] "T_{bp}" represents auxiliary coming time.

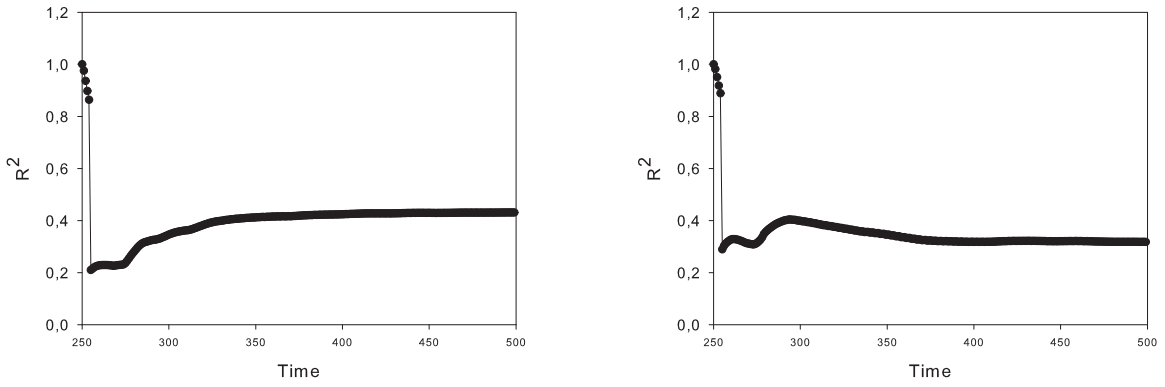
10.2 Sensitivity analysis

The sensitivity of the outputs to the parameters of the tri-trophic ecosystem model is first investigated in this section, in order to analyze the crucial factors that result in the outbreak of insects and to analyze the uninfluential factors to the model outputs, which help simply the model. As the purpose of modeling the tri-trophic ecosystem is to manage pests and to protect plant, the most interesting model output is total green leaf biomass, which is also an indicator of population dynamics. Moreover, the sensitivity of model outputs might vary with time. Hence, we investigate the sensitivity of total green leaf biomass with respect to time in the following subsections. In addition, the results of sensitivity of population dynamics to model parameters are also investigated to show that plant growth and population dynamics are interrelated.

10.2.1 Model linearity degree

Before sensitivity analysis, we first investigate model linearity degree, which is a criterion for choosing which sensitivity analysis methods should be used. One approximate indicator to assess model linearity degree is model coefficient of determination related to linear approximation. The bigger the coefficient value is, the better the model linearity degree is. The coefficient value associated with total green leaf biomass with respect to time is shown in Fig.10.1, where pests come at time 250 and auxiliaries come at time 300. No matter there are auxiliaries or not, the coefficient is small at any time, except the first five time steps when there are only pest adults of the initial population due to the reproduction rate and the age distribution of pests in the initial population ($ai = 0$, $bi = 100$, $ar = 16$, $br = 24$). It reveals that the model is highly nonlinear and parameter interactions cannot be neglected. Hence, the kind of sensitivity analysis methods highly suitable to linear models is not adapted to the present model. Therefore, we turn to a global method, which accounts for all the contributions of the parameters to the variances of outputs. As the tri-trophic ecosystem model is computationally expensive, Morris method (Morris [1991]) is chosen to assess the sensitivity of model outputs to model parameters.

Morris method is computationally cheap, model-independent, and can be used to identify non-influential parameters in the model. The means and standard deviations of the absolute values of outputs with respect to each parameter are Morris measures. Parameters can thus be ranked in order of importance by Morris measures. Bigger the values of Morris measures, more important the parameter is to model outputs. The procedure to compute Morris measures is as follows. Each parameter value is sampled randomly from h feasible values by discretizing its feasible range into h sections uniformly. l trajectories are randomly generated. Each trajectory is composed of $m+1$ sets of parameters, m being the number of model parameters. Between any two parameter sets of each trajectory, only one parameter varies, while among all the parameter sets



(a) (b)

Figure 10.1: Model coefficient of determination (R^2) for total green leaf biomass with time (a) without consideration of the interaction of auxiliaries (b) with consideration of the interaction of auxiliaries.

of each trajectory, each parameter varies once. As the number of model evaluations using Morris method is $l(m+1)$, Morris method is suitable to analyze computationally expensive models, which is the case for the tri-trophic ecosystem model that we study. The detailed information of Morris method is given in Morris [1991].

10.2.2 Sensitivity analysis in the absence of auxiliaries

In this subsection, we first analyze the sensitivity of model outputs to the parameters of pests, without considering the interaction of auxiliaries. The result of the sensitivity analysis is as shown in Fig.10.2. We consider three different stages: initial stage from the beginning of pest attack (time 250) to time 254, short-term stage (from time 255 to time 350 approximately) and long-term stage (from time 350 approximately to the end of plant growth). For the initial stage, as shown in Fig.10.2, the most important parameters for total green leaf biomass are the pest initial population size ($|F0(T_b)|$), maximal viability rate of pest adults (c) and coefficients of the age distribution function (ai, bi), and the least influential parameters are egg viability rate (C_{eg}) and coefficients of reproduction rate function (ar,br). The hatching period of eggs is 4 time steps. Hence, during the first 4 time steps, only adults eat leaves. Moreover, no matter how many pest eggs are, they do not impact the variation of leaf biomass for photosynthesis. Thereby, the parameters related to eggs have no impact on green leaf biomass, and the parameters related to the initial population (e.g. the initial population size, adult viability rate) are the most important. The number of eggs laid on plant is the result of the reproduction rate of female adults and of their ages. In addition, it is also controlled by the mean value of the Poisson distribution, which determines the final number of

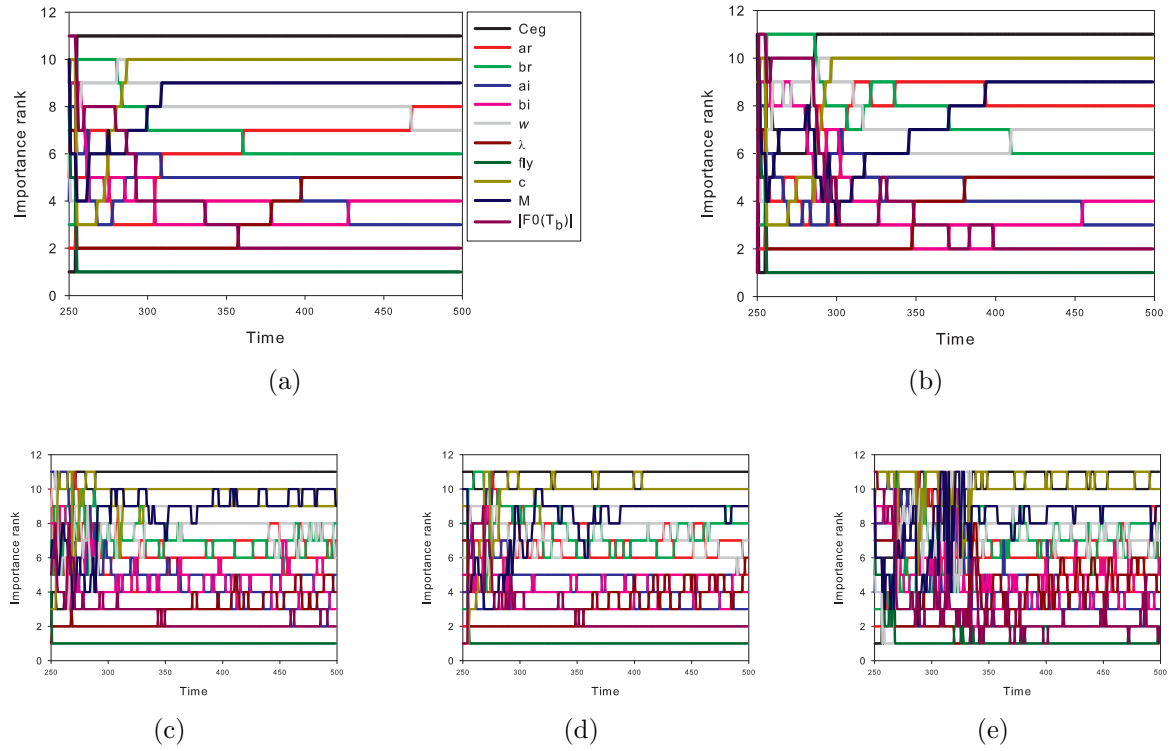


Figure 10.2: Importance ranks of parameters with respect to time, without consideration of the interaction of auxiliaries, based on expectation of (a) total green leaf biomass (b) plant biomass increment (c) total number of pest eggs (d) total number of pest juveniles and (e) total number of pest adults. 11 parameters are involved. In the axis named "Importance rank", 11 represents the most important parameter, while 1 represents the least influential one.

eggs that plant accepts. Therefore, the most important parameters for the number of pest eggs are the coefficients that control the mean of the Poisson distribution (ω), reproduction rate (ar, br) and age distribution function (ai, bi). The least influential parameter is the coefficient that controls the strategy of adults to eat leaves (fly). As during the initial stage there is no juvenile in pest population, the variation of total green leaf biomass is the result of the number of pest adults. Therefore, the result of the importance of the parameters for the number of pest adults is the same as the one for total green leaf biomass.

The initial population size is more important to the total green leaf biomass in the short term than in the long term. For the short-term stage, the most important parameters for total green leaf biomass are egg viability rate (C_{eg}), maximal viability rate of adults (c), the coefficients that control the mean of the Poisson distribution (ω) and reproduction

rate (ar, br), and the least influential ones are the coefficients that control the adult viability (λ) and the strategy of adults to eat leaves (fly).
For the long-term stage, the most important parameters for total green leaf biomass are egg viability rate (C_{eg}), maximal adult viability rate (c) and fecundity (M), and the least influential ones are the coefficient that controls the strategy of adults to eat leaves (fly) and the initial population size $(|F0(T_b)|)$. As the longevity of adults is finite, the initial population only has an effect for a short period. Since leaf damage results from pest juveniles and adults who evolve from eggs, the viability rates of eggs, juveniles and adults, and the fecundity are the most important.
On one hand, the total green leaf biomass is one of the decisive factors of plant biomass increment through photosynthesis. On the other hand, the quantity of the total green leaf biomass is the result of the interaction of pest eggs, juveniles and adults. Therefore, the results of parameter sensitivities for plant biomass increment and for pest population dynamics (eggs, juveniles and adults) are consistent with the results for the total green leaf biomass, except the initial stage where no juveniles are involved, adults are all from the initial population, and eggs have not become juveniles or adults.
Except the first 4 time steps, the most important parameters in the tri-trophic ecosystem model are egg viability rate (C_{eg}) and maximal viability rate of adults (c). The coefficient that controls the strategy of adults to eat leaves is unimportant for all the stages. The importance of the initial population size depends on the stage that we consider. As time increases, it becomes less and less important.

