

PROJET ENJEU : SANTÉ ET BIOTECHNOLOGIES

RAPPORT FINAL

Développement d'outils mathématiques pour l'agriculture de précision

Client et Référent Pédagogique

Pierre CARMIER
Paul-Henry COURNÈDE

P2018 : groupe SBT-11

John DE WASSEIGE
Nayef DERWICHE
Alexis MATHEY
Daina ZHENG

10 juin 2015

Table des matières

Résumé	2
1 Introduction	3
2 Synthèse de l'étude documentaire	4
3 Méthodologies	5
3.1 Outils de travail collaboratif	5
3.2 Gestion de la bibliographie	5
3.3 Organisation et partage des tâches	6
3.4 Planning	6
3.5 Difficultés rencontrées	7
4 Code et modèle	9
4.1 Description du modèle	9
4.2 Description du fonctionnement du code	12
5 Résultats	13
5.1 Résultats	13
5.2 Pour aller plus loin	17
Conclusion	19
Remerciements	20
Bibliographie	21
A Physiologie des plantes	22
A.1 Généralités	22
A.2 Photosynthèse	22
B Histoire de la modélisation de la croissance des plantes	24
B.1 Débuts de la botanique et de l'agronomie	24
B.2 Les premiers modèles	25
C Mise en place des modèles de croissance des plantes	27
C.1 Caractéristiques propres aux modèles de croissance de plantes	27
C.2 Bonnes pratiques en modélisation	27
D Quelques modèles de la croissance des plantes	29
E Tableaux des variables du modèle informatique	35
F Planning	37
G Article modèle LNAS wheat	37

Résumé

Ce rapport présente nos résultats et nos perspectives du projet et s'articule autour de cinq parties.

Dans un premier temps, nous introduisons le contexte global du projet en mettant l'accent sur l'objectif poursuivi par celui-ci.

On expose ensuite une synthèse de l'*étude documentaire* en décrivant brièvement chacun des thèmes abordés dans celle-ci. On rappellera notamment quelques notions de physiologie des plantes, un bref historique de la modélisation des plantes ainsi que les principaux modèles actuels et leurs caractéristiques communes.

La partie suivante décrit l'ensemble des méthodes utilisées pour mener à bien notre projet. On décrira d'abord les outils de travail que nous avons utilisés afin de collaborer efficacement. On rappellera ensuite la façon dont nous avons géré notre bibliographie. Un planning de la répartition des tâches à travers le semestre sera ensuite donné. On terminera cette partie en décrivant de façon exhaustive les difficultés que nous avons rencontrées jusqu'à maintenant.

Avant de décrire les résultats obtenus, nous rappellerons de manière synthétique le fonctionnement de notre modèle. On expliquera ensuite comment nous avons implémenté celui-ci en JULIA en détaillant les variables et fonctions utilisées.

La dernière partie reprendra d'une part nos résultats ainsi qu'une analyse de ceux-ci et de leur pertinence par rapport à ce que l'on observe expérimentalement, et d'autre part les perspectives à long terme que nous proposons pour le projet.

1 Introduction

Au cours de ce projet, nous nous sommes intéressés au *développement d'outils mathématiques pour l'agriculture de précision* en collaboration avec le laboratoire MAS (Mathématiques et Informatique pour la Complexité et les Systèmes).

Dans le contexte environnemental actuel, et alors que 2 milliards de personnes sont en sous-nutrition dans le monde, la modélisation de la croissance des plantes prend tout son sens.

En effet, le problème des ressources *énergétiques* et *alimentaires* est un sujet crucial du 21ème siècle. Il faudra être capable de nourrir plus de 9 milliards d'humains en 2050 [14]. De plus, les ressources fossiles et l'eau douce viennent à manquer dans de nombreuses régions agricoles (comme en Californie). Cependant, l'agriculture nécessite un apport d'eau important et consomme beaucoup d'énergie. L'agriculture est ainsi responsable de l'émission de près de 20% des gaz à effet de serre tous les ans [11].

Les progrès de l'informatique et des mathématiques permettent néanmoins d'espérer une amélioration nette des rendements. L'objectif final est clair : nourrir plus de monde, en minimisant l'utilisation de pesticides, d'engrais et d'eau, et donc limiter l'impact environnemental et le réchauffement climatique. En effet, les capteurs, les satellites, les drones... permettent de collecter des données de plus en plus précises. Dans un futur proche, des modèles de plus en plus efficaces permettront sans doute de tirer parti de ces données pour améliorer considérablement les rendements.

L'objectif premier de notre projet était d'implémenter un modèle, le modèle LNAS appliqué au blé, sur la plateforme PyGMAIion-Julia du laboratoire Digi-plant. Cette implémentation constituait pour ainsi dire le livrable qui était attendu par notre client Pierre Carmier. Une fois implémenté, nous avons pu utiliser ce modèle pour simuler les caractéristiques d'un champ de blé. Par exemple, le modèle permet d'obtenir l'évolution de la biomasse des grains de blé (quantité d'intérêt) dans le champ.

Notre client Pierre Carmier, ainsi que notre référent pédagogique Paul-Henry Cournède sont des chercheurs au laboratoire MAS. Ce laboratoire travaille notamment sur les modèles mathématiques de croissance des plantes, en collaboration avec la startup Digi-plant.

À terme, nous devrions également réaliser une analyse de sensibilité qui permettrait de quantifier l'influence des paramètres qui interviennent dans le modèle. Nous avons aussi en projet de modifier quelques aspects du modèle, en rajoutant des phénomènes comme la diffusion de l'eau dans le sol.

2 Synthèse de l'étude documentaire

Cette partie se présente comme une synthèse de l'analyse bibliographique que nous avons réalisée en décembre. Le lecteur intéressé trouvera une description plus détaillée des paragraphes suivants dans les annexes [A](#), [B](#), [C](#) et [D](#).

Nous nous étions d'abord intéressés à la *physiologie* de la plante. Cela nous a permis d'appréhender les éléments essentiels à son développement. L'élément crucial qui permet le développement d'une plante est la *photosynthèse*. En effet, la plante possède cette capacité extraordinaire de synthétiser de la matière organique à partir de matière minérale. Parce qu'ils interviennent dans la réaction de photosynthèse, l'eau, le soleil... sont des éléments clés de la modélisation du développement d'une plante.

Grâce à l'*histoire* de l'étude des plantes, nous avons également pu comprendre le contexte dans lequel la modélisation des plantes s'inscrit aujourd'hui. C'est l'essor des mathématiques et de l'informatique qui permet d'envisager demain une modélisation efficace, avec des applications comme l'agriculture de précision. En effet, les mathématiques ont peu à peu formalisé la connaissance des plantes pour obtenir des modélisations plus précises. On rappellera que la production de biomasse Q est proportionnelle

1. au LAI, pour Leaf Area Index, qui correspond au ratio entre la surface totale supérieure des feuilles vertes et la surface de sol sur laquelle se développe la culture,
2. au PAR, pour Photosynthetically Active Radiation, qui correspond à l'énergie utile de la lumière incidente,
3. à I , la lumière interceptée, donnée par

$$I = 1 - e^{-k \cdot \text{LAI}}$$

4. au LUE, pour Light Use Efficiency, qui est le facteur d'efficacité énergétique. selon

$$Q = \text{LUE} \times \text{PAR} \times I$$

On rappellera également le concept de *temps thermique* qui correspond à l'accumulation de températures dépassant un certain seuil :

$$\tau^{(n+1)} = \tau^{(n)} + \max[0, \underline{T^{(n)}} - T_c],$$

où $\underline{T^{(n)}}$ est l'écart de température constaté et T_c est le seuil de variation de température qu'on impose.

Nous nous étions également intéressés aux *caractéristiques* propres à la modélisation des plantes, afin de comprendre les principales difficultés qui empêchent une utilisation globale des modèles de croissance des plantes. Parmi eux, on peut citer : l'estimation difficile des paramètres d'un modèle, des techniques sophistiquées en informatique et en mathématique sont nécessaires et une diversité importante sans benchmarking des modèles ne permet pas de choisir facilement celui qui sera le plus adapté.

3 Méthodologies

Nous présenterons dans cette section l’organisation générale du groupe et les outils qui nous aidé tout au long de ce projet. On abordera d’abord les outils que nous avons utilisés pour partager notre travail le plus efficacement possible. On présentera ensuite notre manière de gérer la bibliographie ainsi que les logiciels utilisés pour ne pas perdre les sources visitées. On détaillera ensuite l’organisation interne du groupe, et plus précisément le système mis en place pour optimiser la communication. On décrira finalement les problèmes et les difficultés auxquels nous avons été confrontés tout au long de ce projet.

3.1 Outils de travail collaboratif

Pour des raisons pratiques et esthétiques, nous avons décidé d’écrire nos rapports en \LaTeX . Il s’agissait donc de trouver la meilleure façon de partager le code source et de pouvoir contrôler les changements apportés au document. Une première idée pourrait être d’utiliser ShareLaTeX qui propose une plate-forme de compilation en ligne ainsi qu’un système de gestion de versions assez simple à utiliser. Nous n’avons pas choisi cette solution notamment pour les raisons suivantes. L’utilisateur doit être connecté dès qu’il veut travailler sur le projet, le système de compilation est assez lent et l’utilisateur n’est pas libre d’utiliser son éditeur de texte ou son visualisateur de PDF favori.

Pour palier aux problèmes décrits ci-dessus, le logiciel `git` associé à GitHub est une très bonne alternative. Il permet en effet à chaque membre du groupe de travailler sans être connecté ainsi que d’utiliser son éditeur et compilateur favori. Chaque membre travaille donc de son côté en faisant des *commits* et lorsqu’il juge que son travail est utile pour les autres, il *push* sur le serveur. L’algorithme de fusion, *merge*, permet également de fusionner intelligemment les lignes d’un fichier qui ont été modifiées par plusieurs membres. Le dernier point à souligner est que `git` permet une gestion des branches, particulièrement pratique lorsqu’on veut développer une partie du projet sans risquer de créer des erreurs dans le programme principal.

Nous combinons donc ces deux outils pour

1. implémenter le modèle LNAS blé dans la plateforme Pygmalion en `JULIA` pour le client dont le code source est sur la plateforme GitLab,
2. rédiger l’étude documentaire en partageant le code \LaTeX à l’aide d’un dossier sur GitHub¹.

3.2 Gestion de la bibliographie

Pour la gestion de la bibliographie au sein du document, nous utilisons le package `biblatex`. Celui-ci permet d’écrire l’ensemble de nos références dans un fichier `.bib` sous la forme suivante.

1. <https://github.com/jdewasseige/projet-sbt11>

```
@online{histoire_mod_plantes,
  title = {Une histoire de la modélisation des plantes},
  author = {Philippe de Reffye and Marc Jaeger
    and Paul-Henry Cournède},
  url = {https://interstices.info/jcms/c_38032/une-histoire-de-
    la-modelisation-des-plantes},
  year = {2009},
  month = "04",
}
```

La mise en page est alors automatique en fonction des informations fournies et le rendu de l'exemple est présenté ci-dessous.

- [3] Philippe de REFFYE, Marc JAEGER et Paul-Henry COURNÈDE. *Une histoire de la modélisation des plantes*. Avr. 2009. URL : https://interstices.info/jcms/c_38032/une-histoire-de-la-modelisation-des-plantes.

Cela paraît à priori assez lourd d'écrire soi-même toutes les informations en suivant cette syntaxe mais il existe des logiciels comme Zotero qui font le travail à notre place. Les sources trouvées sur Google Scholar peuvent également être exportées aisément au format BibTex.

3.3 Organisation et partage des tâches

Afin de communiquer et de prévoir notre travail, nous avons essentiellement utilisé Slack², qui est un logiciel de plus en plus utilisé pour les travaux de groupe ainsi que dans les start-ups. Il permet d'éviter de devoir alterner entre plusieurs applications comme les mails, DropBox et Twitter, puisqu'il permet d'être connecté à celles-ci au sein de l'application. On peut également créer plusieurs *channels* pour séparer la communication entre les différentes tâches. Par exemple dans ce projet nous avons les *channels* suivantes : **general**, **etude-documentaire**, **planning** et **dev_plate-forme**. On trouve aussi un système d'historique et de gestion de fichiers efficace.

3.4 Planning

L'organisation de nos travaux au cours de ce semestre est présentée en annexe F dans le tableau 14 (page 37). Il contient ce qui devait être fait en théorie et à quel moment, certaines semaines voyaient leurs tâches réalisées entièrement tandis que d'autre non. Nous présentons donc pour la dernière partie de cette section les différentes complications rencontrées.

2. <https://slack.com/>

3.5 Difficultés rencontrées

Au cours de ce projet, nous avons été confrontés à de nombreuses difficultés.

3.5.1 Difficultés de gestion du temps

Le temps fut une difficulté récurrente. Nous n'étions sans doute pas prêts à gérer de nous-mêmes notre temps. C'est ainsi que nous avons réalisé l'opportunité que constitue ce type de projet. Jusqu'à maintenant, nous avons été confrontés la plupart du temps à des questions précises dans des examens, dont l'emploi du temps nous est imposé à l'avance. Ici, nous devons de nous même organiser notre temps. Nous pourrions nous dédouaner en soulignant l'emploi du temps chargé de nos études à Centrale. Mais ce projet est là pour nous rappeler qu'une organisation préalable et une répartition efficace des tâches doivent permettre d'éviter de subir le temps. Et cela s'apprend grâce à des projets comme celui-ci.

3.5.2 Difficultés informatiques

Nous avons également été confrontés à des problèmes informatiques lors de l'implémentation et de l'utilisation du modèle.