10.2.3 Sensitivity analysis in the presence of auxiliaries

In this subsection, auxiliaries come at time 280. Hence, before time 280, all the parameters related to auxiliaries are unimportant. After time 280, regarding total green leaf biomass, whatever the stage is, the most important parameters are pest egg viability rate (C_{eg}), maximal viability rate of pest adults (c) and auxiliary egg viability rate (CP_{eg}), whereas the least influential ones are the coefficient that controls the strategy of pest adults to eat leaves (fly), the coefficients of the initial age distribution function for auxiliaries ($PEai$,$PEbi$,$PWai$,$PWbi$). The importance ranks of parameters are shown in Fig.10.3. Auxiliary population size is dependent on pest population size, since pests are hosts of auxiliaries. Hence, parameters related to pests are more important than those related to auxiliaries.

10.3 Parameter identification

GreenLab model parameters controlling plant growth in the absence of pests can be estimated independently as listed in Table 10.7. Moreover, we do not consider auxiliaries in the first place. We only study the parametric identification of pest dynamics and of

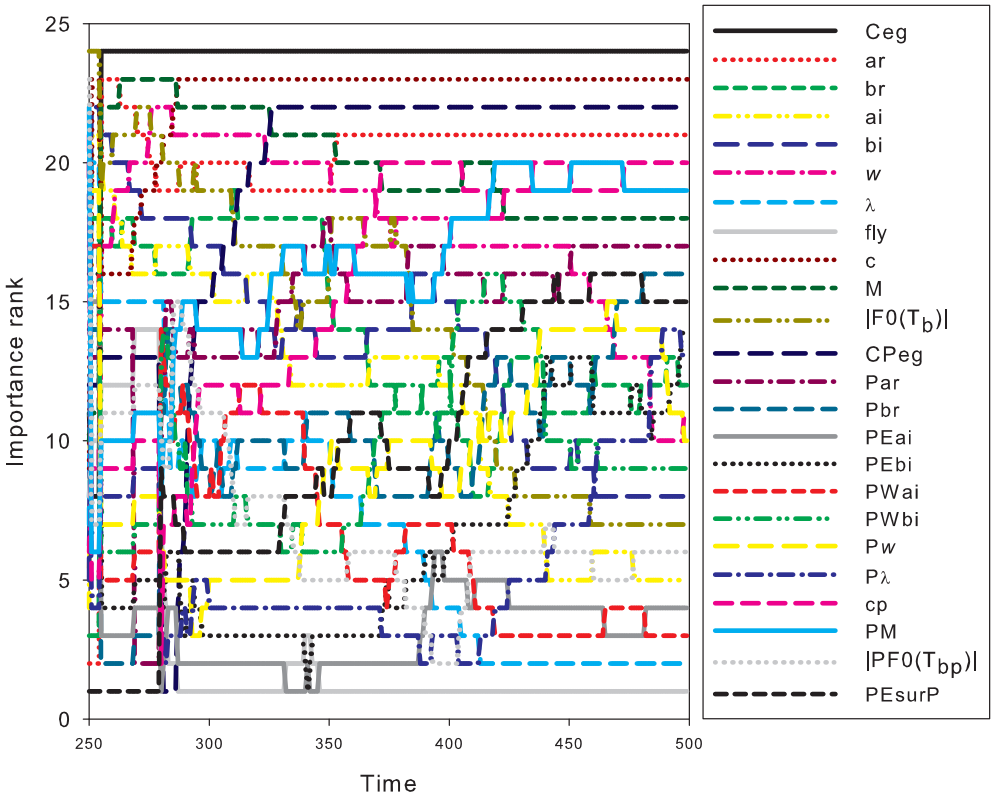


Figure 10.3: Importance ranks of parameters with respect to time for total green leaf biomass, with consideration of the interaction of auxiliaries. 24 parameters are involved. In the axis named "Importance rank", 24 represents the most important parameter, while 1 represents the least influential one.

plant-pest interactions in this section. To study the problem of parameter identification, we generate virtual experimental data as target data by simulating the tri-trophic ecosystem model with given parameter values. The target data include global data and detailed data. The global data consist of the total green leaf biomass of four sampled palm leaves initiated at time 150, 300, 360 and 420, the total number of eggs, juveniles and adults of pests on each sampled palm leaf, the total number of eggs, juveniles and adults of pests on oil palm. The detailed data consist of the number of eggs and juveniles of each age on each sampled leaf.

There are 11 parameters in the model without considering auxiliaries. We fit the present model from target data with respect to each parameter, supposing the other parameter values are known. The estimated parameter values are listed in Table 10.3. 8 out of all the parameters are estimated well. The simulation results with estimated value of one of the 8 parameters, maximal viability rate of pest adults (c), are shown in Fig.10.4.

While the other 3 parameters cannot be estimated, i.e. whatever the initial value is, it does not converge to its proper value. Compared with the results of sensitivity analysis, we found that these parameters that cannot be estimated are the ones with low sensitivity coefficients (i.e. less influential parameters). Even though the importance of some parameters out of all the parameters that can be estimated depends on time, for example, the initial population size $|F0(T_b)|$ is not important for the long-term stage, they can also be estimated well since the target data include the information at different stages.

Table 10.3: Result of parameter identification of the tri-trophic ecosystem model

Parameter	Initial value	Proper value	Estimated value
C_{eg}	0.7	0.9	0.8993
c	0.75	0.9	0.9000
ar	11	16	15.9824
br	20	24	24.0121
ω	0.2	0.5	0.5000
M	100	150	150
$\|F0(T_b)\|$	85	100	100
λ	1	0	0
fly	1	0	–[1]
ai	5	0	–
bi	90	100	–

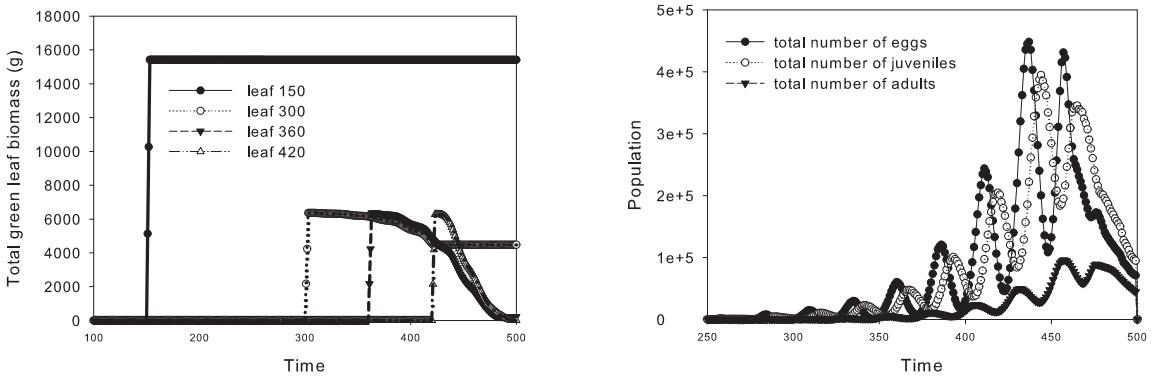
[1] "–" represents that the parameter cannot be estimated.

10.4 Optimization of pest management techniques

In this section, we investigate the problem of optimizing pest management techniques based on the simplified tri-trophic ecosystem model. According to the results of sensitivity analysis and parameter identification, we simplified a little the tri-trophic model by eliminating the coefficient that controls the strategy of leaf eating by pest adults. We suppose that the pest adults first eat the leaves where the juveniles they are evolved from were, and then they fly away to lay eggs. It not only simplifies the model parametrization but also simplifies the model computation.

10.4.1 Optimization of biological control

A method for pest management is biological control on pests by spreading a great number of auxiliaries on plants at certain time. It raises two difficulties: (1) the number of auxiliaries and (2) the treatment time. If auxiliaries are not deposited at appropriate



(a) (b)

Figure 10.4: Simulation results with respect to the estimated value of maximal viability rate of pest adults, compared with the virtual experimental data of (a) total green leaf biomass (g) (b) total number of pest eggs, juveniles and adults on plant. lines represent simulation results, and symbols represent virtual experimental data.

time, they will have no positive effect on the pest population, i.e. pests become extinct. Fig.10.5 illustrated the effect of the treatment time on the population dynamics of pests. When the auxiliary coming time is at time 279, both of the population of pests and auxiliaries reach equilibrium, as shown in Fig.10.5(a) and Fig.10.5(b), but there are still a great number of pests on plant, who will destroy plant leaves. Hence, it has no effect for eliminating pest population if the auxiliaries are deposited at time 279. When the auxiliary coming time is advanced at time 268, the pest population becomes extinct due to parasitism, and so does the auxiliary population since there is no host to feed on, as shown in Fig.10.5(c) and Fig.10.5(d). The results shown in Fig.10.5 are associated with the pest and auxiliary parameter values listed in Table 10.4 and Table 10.5.

Fig.10.5 reveals that it is crucial to determine the time when biological control is applied, in order to enhance its efficiency. A slight delay or advance of application will lead to significant different population dynamics. Hence, on this section, we optimize the number of deposited auxiliaries and the treatment time in order to reduce pests and to keep maximal green leaf biomass. Moreover, all auxiliaries are those deposited by man-made operations for biological control. The optimization problem of maximizing total green leaf biomass is formulated in Eq.10.1.

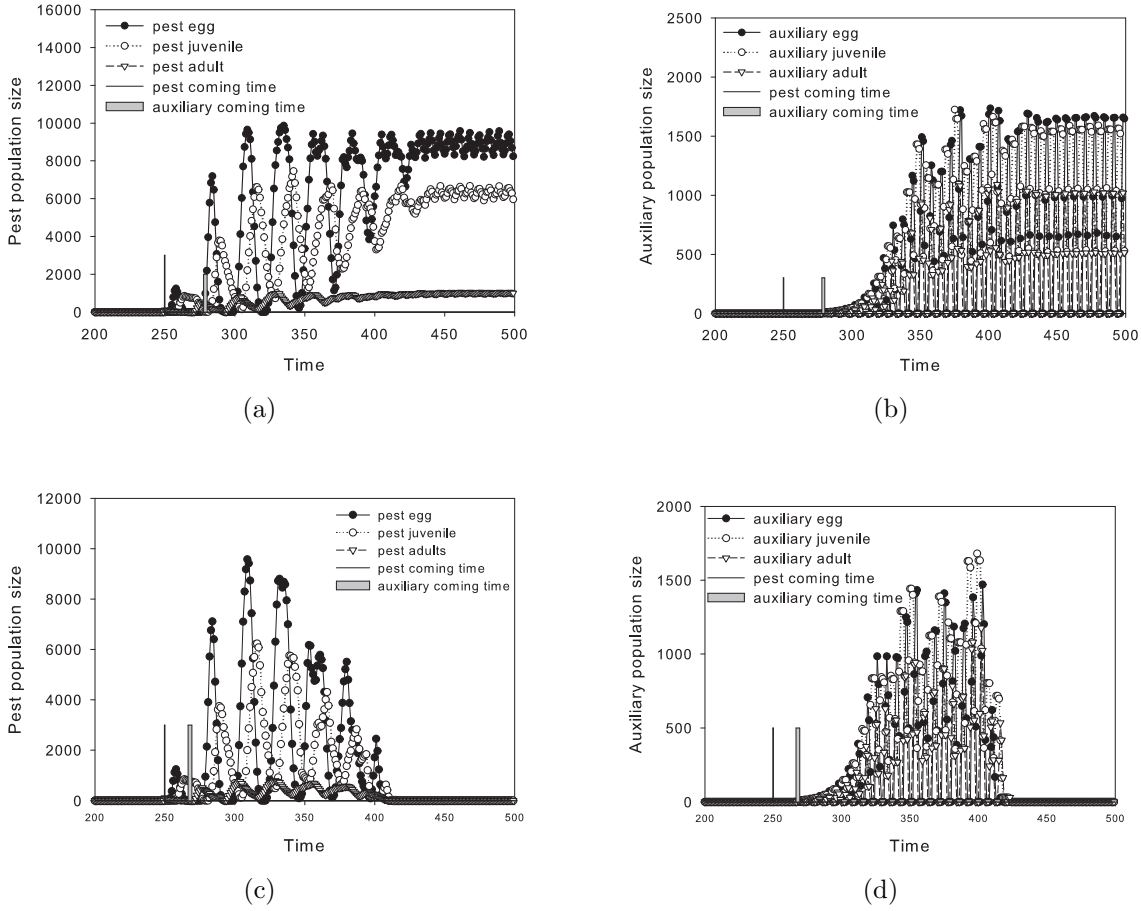


Figure 10.5: Interaction between pest and auxiliary dynamics with respect to auxiliary coming time. (a) Pest population reaches equilibrium when auxiliaries come at time 279 (b) Auxiliary population reaches equilibrium when auxiliaries come at time 279 (c) Pest population decreases when auxiliaries come at time 268 (d) Auxiliary population decreases when auxiliaries come at time 268

$$\begin{aligned} J &= \sum_{k=t-ta+1}^{t} X_f\left(k,t\right) \\ s.t. \quad & mod(k,Ra) \equiv 0 \\ & X_f(k,t) = g\left(X_f(k,t-1), PEF0, PWF0, T_{bp}\right) \end{aligned} \tag{10.1}$$

where J is the objective function of the optimization problem, which is the function of total green leaf biomass; ta is leaf functioning duration; $X_f(k,t)$ is biomass of the leaf initiated at time k when plant age is t, for photosynthesis; $mod(k,Ra)$ represents

Table 10.4: Parameter values for pest

Parameter	Definition	Value
C_{eg}	viability rate of pest eggs	0.88
C_{ju}	viability rate of pest juveniles	0.88
c	maximal viability rate of pest adults	0.88
λ	coefficient that determines the viability rate of pest adults	0
$\|F0(T_b)\|$	initial population size	100
ai, bi	coefficients of the discretized beta function that determines the age distribution of the initial population	$0, 100$
M	fecundity	150
ar, br	coefficients of the discretized beta function that determines the reproduction rate	$16, 24$
ω	coefficient that controls the mean of the Poisson distribution	0.5
T_b	pest coming time	250 TU

Table 10.5: Parameter values for auxiliaries

Parameter	Definition	Value
CP_{eg}	viability rate of auxiliary eggs	0.56
C_{ju}	viability rate of auxiliary juveniles	0.56
cp	maximal viability rate of auxiliary adults	0.56
$P\lambda$	coefficient that determines the viability rate of auxiliary adults	0
$\|PF0(T_{bp})\|$	initial population size of auxiliaries	10
$PEai, PEbi$	coefficients of the discretized beta function that determines the age distribution of the initial population	$0, 100$
PM	fecundity of auxiliaries (Eq.(3.2))	15
Par, Pbr	coefficients of the discretized beta function that determines the reproduction rate	$2, 3$
$P\omega$	coefficient that controls the mean of the Poisson distribution	1

modulo operation, finding the remainder of division of k by Ra; Ra is the ratio of growth cycle in GreenLab to the time unit in the tri-trophic ecosystem model; $PEF0$ and $PWF0$ are the initial population sizes of egg auxiliaries and juvenile auxiliaries, respectively; T_{bp} is time of spreading auxiliaries. The detailed formula of the present model is given in chapter 3.

The problem is a multimodal optimization problem. With two different optimal results of the initial population size of auxiliaries and of the treatment time by the PSO algorithm listed in Table 10.6, the same value of objective function of total green leaf biomass is obtained. With the optimal parameter values, pests are all killed by auxiliaries and plant is protected. The model parameter values on which the optimization procedure is based are listed in Table 10.7. As auxiliaries are all deposited by man-made

operations, $|PF0(T_{bp})| = 0$. From the optimal results, we found that to protect plant, auxiliaries should be deposited as early as possible.

Table 10.6: Optimal parameter values of the optimization problem of biological control.

Parameter	Optimal value		Range
	Experiment I	Experiment II	
$PEF0$	3602	537	$[1, 10000]$
$PWF0$	1902	4451	$[1, 10000]$
T_{bp}	254	292	$[250, 500]$[1]

[1] Pests come at time 250.

Table 10.7: Parameter values of the tri-trophic ecosystem model on which the optimization procedure is based

Parameter	Value	Parameter	Value
C_{eg}	0.9	CP_{eg}	0.7
C_{ju}	0.9	CP_{ju}	0.7
c	0.9	cp	0.7
ar	16	Par	2
br	24	Pbr	3
ω	0.5	$P\omega$	1
M	150	PM	15
$\|F0(T_b)\|$	100	$\|PF0(T_{bp})\|$	0
λ	0	$P\lambda$	0
fly	0	$PEai$	0
ai	0	$PEbi$	100
bi	100	$PWai$	0
$PWbi$	100	$PEsurP$	0.5

10.4.2 Optimization of chemical technique–pesticide application

One of the techniques to control pest population is to apply pesticides. In this section, we optimize the application time of pesticides, in order to obtain maximal total green leaf biomass. The optimization problem is solved by the Particle Swarm Optimization algorithm. Pesticides are supposed to only kill pest adults as well as auxiliary adults in this thesis. Juveniles and eggs are protected by leaves. We assume that pesticides kill 95% of adults, and 5% of adults are resistant to pesticides. The adults of pests and auxiliaries die once pesticides are applied within one time step. If the application time of

pesticides is restricted to be once, the effect of the pesticide application time on the total green leaf biomass with the assumptions mentioned above is shown in Fig.10.6, where only pests are considered and no auxiliaries are involved. The parameter values related to pests are listed in Table 10.4, except that the viability rates of pest eggs,juveniles and adults are identically 0.9.
From Fig.10.6, we found that the protection effect of pesticides on plant is strongly dependent on the application time. If pesticides are not applied at an appropriate time, they are useless and wasted. There is no protection effect on the total green leaf biomass. Compared the results shown in Fig.10.6, the most efficient time for pesticide application is at time 276. The corresponding pest population dynamics and the variation of the total green leaf biomass are shown in Fig.10.7. Even though at the most efficient time, a single pesticide application is not sufficient to avoid severe plant damages. At time 276, even though there is no egg and all pest adults are killed by pesticides, there are juveniles alive, which will become adults at the following time and lead to a new outbreak of pests, as shown in Fig.10.8(b): the adults at time 277 have evolved from the juveniles before pesticide application.

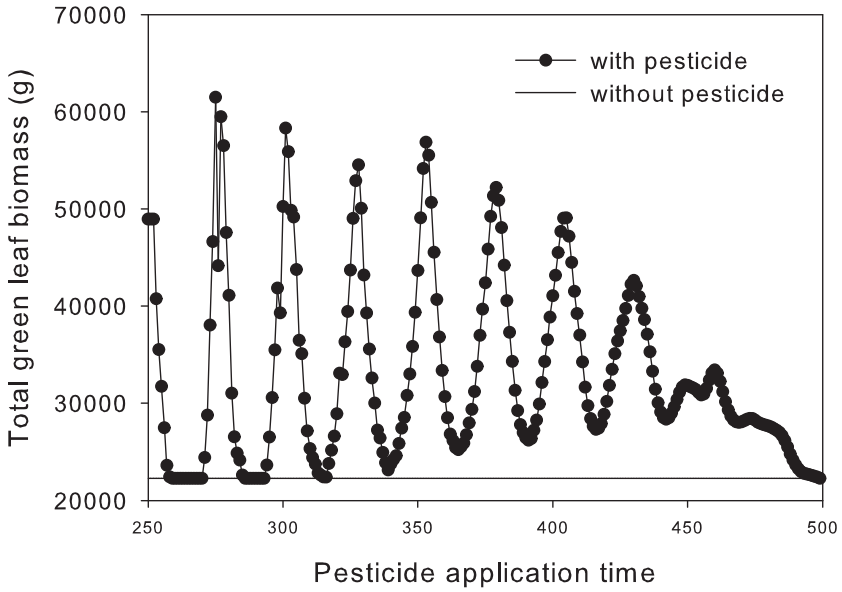


Figure 10.6: Effect of pesticide application on the total green leaf biomass at the end of plant growth. Pesticide is applied once.

From the simulation results shown in Fig.10.6, we found that it is not sufficient to apply pesticides once in order to protect plant. However, there are infinite possibilities to apply pesticides more than once. Moreover, pesticide application has some risks for auxiliaries (Trichilo and Wilson [1993]), and pesticides may lead to environmental

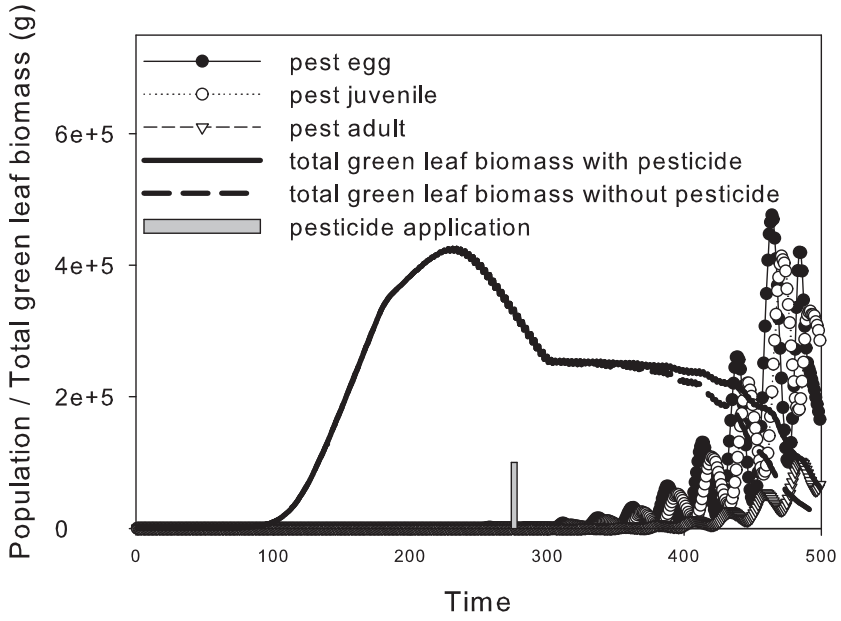


Figure 10.7: Pest population dynamics and the variation of the total green leaf biomass, where pesticide is applied once at time 276.