- Un bogue a empêché au début de nos premières tentatives le lancement de nos programmes. En effet, les chercheurs du laboratoire MAS qui ont développé la plateforme utilisaient les fonctions directement sur leur répertoire, alors que nous y accédions depuis le dossier principal. Il y avait un problème de chemin relatif. Il a donc fallu détecter l'origine de ce bogue avant de pouvoir commencer à utiliser nos programmes.
- De plus, des erreurs classiques sur notre implémentation empêchaient également le bon déroulement de nos utilisations du modèle : coquilles, oublis, fautes de frappe... L'impossibilité d'utiliser la simulation à cause du bogue précédemment décrit a rendu leur détection plus ardue.

3.5.3 Difficultés théoriques

Enfin, le modèle directement implémenté à l'aide du document qui contenait les formules théoriques du modèle, et que nous avons implémenté, conduisait à des valeurs qui n'étaient pas cohérentes. Les connaissances de Pierre Carmier sur les modèles de croissance des plantes nous ont permis de savoir ce qui était cohérent dans nos résultats et ce qui ne l'était pas. Nous avons donc avec l'aide de notre client ajusté les formules que nous utilisons :

- Dans un premier temps, la croissance de la plante était trop faible. Il a fallu changer les fonctions log-normales dans les fonctions d'allocation, de remobilisation et de sénescence. La formule donnée dans le document utilisait une définition différente de l'écart type que la formule qui correspondait aux paramètres que Pierre Carmier nous avait fournis. Nous avons donc utilisé di-

rectement la formule déjà implémentée dans le modèle LNAS betterave qui correspond bien aux conventions utilisées.

- Une fois ceci corrigé, nous obtenions cette fois des récoltes beaucoup trop élevées, plus de 3000 g m^{-2} en grains et environ 16 de LAI, alors que les valeurs attendues sont respectivement en ordre de grandeur 1000 g m^{-2} et entre 5 et 6. Cela était dû à l’absence de prise en compte dans le modèle théorique du temps de montaison, temps à partir duquel la tige commence à croître de manière très rapide.

En effet, nous procédions comme suit pour l’allocation de la biomasse produite par photosynthèse aux différents organes. On calcule d’abord le coefficient d’allocation des grains. Il est nul au départ et tend rapidement vers 1 lorsque la plante arrive à maturité. La biomasse restante était ensuite allouée à parts égales entre les 3 g m^{-2} compartiments restants (tiges, racines, feuilles) (sauf en cas de stress thermique ou hydrique). Ceci est peu réaliste. Nous avons donc introduit un temps de montaison, tout d’abord de manière assez abrupte. Il s’agit de ne rien allouer à la tige au départ, puis de lui en allouer de plus en plus à partir du temps de montaison. Nous nous sommes contentés en première approximation de décrire deux régimes, un régime pré-montaison et un régime post-montaison, avec une transition instantanée. Plus tard, nous comptons implémenter une transition progressive avec une loi lognormale et rentrer ces nouveaux paramètres dans le modèle. Cela a permis d’avoir des valeurs plus réalistes pour le LAI et la quantité de grain est descendue à 2500 g m^{-2} .

- Mais nous obtenions toujours une trop grande quantité de grain (2500 g m^{-2}). Cela était dû au fait qu’après la maturité de la plante, toute la biomasse de la tige était réallouée aux grains, ce qui n’est pas réaliste non plus. En effet, une partie des tiges subit une sénescence trop tôt avant de pouvoir se transformer en grains de blé. On crée donc un compartiment “tige jaune”, suivant le modèle qu’on a feuille jaune/feuille verte. La biomasse de ce compartiment ne sera jamais allouée aux grains. Après ajout de ce compartiment au modèle, nous avons obtenu des valeurs de grain finales légèrement en dessous de 2000 g m^{-2} , ce qui était encore trop. Nous regarderons après la soutenance si le modèle peut encore être affiné, mais surtout nous essaierons de jouer sur les paramètres du modèle pour obtenir des valeurs encore plus cohérentes.

4 Code et modèle

4.1 Description du modèle

Le modèle LNAS blé est une version plus complexe du modèle LNAS betterave. Le modèle est constitué de deux compartiments principaux : le compartiment circulation de biomasse et le compartiment simulation de l'eau. Dans cette section nous ferons une explication des principes généraux et une description qualitative du modèle. Le lecteur intéressé trouvera tous les détails en annexe G dans la référence que nous avons nous-même utilisé pour implémenter le modèle.

4.1.1 Circulation de biomasse

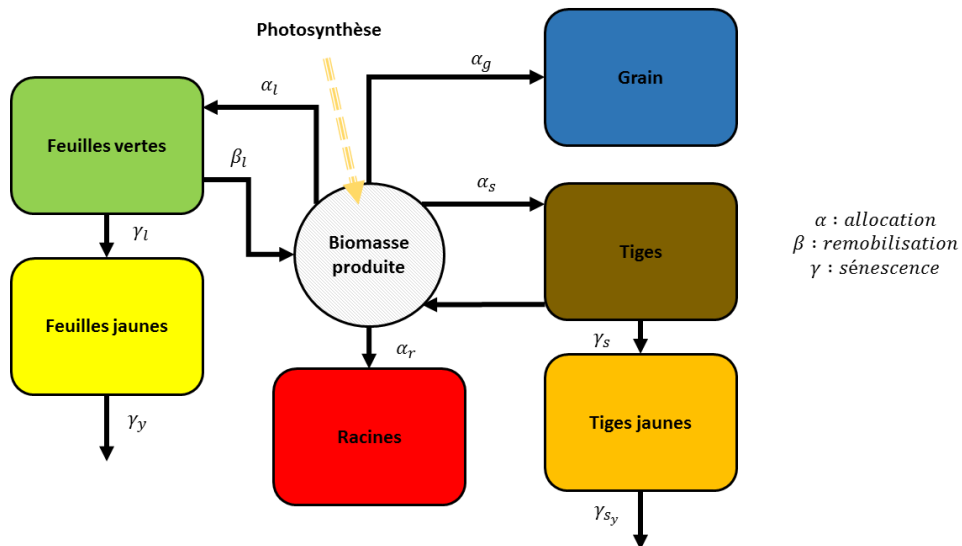


FIGURE 1 – Schéma de la circulation de biomasse.

Le modèle va à chaque itération calculer la quantité de biomasse produite, l'ajouter à un pool de biomasse. Ensuite il va calculer des coefficients d'allocation, de remobilisation et de sénescence qui vont déterminer la circulation de biomasse entre les différents compartiments, la loi de conservation voulant que toute la biomasse soit allouée au final. Les fonctions d'allocations déterminent la part de biomasse distribuée à chaque compartiment. Les fonctions de sénescence représentent un processus de vieillissement. Les fonctions de remobilisation représentent la part de biomasse des compartiments retournée au pool de biomasse pour être à nouveau allouée.

Les étapes sont les suivantes :

- La biomasse est produite par photosynthèse et la quantité créée est calculée grâce à la loi de Beer-Lambert. Un coefficient de stress hydrique et un coefficient de stress thermique peuvent réduire la production.