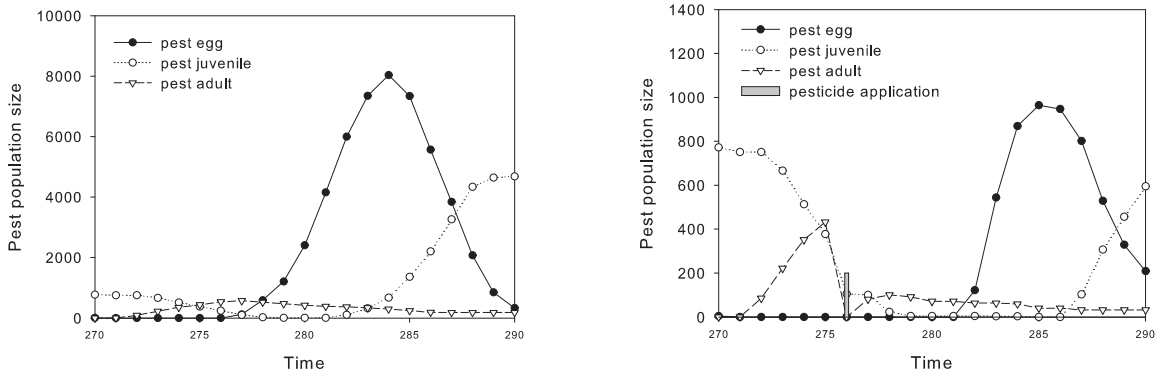
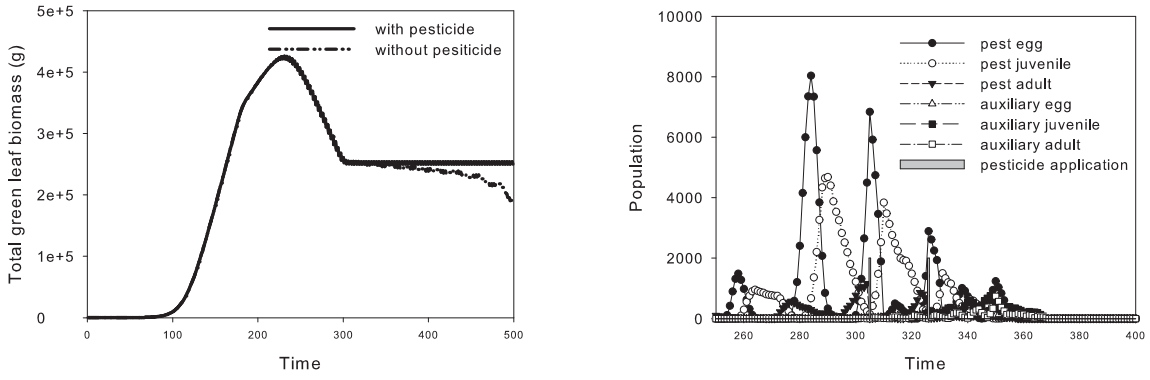


Figure 10.8: Pest population dynamics (a) without pesticide (b) with pesticide applied at time 276.

pollution. Hence, considered the above issues, the number of pesticide applications is restricted to two in this subsection. The parameter values of the present model on which the optimization procedure is based are listed in Table 10.7.

Pests come at time 250, and auxiliaries come at time 300. If we do not set any constraint

on application time, the optimal application time of pesticides should be obviously at time 250 and 251. It is not possible to catch the trace of pests once they come in reality. Hence, we set a constraint that after 10 time steps of pest coming, pests can be traced and pesticides could be applied. Under this constraint, the optimal application time of pesticides is at time 305 and 326. The protected effect is shown in Fig.10.9. With the combination of auxiliries and pesticides, pests disappear and plant is protected efficiently.



(a) (b)

Figure 10.9: Simulation results with the optimal pesticide application. (a) Total green leaf biomass (g) (b) Pest and auxiliary population dynamics with pesticides

10.5 Conclusion

In literature (Trichilo and Wilson [1993],Mariau and Lecoustre [2004]), the viability rate of pests is regarded as the crucial factor that results in the outbreak of pests. This fact is confirmed by the results of the sensitivity analysis of the present model. A slight variance of viability rate will result in a significant variance of total green leaf biomass. Hence, to enhance model accuracy, pest viability rate needs to be well adjusted from experimental data. Moreover, the results of sensitivity analysis help us simplify the model.

Comparing the results of sensitivity analysis and parameter identification, we found that parameters that contribute greatly to model output variation can be identified. The least influential parameters are difficult to estimate. Variations of their values have slight or even no effect on output variation. Hence, for given observation data, their values spread in a broad range.

The study on parameter estimation also revealed the difficulty of estimating the whole set of parameters simultaneously. It leaves space for model improvement. The opti-

mization results revealed that to protect plants, techniques should be applied as early as possible, with the consideration of not eliminating auxiliaries.

Chapter 11

Conclusions and perspectives

The thesis investigated theoretically and systematically the effect of model parameters and environmental biotic exogenous factors on plant yield, based on the functional-structural plant growth model GreenLab by using optimization and optimal control techniques. The thesis developed an insect population dynamics model and implemented its interaction with GreenLab, in order to study the plant growth with the interaction of environmental biotic exogenous factors. As GreenLab used a sink-source submodel to simulation plant biomass production and partition, it is able to analyze the plant growth behavior from the source-sink relation's point of view using GreenLab. The optimal results for different optimization problems revealed that the optimal plant yield is the result of the optimal balance between sources and sinks. Even though the work in this thesis is theoretical study and the results need to be validated by experiments in further step, the results for the plant species studied in this thesis are consistent with the results found by researchers through experimental observations, which lays a foundation for the application and extension of the work in this thesis. To complete the investigation about plant yield optimization, the following issues could be considered in further extension work of the thesis.

- GreenLab model selection

  Applications of optimization and optimal control have been done on the primary variant of GreenLab, i.e. determinate plant development. Different optimization problems were investigated, in order to enhance the yield of different kinds of economic valuable organs. The optimal results derived from the problems revealed the similar physiological information, even though the optimization objectives and the kind of problems are different. These similar information thus help us understand plant growth mechanism, i.e. source-sink relations. On the basis of the primary work in this thesis, the optimization could be applied to more complex variants of the GreenLab model, to valid the primary results of optimal physiological information, or to explore new outcome that may help deeply understand plant growth.

- New methodologies of model analysis and for applications

  Plant growth is influenced by both genotype and environmental conditions, which is widely recognized and accepted, as introduced in chapter 1.2 for English version (and chapter 1.1 for French version). Plant genes may mutate in order to adapt the changes of environmental conditions. It is acknowledged as plant evolution. The self-adjustment of plant to environment reflects that plant growth is a learning process and plants possess intelligence like animals (Trewavas [2003]). Hence, plant growth is plastic in terms of both physiology (e.g. transpiration, assimilation and biomass partition) and morphology (e.g. organ size). Therefore, the inputs (i.e. model parameters) and outputs (e.g. plant yield) of the GreenLab model are not uncertain. They are stochastic variables.

  So far, the methods for parameter estimation under given observations of plant output (e.g. plant yield, plant geometrical information (length and diameter of plant organs)) used on GreenLab are non-linear least square approach and optional Particle Swarm Optimization. Theoretically speaking, the results derived by using these methods are numerical and determinate. Practically speaking, non-linear least square method is iterative, gradient-based. The optimal results derived from it depend on the selection of initial values. Different initial values result in different final estimated values, even the initial values are slight different. The worst situation will occur that if the initial values are far from the proper ones, the method will not converge. The criterion to choose the estimated values as the appropriate ones is root mean square error, i.e. to choose the estimated values with which the summation of the differences between the simulated outputs of the model and the corresponding observation data is minimum. However, this criterion just guarantees that the chosen estimated values are appropriate in the mathematical point of view, not in the botanical or biological point of view.

  Bayesian-based methods provide a quantification of uncertainty on the basis of probabilities, while not numerical determinate values. It will provide posterior probabilities of concerned factors based on the prior probabilities of the concerned factors and the conditional probabilities of observation data given specific values of the concerned factors. Considered the stochastic properties of plant growth and the drawbacks of the methods currently used, Bayesian-based methods may be possibly imposed to the GreenLab model for model analysis and further applications.

- Sample limitation

  Besides understanding physiological mechanisms of plant growth, the objectives of modeling plant growth physiologically and architecturally is to enhance plant yield, which is the most interesting and more significant for model application and is more attractive for people in many fields. In this thesis, we have demonstrated

how optimization problems of yield improvement were formalized and optimization algorithms were used on GreenLab for breeding of ideotype. We assumed that the concerned parameters that we optimized are genetic, which are less influenced by environmental conditions. However, the assumption has not been well proved yet. For maize, the parameters are shown stable to different environmental conditions (temperature, planting density) (Guo et al. [2006],Ma et al. [2007],Ma et al. [2008]), which is chosen as the control plant to optimize fruit yield in this thesis. While for tomato, some of parameters vary with environmental conditions (Dong et al. [2008]). To investigate genetic property of parameters, huge number of experiments with different genotypes and different environmental conditions are needed. Moreover, for each experiment, huge number of plants are needed to be measured. There are dozen of parameters related to physiological processes in GreenLab. To cover the variable space, number of thousands of data are required. This is the challenge for genetic study of GreenLab. It is also a common challenge in statistical learning, which is called curse of dimensionality. Moreover, the time consumed by GreenLab is another limitation. Hence, if statistical learning methods are used, the methods that are suitable for computationally expensive models and for small number of samples are needed to be explored and tested, as well as sensitivity analysis methods.

Considered the issues mentioned above, the extended GreenLab model will be able to not only predict plant yield but also evaluate the reliability of prediction results. Based on the evaluation results, more flexible strategies for ideotype breeding and cultivation could be made, and the optimization of plant yield could thus be implemented. Through the further investigation, the theoretical study of GreenLab will turn to the practical study. Hence, the GreenLab model will provide direct decision aid in agriculture.

Part IV

Appendix

Appendix A

Notations

Table A.1: Parameters of GreenLab

CA	chronological age
PA	physiological age
P_m	maximal PA
$m_{pq}(n,t)$	metamer of CA n and PA p bearing axillary buds of PA q when plant age is t
$s_p(t)$	bud of PA p when plant age is t
$S_p(n,t)$	structure of CA n with a basis metamer of PA p when plant age is t
$u_{pq}(n)$	number of metamers of PA p with lateral buds of PA q, initiated at growth cycle n
$b_{pq}(n)$	number of lateral axes of PA q initiated at growth cycle n on a growth unit of PA p
$Q(t)$	biomass increment of an individual plant at growth cycle t
$D(t)$	total biomass demand of an individual plant at growth cycle t
$E(t)$	environmental factor at growth cycle t
μ	light use efficiency
Sp	characteristic projection area of an individual plant
k	light interception coefficient (Beer-Lambert extinction coefficient)
$S(t)$	total green leaf surface area at growth cycle t
$QB(t)$	total green leaf weight at growth cycle t
slw	specific leaf weight
pg	primary growth
sg	secondary growth
rg	secondary growth of internodes
$p_p^o(j)$	sink value of an organ o of CA j and PA p for biomass
P_p^o	sink strength (amplitude) of an organ o of PA p
P_0^{rg}	parameter of constant demand of biomass for secondary growth
P_1^{rg}	mass sink for secondary growth
P_{sg}^b	parameter of constant demand of biomass for leaf secondary growth
o	blade (b), petiole (s), internode (e), female (f) and male (fm) organs
t_x^o	expansion duration of organs o
ta	leaf functioning duration
$N_p^o(t)$	number of organ o of PA p initiated at growth cycle t
$N^b(t)$	number of leaves alive at growth cycle t
$NL_p^b(t)$	number of living leaves above the metamer of PA p initiated at growth cycle t
$\Delta q_p^o(t,j)$	biomass increment of an organ o of PA p and CA j when plant age is t
$q_p^o(t,j)$	accumulated biomass of an organ o of PA p and CA j when plant age is t
$\Delta q_p^{bsg}(t,j)$	biomass increment for the thickening of the leaf of PA p and CA j when plant age is t
$q_p^{bt}(t,j)$	total biomass for the thickening of the leaf of PA p and CA j when plant age is t
$BS_p(t,j)$	surface area of the leaf of CA j and PA p when plant age is t

Table A.2: Plant parameters (*continued*)

a^o, b^o	coefficients of the sink variation function (Beta function)
γ	coefficient that controls the importance of the ratio Q/D on internode secondary growth
λ	coefficient that controls the proportion of biomass for internode secondary growth allocated downwards
$l_p(t)$	length of the internode of PA p initiated at growth cycle t
E_y	structural Young's modulus

Table A.3: Parameters of population dynamics

T	coming time of pests or auxiliaries
$F0(i,T)$	number of adults of age i in the initial population
$\|F0(T)\|$	initial population size
ai, bi	coefficients of age distribution function of adults in the initial population
ar, br	coefficients of reproduction function
M	fecundity per female adult
s	adults (ad), eggs (eg), juveniles (ju)
$NT_s(i,t)$	number of individuals at stage s of age i at time t
t_{ad}	maximal life span for adults
t_{eg}	hatching duration of eggs
t_{ju}	longevity of juveniles
$N'_{eg}(t)$	total number of eggs laid at time t
$NI_{eg}(t)$	final number of pest eggs plant accepts at time t
$NPE_{eg}(t)$	final number of auxiliary eggs on plant at time t
η	mean number of the eggs laid on each host
ω	coefficient that controls the mean value of the Poisson distribution
$C_s(t)$	viability rate of an individual at stage s at time t
λ	coefficient that controls the variation of the viability rate with respect to the resource
α	leaf surface area eaten by a juvenile during the whole life
β	leaf surface area eaten by an adult during the whole life
$PI_s(k,i,t)$	number of pest individuals at stage s of age i on the leaf initiated at time k when plant age is t
$PE_s(k,j,i,t)$	number of egg auxiliary individuals at stage s of age j on the pest of age i, which is on the leaf initiated at time k when plant age is t
$PW_s(k,j,i,t)$	number of juvenile auxiliary individuals at stage s of age j on the pest of age i, which is on the leaf initiated at time k when plant age is t

Table A.4: Parameters of optimization algorithm

v_{ij}^k	j^{th} coordinate component of the velocity of the i^{th} particle at iteration k
ω^k	inertia weight at iteration k
ω_{start}	initial value of inertia weight
ω_{end}	final value of inertia weight
c_1, c_2, c_3	acceleration coefficients
r_1, r_2, r_3	uniformly distributed random values between 0 and 1
B_{ij}	j^{th} coordinate component of the best position recorded by the i^{th} particle during the previous iterations
B_{gj}	j^{th} coordinate component of the best position of the global best particle in the swarm, which is marked by g
B_{rj}	j^{th} coordinate component of the best position recorded by a random selected particle r during the previous iterations
Bl_{ij}	j^{th} coordinate component of the local guide best position of particle i
x_{ij}^k	j^{th} coordinate component of the current position of particle i at iteration k

Bibliography

T. Alméras, E. Costes, and J. Salles. Identification of biomechanical factors involved in stem shape variability between apricot tree varieties. *Annals of Botany*, 93:455–468, 2004.

P. Ancelin, B. Courbaud, and T. Fourcaud. Development of an individual tree-based mechanical model to predict wind damage within forest stands. *Forest Ecology and Management*, 203:101–121, 2004a.

P. Ancelin, T. Fourcaud, and P. Lac. Modelling the biomechanical behaviour of growing trees at the forest stand scale. Part 1: Development of an incremental transfer matrix method and application to simplified tree structures. *Annals of Forest Science*, 61: 263–275, 2004b.

L. Angelis and G. Stamatellos. Multiple objective optimization of sampling designs for forest inventories using random search algorithms. *Computers and Electronics in Agriculture*, 42:129–148, 2004.

D. Angus and T. Hendtlass. Dynamic ant colony optimisation. *Applied Intelligence*, 23 (1):33–38, 2005.

E. Assmann. *The principles of forest yield study.* Pergamon Press, Oxford, 1970.

P. Baenziger, W. Russell, G. Graef, and B. Campbell. Improving lives: 50 years of crop breeding, genetics, and cytology (c-1). *Crop Sci.*, 46:2230–2244, 2006.

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. *Ann. For. Sci.*, 57:571–585, 2000.

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

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

M. Bertamini, K. Muthuchelian, M. Rubinigg, R. Zore, and N. Nedunchezhian. Low-night temperature (LNT) induced changes of photosynthesis in grapevine (*Vitis Vinifera* L.) plants. *Plant Physiology and Biochemistry*, 43:693–699, 2005.

A. Blum. Drought resistanc, water-use efficiency, and yield potential – are they compatible, dissonant, or mutually exclusive. *Australian Journal of Agricultural Research*, 56:1159–1168, 2005.

L. Borrás and M. Otegui. Maize kernel weight response to postflowering source-sink ratio. *Crop Sci.*, 41:1816–1822, 2001.

L. Borrás, J. A. Cura, and M. Otegui. Maize kernel composition and post-flowering source-sink ratio. *Crop Sci.*, 42:781–790, 2002.

D. Brown, H. Ferris, S. Fu, and R. Plant. Modeling direct positive feedback between predators and prey. *Theoretical Population Biology*, 65:143–152, 2004.

F. Brüchert, G. Becker, and T. Speck. The mechanics of Norway spruce [Picea abies (L.) Karst]: mechanical properties of standing trees from different thinning regimes. *Forest Ecology and Management*, 135:45–62, 2000.

B. Buddadee, W. Wirojanagud, D. Watts, and R. Pitakaso. The development of multi-objective optimization model for excess bagasse utilization: A case study for thailand. *Environmental Impact Assessment Review*, 28:380–391, 2008.

G. Buffoni and G. Gilioli. A lumped parameter model for acarine predator-prey population interactions. *Ecological Modelling*, 170:155–171, 2003.

M. Burger. Iterative regularization of a parameter identification problem occurring in polymer crystallization. In *Inverse Problems*, pages 943–969, 2002.

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

S. Cerasoli, T. M. Wertin, M. A. McGuire, D. Aubrey, and R. O. Teskey. The effects of short term temperature fluctuations on carbon balance and growth in poplar saplings. In *94th ESA ANNUAL MEETING*, Albuquerque, New Mexico, August 2009. Albuquerque Convention Center. http://esameetings.allenpress.com/2009/Paper20996.html.

M. Chelle. Phylloclimate or the climate perceived by individual plant organs: What is it? how to model it? what for? *New Phytologist*, 166:781–790, 2005.

C. Cilas, A. Bar-Hen, C. Montagnon, and C. Godin. Definition of architectural ideotypes for good yield capacity in coffea canephora. *Annals of Botany*, 97:405–411, 2006.

M. Clerc. Discrete particle swarm optimization - illustrated by the traveling salesman problem. 2002. http://www.mauriceclerc.net.

Y. Collette. TSP-2.0. 2009. http://code.google.com/p/scilab-mip/downloads/list.

P. Cournède. Syst'eme dynamique de la croissance des plantes. Technical report, Université de Montpellier II, 2009.

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*, 82:427–438, 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*, 101:1207–1219, 2008.

P. de Reffye. *Modélisation de l'architecture des arbres par des processus stochastiques. Simulation spatiale des modeles tropicaux sous l'effet de la pesanteur. Application au coffea robusta.* PhD thesis, l'université de Paris sud Centre d'Orsay, Paris, France, 1979.

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

D. de Vienne. Les marqueurs moléculaires en génétique et biotechnologies végétales. Technical report, INRA Editions, Paris, 1998.

C. Deleuze and F. Houllier. A transport model for tree ring width. *Silva Fennica*, 31 (3):239–250, 1997.

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.

S. Dencic. Designing a wheat ideotype with increased sink capacity. *Plant Breeding*, 112:311–317, 1994.

M. Denne, C. Atkinson, and R. Dodd. Quantification of Trends in Wood Production within Trees of Silver Birch (Betula pendula Roth.). *Annals of Botany*, 73:655–664, 1994.

D. Dent. *Integrated pest management*. Chapman & Hall, 1995. ISBN 0412573709.

J. Dieleman, L. Marcelis, A. Elings, T. Dueck, and E. Meinen. Energy saving in greenhouse: Optimal use of climate conditions and crop management. *Acta Hort. (ISHS)*, 718:203–210, 2006.

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.

M. Dingkuhn, D. Luquet, A. Clément-Vidal, L. Tambour, H. Kim, and Y. Song. Is plant growth driven by sink regulation? In J. Spiertz, P. Struik, and H. Van Laar, editors, *Scale and complexity in plant systems research: gene-plant-crop relations*, pages 157–170, 2007.

C. Donald. The breeding of crop ideotypes. *Euphytica*, 17:385–403, 1968.

Q. Dong, G. Louarn, Y. Wang, J. Barczi, and P. de Reffye. Does the structure-function model greenlab deal with crop phenotypic plasticity induced by plant spacing? a case study on tomato. *Annals of Botany*, 101:1195–1206, 2008.