- Des coefficients d'allocation sont ensuite calculés selon des lois log-normales, avec pour paramètres un temps caractéristique, à partir duquel l'allocation commence à être significative, une espérance et une variance. Typiquement, l'allocation aux racines et aux feuilles domine au départ, puis c'est celle aux tiges et vers la maturité de la plante c'est celle au grain.
- Des coefficients de sénescence sont calculés selon des lois log-normales de la même façon. La sénescence des feuilles commence vers la fin de la croissance de la plante.
- Enfin, des coefficients de remobilisation sont calculés toujours selon des lois log-normales. La remobilisation intervient de façon intense vers la fin pour les feuilles et les tiges en faveur du grain.

La valeur critique qui donne le rythme d'évolution de la plante n'est pas le temps réel mais le temps thermique qui correspond à l'accumulation de températures dépassant un certain seuil :

$$\tau^{(n+1)} = \tau^{(n)} + \max[0, \underline{T^{(n)}} - T_c],$$

L'idée part d'un constat simple : la croissance de la plante est ralentie lorsque la température est basse.

4.1.2 Simulation de l'eau

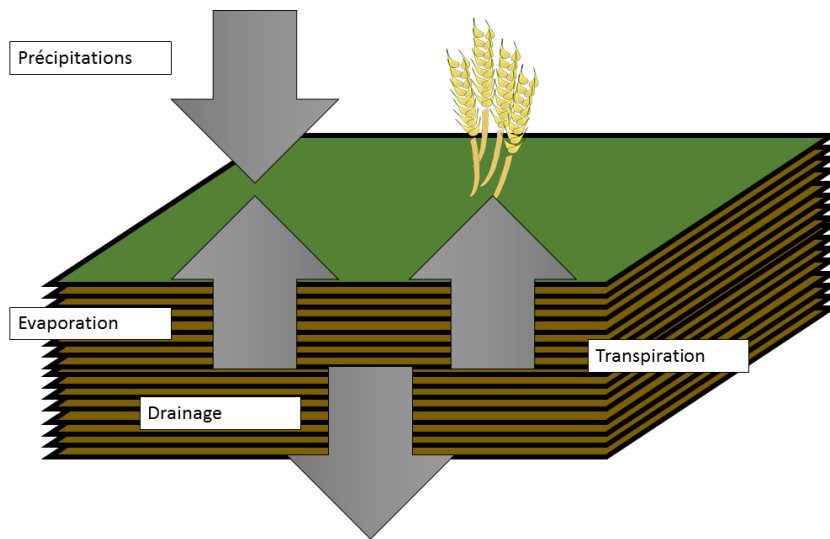


FIGURE 2 – Schéma de simulation de l'eau.

La simulation de l'eau est assez précisément décrite dans le modèle et la figure 2 permet de comprendre intuitivement ce qu'il se passe.

Elle repose sur l'équation d'équilibre de l'eau

$$R^{(n+1)} = R^{(n)} + W^{(n)} - E_s^{(n)} - T_p^{(n)} - d^{(n)}$$

Dans cette équation, $W^{(n)}$ représente les précipitations quotidiennes, $Es^{(n)}$ est la quantité d'eau perdue par évaporation, $Tp^{(n)}$ est la quantité d'eau transpirée par la plante et $d^{(n)}$ est une fonction de drainage qui évacue l'eau ne pouvant être absorbé si le sol est saturé.

On calcule l'évaporation et la transpiration potentielles, qui sont la quantité d'eau pouvant être évaporée et transpirée en fonction de l'environnement (évapotranspiration du milieu, ensoleillement...). Puis l'évaporation et la transpiration "max" qui sont celles requises étant donnée la plante et l'environnement (profondeur des racines, ensoleillement...). On regarde donc la transpiration et l'évaporation qui aura effectivement lieu, qui sera le minimum entre le potentiel et ce qui est demandé. Si la transpiration requise pour la plante n'est pas satisfaite, il y aura un stress hydrique.

Une fois l'évaporation et la transpiration déterminée on rajoute au sol l'eau apportée par les précipitations puis on enlève l'eau évaporée puis transpirée. Les apports et pertes d'eau se font en FIFO (First In, First Out) ou "piles" : l'eau est d'abord rajoutée sur les couches supérieures jusqu'à saturation à l'humidité maximale (humidité au delà de laquelle le sol ne peut plus absorber d'eau) avant d'être rajoutée aux couches les plus basses, de même, on retire l'eau des couches supérieures d'abord jusqu'à atteindre l'humidité minimale (humidité en deçà laquelle on ne peut plus tirer d'eau), avant de retirer l'eau des couches les plus basses. Tout excès éventuel d'eau après opérations est drainé.

4.2 Description du fonctionnement du code

Dans un premier temps, nous avons construit un modèle informatique à partir du modèle mathématique fourni [3].

Il s'agissait d'abord d'affecter à chaque variable du modèle un représentant dans notre programme, en rendant ces derniers assez explicites afin d'être plus efficace pour l'implémentation des fonctions. Les tableaux 1, 2 et 3, dans l'Annexe E, contiennent respectivement les variables d'état, temporaires, auxiliaires et principales, les variables *environnementales* et les *paramètres* de la tige, du sol et des racines.

La deuxième partie du travail consistait à définir les différentes fonctions (appelées “modules” selon les conventions Pygmalion) qui mettent à jour les variables d'état et paramètres du temps n au temps $n + 1$. L'implémentation de chacune d'entre elles est basée sur l'algorithme générique 1. On retrouve à l'entrée les variables d'états **xn** et **xnplus1**, le temps **n**, les variables de contrôles **u**, les paramètres **p**. On va ensuite mettre à jour une composante de **xnplus1** par rapport à **xn** en fonction de **n**, **u** et **p** en suivant la description du modèle.

Algorithm 1 Algorithme générique qui sert de base pour l'implémentation des fonctions. La fonction f n'est pas définie mais sert de placeholder pour représenter les opérations nécessaires à la mise à jour de **xnplus1**.

```

1: procedure GENERIC(int  $n$ , State Vector  $xn$ , Control Vector  $u$ , Parameters Vector  $p$ , State Vector  $xnplus1$ )
2:   begin
3:      $xnplus1.composante \leftarrow f(p.composante, u.composante, xn.composante)$ 
4:   end
5:   return None
6: end procedure

```

À titre d'exemple, on présente la fonction `get_pot_evaporation` qui met à jour l'évaporation requise en fonction des conditions environnementales selon l'équation 1.

$$Espot^{(n)} = K_s ET0 e^{-\lambda LAI^{(n)}} \quad (1)$$

On obtient en JULIA l'implémentation suivante

```

1 function get_pot_evaporation!( $n$ ,  $xn$ ,  $u$ ,  $p$ ,  $xnplus1$ )
2    $xnplus1.soil\_req\_evaporation =$ 
3      $p.k\_s * u.ET0[n] * \exp(-p.lambda * xn.leaf\_area\_index)$ 
4 end

```

Il reste finalement deux fonctions particulières qui ne suivent pas l'algorithme 1.