M. Dorigo and C. Blum. Ant colony optimization theory: A survey. *Theoretical Computer Science*, 344:243–278, 2005.

J. Drouet and L. Pagès. GRAAL: a model of growth, architecture and carbon allocation during the vegetative phase of the whole maize plant: Model description and parameterisation. *Ecological Modelling*, 165:147–173, 2003.

C. Eschenbach. Emergent properties modelled with the functional structural tree growth model ALMIS: Computer experiments on resource gain and use. *Ecological Modelling*, 186:470–488, 2005.

Y. Eshed and D. Zamir. An introgression line population of lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics*, 141:1147–1162, 1995.

K. Ferentinos, K. Arvanitis, H. Tantau, and N. Sigrimis. Special aspects of it for greenhouse cultivation. In *CIGR Handbook of Agricultural Engineering: Information Technology*, pages 294–312, St. Joseph, Michigan, 2006. American Society of Agricultural and Biological Engineers.

M. Fink. Effects of short-term temperature fluctuation on plant growth and conclusions for short-term temperature optimization in greenhouse. *Acta Hort. (ISHS)*, 328:147–154, 1993.

T. Fourcaud and P. Lac. Numerical modelling of shape regulation and growth stresses in trees I. An incremental static finite element formulation. *Trees-Structure and Function*, 13:23–30, 2003.

T. Fourcaud, F. Blaise, P. Lac, P. Castera, and P. de Reffye. Numerical modelling of shape regulation and growth stresses in trees II. Implementation in the AMAPpara software and simulation of tree growth. *Trees-Structure and Function*, 17:31–39, 2003.

T. Fourcaud, X. Zhang, A. Stokes, H. Lambers, and C. Koerner. Plant growth modelling and applications: The increasing importance of plant architecture in growth models. *Annals of Botany*, 101:1053–1063, 2008.

M. Fournier, A. Stokes, C. Coutand, T. Fourcaud, and B. Moulia. Tree biomechanics and growth strategies in the context of forest functional ecology. In A. Herrel, T. Speck, and N. Rowe, editors, *Ecology and biomechanics – a mechanical approach to the ecology of animals and plants*, pages 1–33. Boca Raton: CRC Press Taylor & Francis, 2006.

S. Francisco and M. Ali. Resource allocation tradeoffs in manila's peri-urban vegetable production systems: An application of multiple objective programming. *Agricultural Systems*, 87:147–168, 2006.

U. Gosselke, H. Triltsch, D. RoBberg, and B. Freier. GETLAUS01-the latest version of a model for simulating aphid population dynamics in dependence on antagonists in wheat. *Ecological Modelling*, 145:143–157, 2001.

H. Guo, V. Letort, L. Hong, T. Fourcaud, P. Cournède, Y. Lu, and P. de Reffye. Adaptation of the GreenLab model for analyzing sink–source relationships in Chinese pine saplings. In T. Fourcaud and X. Zhang, editors, *Plant growth modeling, and applications*, pages 236–243. IEEE Computer Society (Los Alamitos, California), 2008.

Y. Guo, P. de Reffye, Y. Song, Z. Zhan, M. Dingkuhn, and B. Li. Modeling of biomass acquisition and partitioning in the architecture of sunflower. In B. Hu and M. Jaeger, editors, *Plant growth Modeling and Applications (PMA'03)*, pages 271–284, Beijing, China, 2003. Tsinghua University Press and Springer.

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.

B. Habekotté. Options for increasing seed yield of winter oilseed rape (brassica napus l.) : a simulation study. *Field Crops Research*, 54:109–126, 1997.

F. Hallé. *Plaidoyer pour l'arbre*. Acte Sud, 2005. ISBN 2-7427-5712-0.

F. Hallé and R. Oldeman. *Essai sur l'architecture et la dynamique de croissance des arbres tropicaux*. Masson, Paris, 1970.

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

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

J. Hanan and A. Hearn. Linking physiological and architectural models of cotton. *Agricultural Systems*, 75:47–77, 2003.

A. Haverkort and C. Grashoff. Ideotyping-potato a modelling approach to genotype performance. In D. MacKerron and A. Haverkort, editors, *Decision Support Systems in Potato Production - Bringing Models to Practice*, pages 199–211. Wageningen Academic, 2004.

S. He, Q. Wu, J. Wen, J. Saunders, and R. Paton. A particle swarm optimizer with passive congregation. *BioSystems*, 78:135–147, 2004.

M. Herndl, C. Shan, P. Wang, S. Graeff, and W. Claupein. A model based ideotyping approach for wheat under different environmental conditions in north china plain. *Agricultural Sciences in China*, 6:1426–1436, 2007.

M. Ho, B. McCannon, and J. Lynch. Optimization modeling of plant root architecture for water and phosphorus acquisition. *Journal of Theoretical Biology*, 226:331–340, 2004.

E. Holland, J. Burrow, C. Dytham, and J. Aegerter. Modelling with uncertainty: Introducing a probabilistic framework to predict animal population dynamics. *Ecological Modelling*, 220:1203–1217, 2009.

G. E. Host, H. W. Stech, K. E. Lenz, K. Roskoski, and R. Mather. Forest patch modeling: using high performance computing to simulate aboveground interactions among individual trees. *Functional Plant Biology*, 35:976–987, 2008.

C. Houck, J. Joines, and M. Kay. A genetic algorithm for function optimization: A matlab implementation. Technical report, North Carolina State University, 1996.

F. Houllier, J. Leban, and F. Colin. Linking growth modelling to timber quality assessment for Norway spruce. *Forest Ecology and Management*, 74:91–102, 1995.

A. Huston and C. Jeffree. *Handbook of Plant Science*, volume 2, chapter Leaf and Internode, pages 93–98. Wiley, 2007.

K. Jayawickrama. Breeding radiata pine for wood stiffness: review and analysis. *Aust. Forest.*, 64:51–56, 2001.

M. Jones, R. Wells, and D. Guthrie. Cotton response to seasonal patterns of flower removal: I. yield and fiber quality. *Crop Science*, 36:633–638, 1996.

P. Kaitaniemi, J. Hanan, and P. Room. Virtual sorghum: visualisation of partitioning and morphogenesis. *Computers and Electronics in Agriculture*, 28:195–205, 2000.

L. Kamizi, P. Adebola, and A. Afolayan. Effects of temperature, pre-chilling and light on seed germination of *Withania somnifera*; a high value medicinal plant. *South African Journal of Botany*, 72(1):11–14, 2006.

M. Kang. *Stochastic Functional and Structural Plant Modelling based on Substructures*. Phd thesis, Institute of Automation, Chinese Academy of Sciences, China, 2003.

M. Kang, E. Heuvelink, and P. de Reffye. Building virtual chrysanthemum based on sink-source relationships: Preliminary results. *Acta Hort. (ISHS)*, 718:129–136, 2006.

M. Kang, P. Cournède, P. de Reffye, D. Auclair, and B. Hu. Analytical study of a stochastic plant growth model: Application to the GreenLab model. *Mathematics and Computers in Simulation*, 78:57–75, 2008a.

M. Kang, J. Evers, J. Vos, and P. de Reffye. The derivation of sink functions of wheat organs using the greenlab model. *Ann. Bot.*, 101:1099–1108, 2008b.

J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proc. IEEE Conf. on Neural Networks*, volume 4, pages 1942–1948, Piscataway, NJ, 1995.

J. Kennedy and R. Eberhart. *Swarm Intelligence*. Morgan Kaufmann Publishers, 2001.

R. King and N. Sigrimis. Computational intelligence in crop production. *Computers and Electronics in Agriculture*, 31:1–3, 2001.

L. Klèová, M. Havrlentová, and J. Faragó. Cultivar and environmental conditions affect the morphogenic ability of barley (*Hordeum vulgare*) scutellum-derived calli. *Biologia, Bratislava*, 59(4):501–504, 2004.

J. Ko, G. Piccinni, and E. Steglich. Using EPIC model to manage irrigated cotton and maize. *Agricultural Water Management*, 96:1323–1331, 2009.

M. Kropff, P. Teng, and R. Rabbinge. The challenge of linking pest and crop models. *Agricultural Systems*, 49:413–434, 1995.

A. Lacointe. Carbon allocation among tree organs: A review of basic processes and representation in functional-structural tree models. *Ann. For. Sci.*, 57:521–533, 2000.

G. R. Larocque, J. S. Bhatti, R. Boutin, and O. Chertov. Uncertainty analysis in carbon cycle models of forest ecosystems: Research needs and development of a theoretical framework to estimate error propagation. *Ecological Modelling*, 219:400–412, 2009.

J. Lasserre, E. Mason, and M. Watt. The effects of genotype and spacing on Pinus radiata [D. Don] corewood stiffness in an 11 – year old experiment. *Forest Ecology and Management*, 205:375–383, 2005.

P. Lauri and E. Costes. Progress in whole-tree architecture studies for apple cultivar characterization at INRA, France - contribution to the ideotype approach. *Acta Hort. (ISHS)*, 663:357–362, 2004.

R. Lecoustre. *Approche mathématique d'un équilibre biologique à trois antagonistes - exemple du palmier à huile, de* Coelanomenodera minuta *Uh. et de ses parasites d'oeufs*. PhD thesis, Université des Sciences et Techniques du Languedoc, France, 1988.

B. Legaspi, R. Carruthers, and J. Morales-Ramos. Functional response as a component of dynamic simulation models in biological control: the *Catolaccus*–boll weevil system. *Ecological Modelling*, 89:43–57, 1996.

V. Letort. *Adaptation du modèle de croissance GreenLab aux plantes à architecture complexe et analyse multi-échelle des relations source-puits pour l'identification paramètrique*. PhD thesis, Ecole Centrale Paris, Paris, France, 2008.

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

V. Letort, P. Mahe, P. 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*, 101:1243–1254, 2008b.

V. Letort, P. Mahe, P. 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, and Applications*, pages 16–21, IEEE Computer Society (Los Alamitos, California), 2008c.

V. Letort, C. P.H., J. Lecoeur, I. Hummel, P. de Reffye, and A. Chiustophe. 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, and Applications*, pages 65–69. IEEE Computer Society (Los Alamitos, California), 2008d.

D. Li, V. Letort, Y. Guo, P. de Reffye, and Z. Zhan. Modeling branching effects on source-sink relationships of the cotton plant. In B. Li and M. Jaeger, editors, *Plant Growth Modeling, Simulation, Visualization and Applications (PMA'09)*. IEEE Computer Society (Los Alamitos, California), 2009. in press.

Z. Li and B. Hu. Genetic algorithm toolbox based on scientific computing language. *Computer Simulation*, 22(10):186–190, 2005.

A. Lindenmayer. Models for plant tissue development with cell division orientation regulated by preprophase bands of microtubules. *Differentiation*, 26:1–10, 1984.