1. `initialize` qui va initialiser les composantes de **xn**,
2. `transition` qui permet de passer au temps suivant et d'exécuter toutes les fonctions.

5 Résultats

5.1 Résultats

Après les divers corrections au modèles citées dans la partie méthodologie, nous avons lancé une simulation. Nous n'avons pas encore eu les données expérimentales et donc nous n'avons pas encore fait l'estimation des paramètres. Nous n'avons pas non plus fait d'étude de sensibilité des paramètres. Les résultats exploités et présentés dans ce rapport sont donc des sorties d'une simple simulation en première approche, bien que déjà satisfaisante. Les sorties sont toutes qualitativement très satisfaisantes et pour beaucoup proches des résultats quantitatifs voulus.

5.1.1 Cadre de la simulation

La simulation se fait sur 275 jours, avec un pas de discrétisation d'un jour. Les paramètres utilisés nous ont été donnés par notre client, Pierre Carmier, ils correspondent à des paramètres réalistes et cohérents avec la littérature scientifique pour ce type de modèle et de plante. Les données environnementales correspondent à des données réelles qui nous ont aussi été fournies par le client.

5.1.2 Résultats considérés

Nous considérerons les sorties suivantes :

- **La masse du grain (fig. 3)** : C'est la donnée la plus importante et la plus intéressante à mesurer qui correspond à la récolte de céréale. On attend une valeur autour de 1000.
- **Le LAI (fig. 4)** : C'est l'indice de surface foliaire, décrit précédemment, il est directement lié à la masse du feuillage vert. On le préfère à ce dernier car les valeurs qu'il atteint classiquement sont connues : on attend une valeur entre 5 et 6.
- **La masse des tiges (fig. 5)** : Son profil est intéressant à suivre car il est lié à deux moments critiques : le temps de montaison (temps à partir duquel la biomasse est allouée aux tiges en priorité) et un second temps de remobilisation de la biomasse de la tige (en même temps que celle des feuilles vertes) qui se transforme en grain.
- **La quantité d'eau dans le sol (fig. 6)** : Une part importante du modèle est la simulation du comportement de l'eau dans le sol. Il est donc intéressant de voir de façon globale comment la quantité d'eau sur la couche du sol considérée évolue.
- **L'indice de stress hydrique (fig. 7)** : Il permet de voir à quel moment la plante manque d'eau et les influences que cela a sur la croissance de la plante.
- **L'indice de stress thermique (fig. 8)** : De même il permet de voir l'influence de la température sur la croissance de la plante.

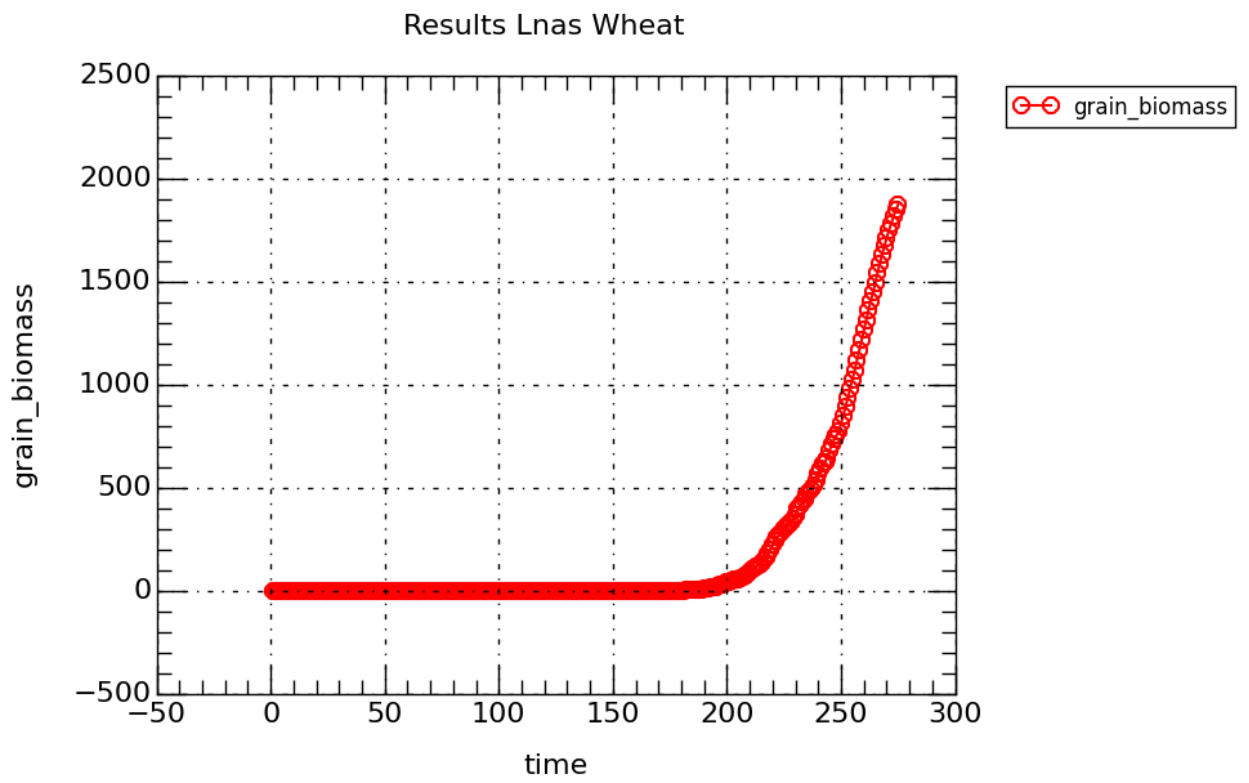


FIGURE 3 – Biomasse de grain en fonction du temps.

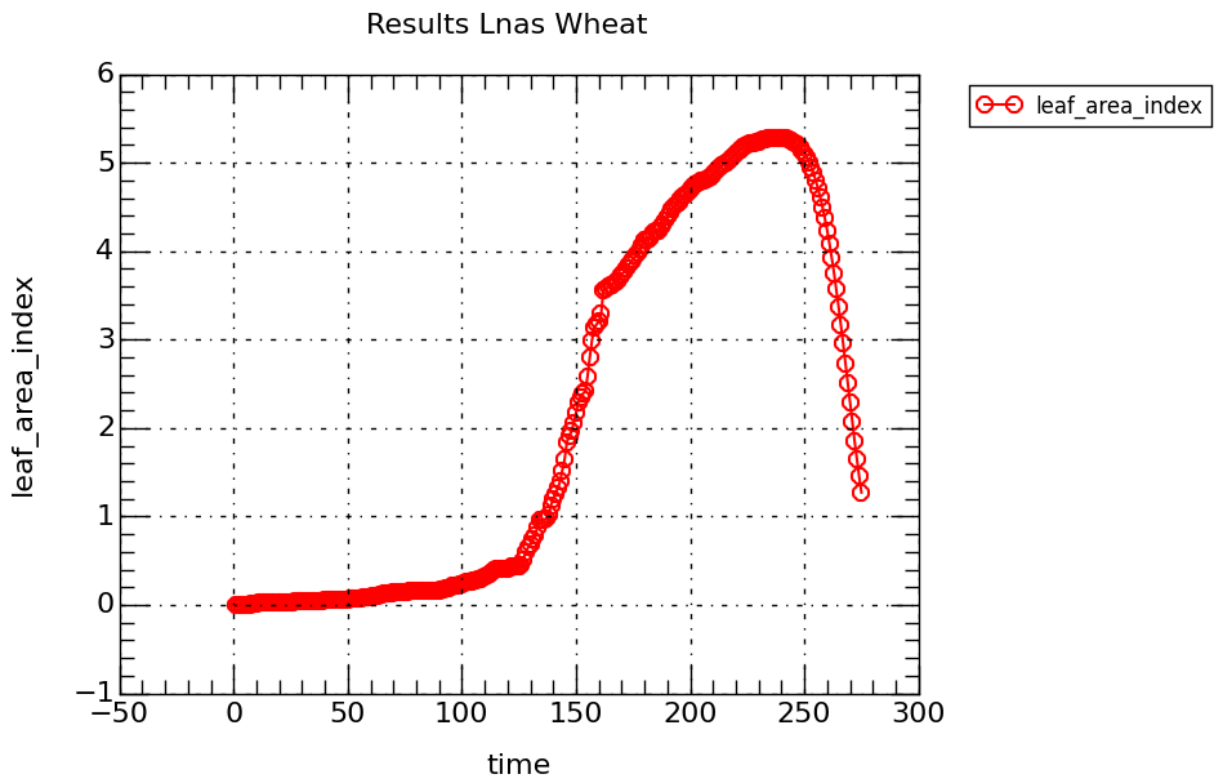


FIGURE 4 – Indice LAI en fonction du temps.

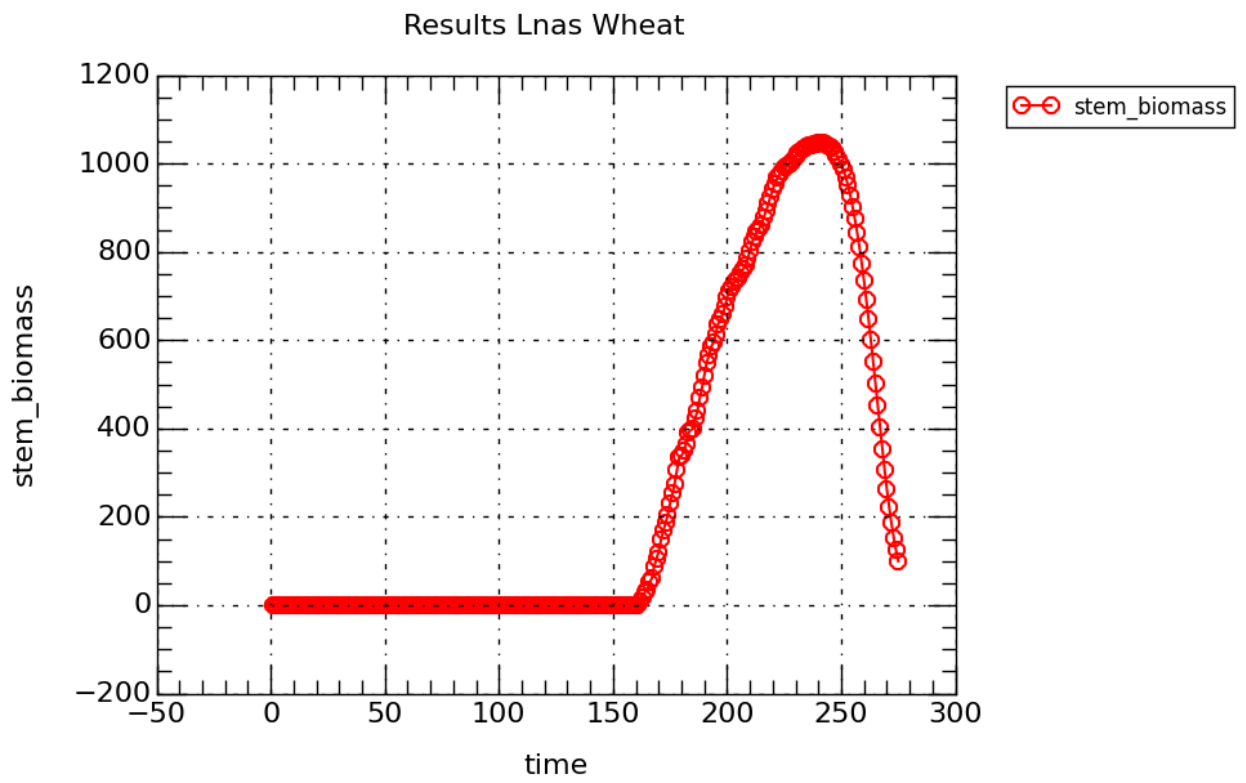


FIGURE 5 – Biomasse de la tige en fonction du temps.