R. Linker, I. Seginer, and P. Gutman. Optimal CO_2 control in a greenhouse modeled with neural networks. *Computers and Electronics in Agriculture*, 19:289–310, 1998.

G. Lopez, R. R. Favreau, C. Smith, E. Costes, P. Prusinkiewicz, and T. M. DeJong. Integrating simulation of architectural development and source-sink behaviour of peach trees by incorporating markov chains and physiological organ function submodels into L-PEACH. *Functional Plant Biology*, 35(10):761–771, 2008.

B. Loveys, I. Scheurwater, T. Pons, and O. Atkin. Growth temperature influences the underlying components of relative growth rate: an investigation using inherently fast- and slow-growing plant species. *Plant, Cell and Environment*, 25:975–987, 2002.

D. Luquet, M. Dingkuhn, H. Kim, L. Tambour, and A. Clement-Vidal. Ecomeristem, a model of morphogenesis and competition among sinks in rice. 1. concept, validation and sensitivity analysis. *Functional Plant Biology*, 33:309–323, 2006.

Y. Ma, B. Li, Z. Zhan, Y. Guo, D. Luquet, P. de Reffye, and M. Dingkuhn. Parameter stability of the functional-structural 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. 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*, 101:1185–1194, 2008.

A. Mäkelä. Implications of the pipe model theory on dry matter partitioning and height growth in trees. *Journal of Theoretical Biology*, 123:103–120, 1986.

D. Mariau and R. Lecoustre. An explanation for outbreaks of Coelaenomenodera lameensis Berti & Mariau (Coleoptera: Chrysomelidae), a leaf miner of oil palm (Elaeis guineensis Jacq.) in West Africa, based on a study of mortality factors. *International journal of tropical insect science*, 24:159–169, 2004.

D. Mariau and J. Morin. La biologie de *Coelaenomenodera elaeidis* Mlk. IV - La Dynamique des Populations du ravageur et de ses parasites. *Oléagineux*, 27:469–474, 1972.

T. Marsh, R. Huffaker, and G. Long. Optimal control of vector-virus-plant interactions: The case of potato leafroll virus net necrosis. *Amer. J. Agr. Econ.*, 83:556–569, 2000.

A. Mathieu. *Essai sur la modélisation des interactions entre la croissance et le développement d'une plante – Cas du modèle GreenLab.* PhD thesis, Ecole Centrale de Paris, Paris, France, 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 *the 5th International Workshop on Functional-Structural Plant Models*, pages 5–1–5–4, Napier, New Zealand, 2007.

A. Mathieu, P. Cournède, V. Letort, D. Barthélémy, and P. de Reffye. A dynamic model of plant growth with interactions between development and functional mechanisms to study plant structural plasticity related to trophic competition. *Annals of Botany*, 103:1173–1186, 2009.

R. McMurtrie. Forest productivity in relation to carbon partitioning and nutrient cycling: a mathematical model. In M. Cannel and J. Jackson, editors, *Trees as Crop Plants*, pages 194–205, Great Britain, Cumbria, 1985. Institute of Terrestrial Ecology.

J. Mock and P. Pearce. An ideotype of maize. *Euphytica*, 24:613–623, 1975.

B. Monograph. *Growth stages of mono-and dicotyledonous plants.* Fereral Biological Research Centre for Agriculture and Forestry, 2 edition, 2001.

T. Morimoto, T. Takeuchi, and Y. Hashimoto. Growth optimization of plant by means of the hybrid system of genetic algorithm and neural network. In *Neural Networks, IJCNN'93-Nagoya Proceedings of* 1993 *International Joint Conference on*, pages 2979–2982, 1993.

T. Morimoto, T. Torii, and Y. Hashimoto. Optimal control of physiological processes of plants in a green plant factory. *Practice*, 3:505–511, 1995.

T. Morimoto, J. De Baerdemaeker, and Y. Hashimoto. An intelligent approach for optimal control of fruit-storage process using neural networks and genetic algorithms. *Computers and Electronics in Agriculture*, 18:205–224, 1997.

T. Morimoto, K. Tu, K. Hatou, and Y. Hashimoto. Dynamic optimization using neural networks and genetic algorithms for tomato cool storage to minimize water loss. *Trans. ASAE*, 46:1151–1159, 2003.

M. D. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33:161–174, 1991.

S. Mostaghim and J. Teich. Strategies for finding good local guides in multi-objective particle swarm optimization (MOPSO). In *IEEE Swarm Intelligence Symposium*, pages 26–33, 2003.

B. Moulia, C. Coutand, and C. Lenne. Posture control and skeletal mechanical acclimation in terrestrial plants: implications for mechanical modeling of plant architecture. *American Journal of Botany*, 93:1477–1489, 2006.

G. Nepveu. The possible status of wood quality in oak breeding programs (Quercus Petreae Liebl and Quercus robur L). *Annals of Forest Science*, 50 Suppl 1:338–394, 1993.

T. Nguyen-Huu, C. Lett, J. Poggiale, and P. Auger. Effect of movement frequency on global host-parasitoid spatial dynamics with unstable local dynamics. *Ecological Modelling*, 197:290–295, 2006.

K. Niklas. Mechanical properties of Black locust (Robinia pseudoacacia L.) wood. Size – and age–dependent variations in sap–and heartwood. *Annals of Botany*, 79:265–272, 1997a.

K. Niklas. Size– and age–dependent variation in the properties of sap– and heartwood in Black locust (Robinia pseudoacacia L.). *Annals of Botany*, 79:473–478, 1997b.

R. W. Pearcy, H. Muraoka, and F. Valladares. Crown architecture in sun and shade environments: assessing function and trade-offs with a three-dimensiional simulation model. *New Phytologist*, 166:791–800, 2005.

S. Peng, G. Khush, P. Virk, Q. Tang, and Y. Zou. Progress in ideotype breeding to increase rice yield potential. *Field Crops Research*, 108:32–38, 2008.

J. A. Pennington and R. A. Fisher. Classification of fruits and vegetables. *Journal of Food Composition and Analysis*, 22(1):s23–s31, 2009.

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. Peters, N. E. Verhoest, R. Samson, M. Van Meirvenne, L. Cockx, and B. De Baets. Uncertainty propagation in vegetation distribution models based on ensemble classifiers. *Ecological Modelling*, 220:791–804, 2009.

H. Pinnschmidt, W. D. Batchelo, and P. S. Teng. Simulation of multiple species pest damage in rice using CERES-rice. *Agricultural Systems*, 48:193–222, 1995.

J. Piqueira, S. de Mattos, and J. Vasconcelos-Neto. Measuring complexity in three-trophic level systems. *Ecological Modelling*, 220:266–271, 2009.

R. Plant, M. Mangel, and L. Flynn. Multiseasonal management of an agricultural pest II: The economic optimization problem. *Journal of Environmental Economics and Management*, 12:45–61, 1985.

E. Poupaert and Y. Deville. Simulated Annealing with estimated temperature. *AI Communications*, 13:19–26, 2000.

G. Prasanna Kumar, B. Srivastava, and D. Nagesh. Modeling and optimization of parameters of flow rate of paddy rice grains through the horizontal rotating cylindrical drum of drum seeder. *Computers and Electronics in Agriculture*, 65(1):26–35, 2009.

R. Pratt. Specialty corns. *Crop Sci.*, 41:1990, 2001.

P. W. Price, C. E. Bouton, P. Gross, B. A. McPheron, J. N. Thompson, and A. E. Weis. Interactions among three trophic levels: Influence of plants on interactions between insect herbivores and natural enemies. *Ann. Rev. Ecol. Syst.*, 11:41–65, 1980.

P. Prusinkiewicz, M. Hammel, J. Hanan, and R. Mech. *Handbook of formal languages, vol. III: Beyond Words*, chapter Visual models of plant development, pages 535–597. 1996.

K. Raju and D. Kumar. Multicriterion decision making in irrigation planning. *Agricultural Systems*, 62:117–129, 1999.

D. Rasmusson. An evaluation of ideotype breeding. *Crop Sci.*, 27:1140–1146, 1987.

P. Reeves and G. Coupland. Response of plant development to environment: control of flowering by daylength and temperature. *Current Opinion in Plant Biology*, 3:37–42, 2000.

M. Renton, P. Kaitaniemi, and J. Hanan. Functional-structural plant modelling using a combination of architectural analysis, L-systems and a canonical model of function. *Ecological Modelling*, 184:277–298, 2005.

G. L. Ritchie and C. W. Bednarz. *Cotton Growth and Development*, chapter Cotton Growth and Development, pages 3–14. Cooperative Extension, 2007.

K. Sastry, G. David, and G. Kendall. *Search Methodologies Introductory Tutorials in Optimization and Decision Support Techniques*, chapter Genetic Algorithms, pages 97–125. Springer US, 2005.

A. Sharov. Modelling insect dynamics. In E. Korpilahti, H. Mukkela, and T. Salonen, editors, *Caring for the forest: research in a changing world*, volume II, pages 293–303, Jyvaskyla, Finland, August 1996. Gummerus Printing.

A. Sharov. *Quantitative Population Ecology - on-line lectures*. http://home.comcast.net/s̃harov/PopEcol/.

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

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

J. A. Snyman. *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*. Springer Berlin, 2005. ISBN 0-387-24348-8.

M. Song and G. Gu. Research on particle swarm optimization: A review. In *Third International Conference on Machine Learning and Cybernetics*, pages 2236–2241, Shanghai, 2004.

H. Spatz and F. Bruechert. Basic biomechanics of self-supporting plants : wind loads and gravitational loads on a Norway spruce tree. *Forest Ecology and Management*, 135:33–44, 2000.

S. Tanksley and R. Susan. Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science*, 277(5329):1063–1066, August 1997.

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

H. Tonnang, L. Nedorezov, and H. Ochanda. Assessing the impact of biological control of *Plutella xylostella* through the application of Lotka-Volterra model. *Ecological Modelling*, 220:60–70, 2009.

A. Trewavas. Aspects of plant intelligence. *Annals of Botany*, 92:1–20, 2003.

P. Trichilo and L. Wilson. An ecosystem analysis of spider mite outbreaks: physiological stimulation or natrual enemy suppression. *Experimental & Applied Acrology*, 17:291–314, 1993.

P. Tripathi, S. Bandyopadhyay, and S. Pal. Multi-objective particle swarm optimization with time variant inertia and acceleration coefficients. *Information Sciences*, 177: 5033–5049, 2007.

TSPLIB95. http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/.

M. Uribelarrea, J. Carcova, M. Otegui, and M. Westgate. Pollen production, pollination dynamics, and kernel set in maize. *Crop Sci.*, 42:1910–1918, 2002.

F. van Evert, P. de Visser, and M. Heinen. Operational optimization of organic fertilizer application in greenhouse crops. *Acta Hort. (ISHS)*, 718:165–172, 2006.

E. van Henten, F. Buwalda, H. de Zwart, A. de Gelder, J. Hemming, and J. Bontsema. Toward an optimal control strategy for sweet pepper cultivation - 2. optimization of the yield pattern and energy efficiency. *Acta Hort. (ISHS)*, 718:391–398, 2006.

G. van Straten, H. Challa, and F. Buwalda. Towards user accepted optimal control of greenhouse climate. *Computers and Electronics in Agriculture*, 26:221–238, 2000.

C. Vega, V. Sadras, F. Andrade, and S. Uhart. Reproductive allometry in soybean, maize and sunflower. *Annals of Botany*, 85:461–468, 2000.

A. Walter and U. Schurr. Dynamics of leaf and root growth: endogenous control versus environmental impact. *Annals of Botany*, 95:891–900, 2005.

F. Wang, M. Kang, H. Han, Q. Lu, V. Letort, Y. Guo, and P. de Reffye. Calibration of topological development in the procedure of parametric identification: application of the stochastic GreenLab model for *Pinus sylvestris* var. *Mongolica*. In B. Li and M. Jaeger, editors, *Plant Growth Modeling, Simulation, Visualization and Applications (PMA'09)*. IEEE Computer Society (Los Alamitos, California), 2009. in press.

X. Wang, F. Divos, C. Pilon, B. Brashaw, R. Ross, and R. Pellerin. Assessment of decay in standing timber using stress wave timing nondestructive evaluation tools: A guide for use and interpretation. Technical report, http://www.fpl.fs.fed.us/documnts/fplgtr/fpl_gtr147.pdf, 2004.

S. O. Web. Scilab the open source platform for numerical computation. http://www.scilab.org/.

J. Weiner. The influence of competition on plant reproduction. In J. Lovett Doust and L. Lovett Doust, editors, *Plant reproductive ecology: patterns and strategies*, pages 228–245, New York, 1988. Oxford University Press.

M. J. West-Eberhard. *Developmental Plasticity and Evolution*. OXFORD University press, New York, 2003. ISBN 0-19-512235-6.

M. Westgate, J. Lizaso, and W. Batchelor. Quantitative relationships between pollen shed density and grain yield in maize. *Crop Sci.*, 43:934–942, 2003.

M. Wiemann and G. Williamson. Radial gradients in the specific gravity of wood in some tropical and temperate trees. *Forest Science*, 35:197–210, 1989a.

M. Wiemann and G. Williamson. Wood specific gravity gradients in tropical dry and montane rain forest trees. *American Journal of Botany*, 76:924–928, 1989b.

D. Woodcock and A. Shier. Does canopy position affect wood specific gravity in temperate forest trees? *Annals of Botany*, 91:529–537, 2003.

L. Wu. *Méthodes Variationelles pour des Modèles Fonction-Structure de Plantes : Identification de Paramètre, Contrôle et Assimilation de Données*. Phd thesis, Université Joseph Fourier, Grenoble, France, 2005.

L. Wu, P. de Reffye, F. Le Dimet, and B. Hu. Optimization of source-sink relationships based on a plant functional-structural model: a case study on maize. In B. Hu and M. Jaeger, editors, *Plant Growth Modeling and Applications (PMA'03)*, pages 285–295, Beijing, China, 2003. Tsinghua University Press and Springer.

L. Wu, P. de Reffye, B. Hu, F. Le Dimet, and P. Cournède. A water supply optimization problem for plant growth based on GreenLab model. *ARIMA*, pages 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.

X. Yin, P. Stam, M. Kropff, and A. Schapendonk. Crop modeling, QTL mapping, and their complementary role in plant breeding. *Agron. J.*, 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*, 9:426–432, 2004.

D. Zamir. Improving plant breeding with exotic genetic libraries. *Natural Reviews Genetics*, 2:983–989, 2001.

Z. Zhan, P. de Reffye, F. Houllier, and B. Hu. Fitting a functional- structural growth model with plant architectural data. In B. Hu and M. Jaeger, editors, *Plant Growth Modeling and Applications (PMA'03)*, pages 236–249, Beijing, China, 2003. Tsinghua University Press and Springer.

Z. Zhan, H. Rey, D. Li, Y. Guo, P. Cournède, and P. de Reffye. Study on the effects of defoliation on the growth of cotton plant using the functional structural model greenlab. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, and Applications*, pages 194–201. IEEE Computer Society (Los Alamitos, California), 2008.

G. Zhang, H. Jiang, G. Niu, X. Liu, and S. Peng. Simulating the dynamics of carbon and nitrogen in litter-removed pine forest. *Ecological Modelling*, 195:363–376, 2006.

X. Zhao, P. de Reffye, D. Barthélémy, and B. Hu. Interactive simulation of plant architecture based on a Dual-Scale Automaton model. In B. Hu and M. Jaeger, editors, *Plant Growth Modeling and Applications (PMA'03)*, pages 144–153, Beijing, China, 2003. Tsinghua University Press and Springer.

B. Zheng. *Evaluation of 3-D Structure and Function and Simulation of Growth for the Canopies of Super High-yielding Rice*. Phd thesis, China Agricultural University, Beijing, China, November 2009.

Y. Zhou and H.-B. Shao. The responding relationship between plants and environment is the essential principle for agricultural sustainable development on the globe. *C.R.Biologies*, 331:321–328, 2008.

Acknowledgements

Great thanks to

- Philippe de Reffye, for his supervision during more than three years; for his courage to my work all the time, which makes me confident to my work and makes my thesis be finished successfully; for his concern when I met problems and was in blue; for his invitation to his home several times, especially for the Christmas day and New year, which was an unforgettable experience during the period when I was in France; for his generous. Especially, he taught me by personal example a lot the morals of being a man.

- Baogang Hu, for his supervision during more than five years, especially for his support when I was in France; for his concern about my research work and my life when I came back to China. His rigorous scholarship and modest character deeply affected me and I benefited a lot from it.

- Paul-Henry Cournède, for his supervision in France; for his valuable suggestions for my manuscripts and his corrections; for the discussion proposed by him about the plan after my defence two years ago. From that day on, I began to think seriously about what the research is, how to do research, and the career plan after the defence.

- Mengzhen Kang. Through the discussion with her, I could understand problems much deeply, and could generate new ideas. Thanks for her concern and understanding when I was in blue.

- Thierry Fourcaud, for his very useful suggestions, and his corrections for the paper that is published in the journal "Silva Fennica".

- Véronique Letort, for her much careful corrections and valuable suggestions of the manuscript; for her great help for me to handle everything essential for living in France when I arrived in France the first time; for her consolation to me When I was in blue. She let me enter the life of the French. I learned a lot from you about the theory of GreenLab from agricultural and botanical points of view. I learned a lot from you about writing papers, correcting papers as a co-author.

- Yuntao Ma, for her data provided to me for the work about the optimization on maize in the thesis, and for her explanations to my questions.

- Vincent Le Chevalier, for helping me handle every programming problem.

- Colleagues in Ecole Centrale Paris: Cédric Loi, Aurélie Cotillard, Thomas Guyard, Xiujuan Wang, Qiongli Wu, Zhongping Li and Guanghui Wang, for the invitation to your homes for leisures or sportive activities, which make the life in France colorful.

- Feng Wang, Dong Li, Sébastian Lemaire, who went to a tea plant nursery in Nanjing and help me analyze the topological structure and growth behavior of tea plant.

- Liu Jia, colleague in CASIA, for her help handling all the issues when I was in France.

- Ling Di, Lei Li and Juan Cao, professors of the department of graduate students in CASIA, and Jean-Hubert Schmitt and Géraldine Carbonel, director and secretory in Ecole Centrale Paris, for their great support and comprehension to my PhD work.

- my father, my mother, my sister and Yufei Han, for their great support for my work and my choice.

Publications

1. **Rui Qi**, Yuntao Ma, Baogang Hu, Philippe de Reffye, and Paul-Henry Cournède. Optimization of source-sink dynamics in plant growth for ideotype breeding: a case study on maize. *Computer and Electronics in Agriculture*, 71:96-105, 2010.

2. **Rui Qi**, Véronique Letort, Mengzhen Kang, Paul-Henry Cournède, Philippe de Reffye, and Thierry Fourcaud. Application of the GreenLab Model to Simulate and Optimize Wood Production and Tree Stability: A Theoretical Study. *Silva Fennica*, 43(3):465-487, 2009.

3. **Rui Qi**, Paul-Henry Cournède, René Lecoustre, and Philippe de Reffye. Tri-trophic Ecosystem Oil palm-Pests-Auxiliaries : I. Modeling and Simulation. In B.G. Li and M. Jaeger, editors, *Plant Growth Modeling, Simulation, Visualization and Applications (PMA'09)*, Beijing, China, 2009. IEEE Computer Society (Los Alamitos, California). in press.

4. **Rui Qi**, Paul-Henry Cournède, and Philippe de Reffye. Tri-trophic Ecosystem Oil palm-Pests-Auxiliaries : II. Sensitivity Analysis, Parameter Identification and Control. In B.G. Li and M. Jaeger, editors, *Plant Growth Modeling, Simulation, Visualization and Applications (PMA'09)*, Beijing, China, 2009. IEEE Computer Society (Los Alamitos, California). in press.

5. **Rui Qi**, Baogang Hu, and Paul-Henry Cournède. PSOTS: A Particle Swarm Optimization Toolbox in Scilab. In *International Workshop on Open-source Software for Scientific Computation*, pages 107-114, Guiyang, China, 2009. IEEE press.

6. **Rui Qi**, Yuntao Ma, Baogang Hu, Philippe de Reffye, and Paul-Henry Cournède. New Approach for the Study of Source-Sink Dynamics on Maize. In *International Symposium on Crop Modeling and Decision Support*, pages 161-168, Nanjing, China, 2008.

7. MengZhen Kang, **Rui Qi**, Philippe de Reffye, and Baogang Hu. GreenScilab:A toolbox simulating virtual plants in the Scilab environment. In Marwan Al-Akaidi editor, *8th International Middle Eastern Simulation Multiconference*, pages 174-178, Alexandria, Egypt, 2006. European Technology Institute, EUROSIS.

8. MengZhen Kang, **Rui Qi**, Philippe de Reffye, and Baogang Hu. Cultivating Virtual Plant in Scilab. In *SCILAB Research, Development and Applications*, pages 167-181, Springer.

9. Mengzhen Kang, Xianwen Wang, **Rui Qi**, and Philippe de Reffye. GreenScilab_Crop, an open source toolbox for plant simulation and parameter estimation. In *International Workshop on Open-source Software for Scientific Computation*, pages 91-95, Guiyang, China, 2009. IEEE press.

10. Mengzhen Kang, Paul-Henry Cournède, Amélie Mathieu, Véronique Letort, **Rui Qi**, and Zhigang Zhan. A Functional-Structural Plant Model: Theory and Applications in Agronomy. In *International Symposium on Crop Modeling and Decision Support*, pages 148-160, Nanjing, China, 2008.