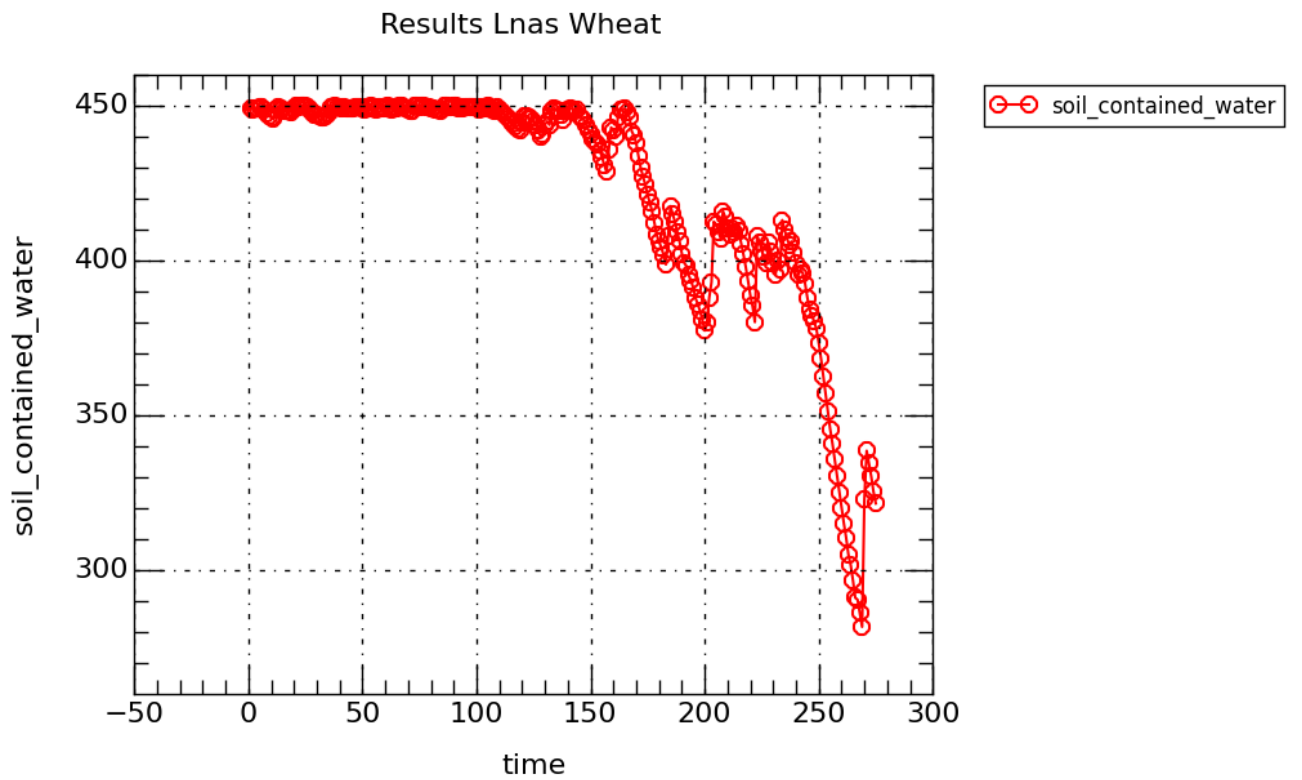


FIGURE 6 – Quantité d'eau dans la couche de sol considérée.

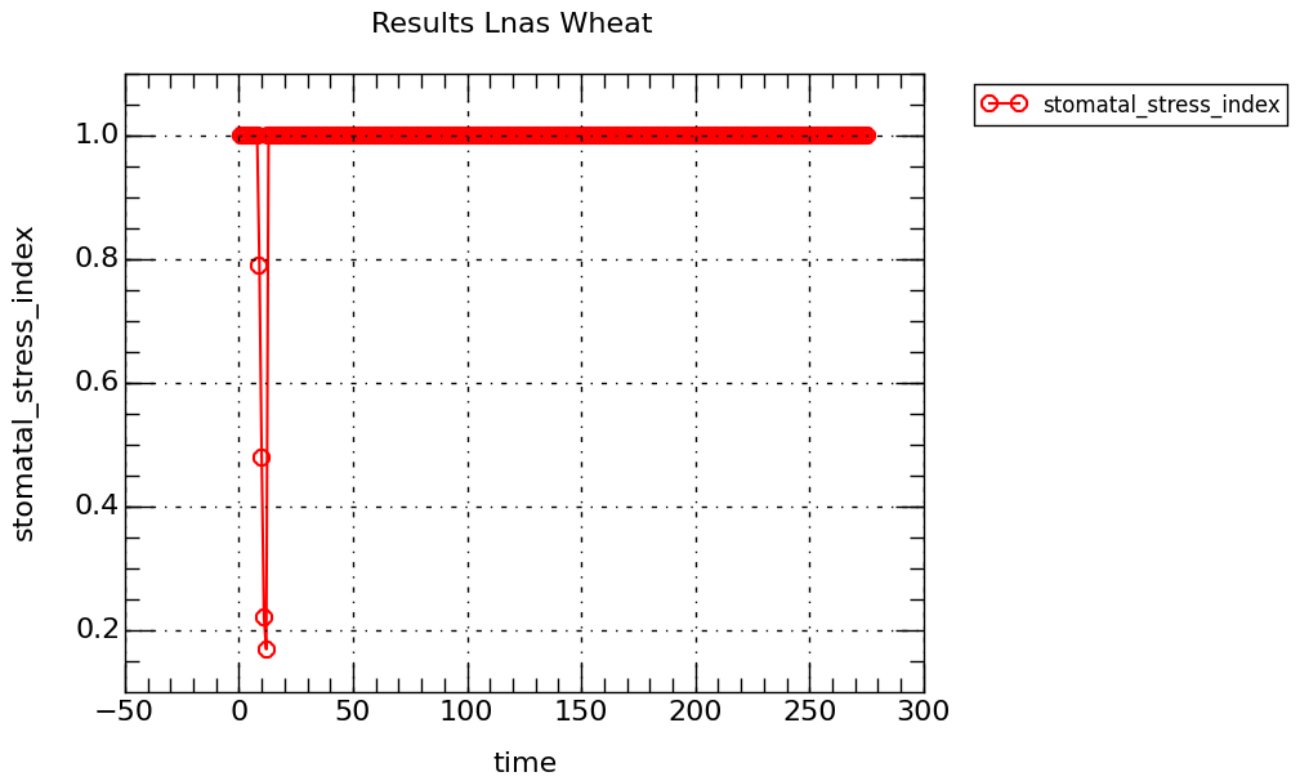


FIGURE 7 – Indice de stress hydrique en fonction du temps.

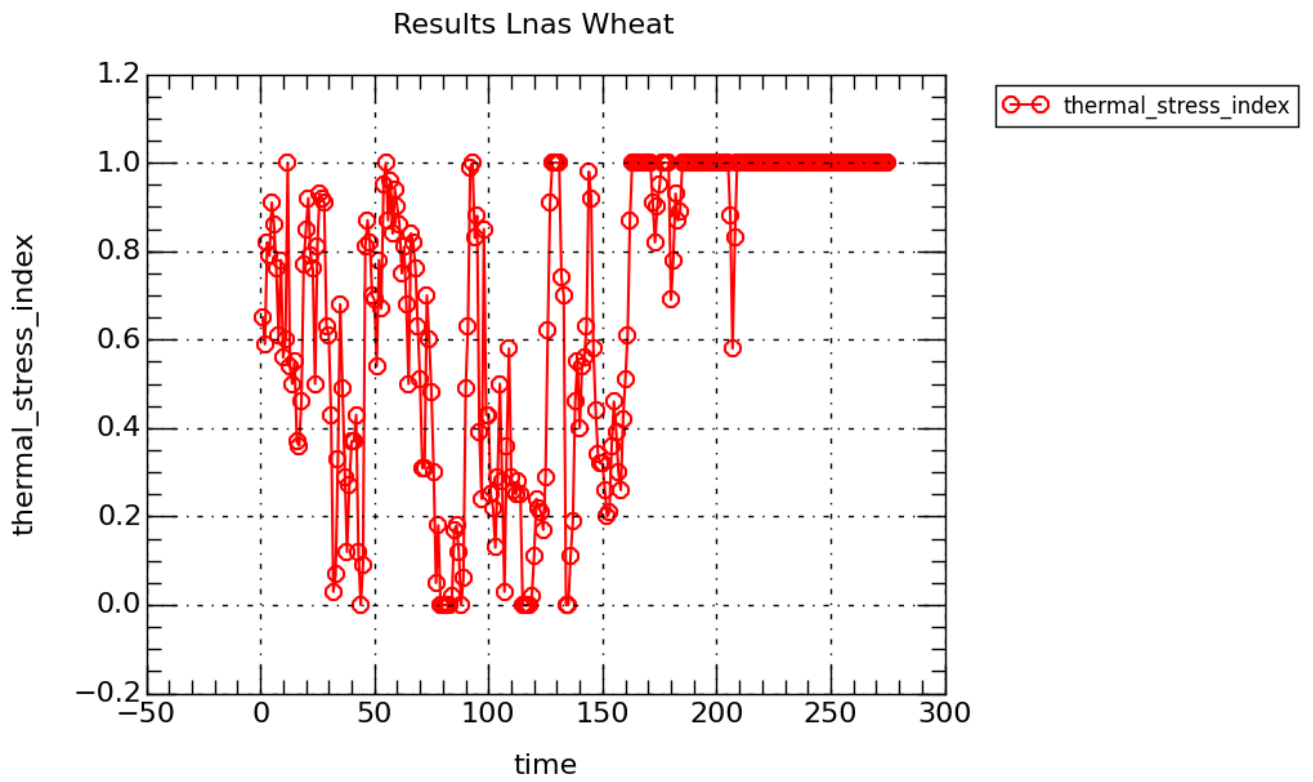


FIGURE 8 – Indice de stress thermique en fonction du temps.

5.1.3 Résultats et commentaires

Pour ces différentes sorties, nous avons eu les résultats suivants :

- **La masse du grain** : Nous avons un comportement qualitatif totalement cohérent : une croissance rapide et tardive de la quantité de grain. Cependant cette croissance dure trop longtemps et la quantité de biomasse du grain au moment de la récolte atteint des valeurs autour de 2000 au lieu de 1000. Les causes plausibles sont expliquées dans la partie méthodologie.
- **Le LAI** : Le résultat est à la fois très cohérent qualitativement et quantitativement. Le profil de la courbe correspond à celui qu'on trouve expérimentalement : croissance très rapide, croissance rapide, pic puis décroissance rapide. Et les valeurs maximum obtenues sont bien entre 5 et 6.
- **La masse des tiges** : Le comportement est cohérent et correspond à celui attendu : une croissance rapide à partir du temps de montaison puis une décroissance rapide avec la remobilisation et la sénescence. Il nous faudrait comparer avec des données expérimentales pour le quantitatif.
- **La quantité d'eau dans le sol** : Les réserves d'eau dans le sol sont larges au début et ne sont mises à mal par la plante que vers la fin alors qu'elle est déjà à maturité. Les paramètres environnementaux ne rendent pas la question de l'eau cruciale dans cette simulation : elle est abondante.
- **L'indice de stress hydrique** : La conclusion précédente est confirmée. La plante ne ressent à aucun moment de stress hydrique, à part un court épisode au départ dû à la longueur insuffisante de la racine mais rapidement corrigée. Pour mieux apprécier et évaluer la partie simulation de l'eau, on pourrait essayer le modèle avec des données environnementales qui présenteraient des périodes de sécheresse qui rendraient la question de l'eau plus cruciale.
- **L'indice de stress thermique** : Plus prévisible et direct que le stress hydrique, il est directement lié à la température. La simulation commence en hiver, le stress peut être ressenti avec des températures inférieures à 10, les épisodes de stress thermique sont donc surtout marqués en début de simulation, où ils peuvent ralentir la croissance de la plante, mais leur occurrence est normale en hiver et pas suffisamment sévère dans nos données pour empêcher la croissance normale de la plante.

Ainsi nous avons obtenu des résultats très satisfaisants qualitativement qui montrent que le modèle est fonctionnel et cohérent. Les comportements de sortie sont ceux attendus. La quantité de grain obtenue en sortie reste trop importante pour être réaliste et nous devons améliorer le modèle sur ce point. Les autres résultats quantitatifs sont satisfaisants.

5.2 Pour aller plus loin . . .

Notre objectif premier, qui était d'implémenter le modèle LNAS appliqué au blé sur la plateforme Pygmalion-Julia, étant atteint, il s'agirait d'affiner le modèle, en l'étudiant et en l'utilisant sur un jeu de données expérimentales pour comparer

les simulations du modèle aux données réelles. Nous avons bien avancé et suivi la méthodologie de bonne pratique en modélisation (voir annexe C).

Nous avons terminé la première étape d’analyse et étudié les comportements généraux du modèle. Une étape intéressante à réaliser pourrait être une analyse de sensibilité, afin de déterminer l’influence de chaque paramètre sur le modèle. En effet, plus de 20 paramètres sont présents dans le modèle, et les déterminer tous nécessite de nombreux jeux de données. Si l’influence de certains paramètres s’avère assez faible, on pourra les remplacer par des constantes prises à leur valeur moyenne pour simplifier le modèle et rendre les estimations plus faciles.

Ensuite il faudrait identifier le modèle. C’est-à-dire en faire une estimation non-paramétrique (structure du modèle, choix de modélisations...) et paramétrique (valeur des paramètres) en le confrontant à un jeu de données expérimentale. L’estimation des paramètres du modèle se décrit comme suit : on rentre un fichier de données expérimentales et on utilise une méthode d’estimation des paramètres (par exemple la méthode des moindres carrés généralisée, qui est déjà implémentée sur la plateforme). Cela permet d’obtenir une estimation des paramètres. Des modifications qualitatives au modèle peuvent être faites à la main ou alors mises sous formes de paramètres booléens (par exemple, “use advanced water simulation = TRUE or FALSE”).

Enfin il s’agirait d’évaluer le modèle. Une fois les paramètres estimés, on peut ensuite utiliser le modèle pour réaliser des simulations et les confronter à des données expérimentales qui n’ont pas servi à déterminer les paramètres pour voir si le comportement qualitatif et quantitatif est satisfaisant : évolutions cohérentes, goodness of fit des données, incertitudes. Si cette étape s’avère concluante, on peut alors tester les capacités prédictives du modèle et en voir ainsi des applications pratiques.

La modélisation n’est pas linéaire et il s’agit de faire des allers-retours entre ces étapes. Par exemple, pour affiner ce modèle, on pourrait décrire de manière plus précise le comportement de l’eau dans le sol. Ainsi, nous pourrions introduire dans le modèle la capillarité de l’eau, qui a tendance à la faire remonter des couches les plus humides aux couches les plus sèches, et décrire plus précisément le drainage de l’eau vers le bas dû tout simplement à la pesanteur. Il s’agira de comparer le comportement du modèle avec et sans ce raffinement, afin de constater ou non son utilité, s’il donne des résultats plus réalistes ou pas. On peut également l’introduire en paramètre booléen comme suggéré plus haut.

Conclusion

Alors que l'intelligence artificielle vient de battre le meilleur joueur du monde de go, plus rien ne semble hors de portée de l'informatique et des mathématiques, reste à appliquer cette révolution à un secteur vieux comme le monde : l'agriculture.

Pour la modélisation de la croissance des plantes, des modèles comme le modèle LNAS permettent d'envisager une simulation efficace et précise des champs de culture et à terme d'augmenter considérablement les rendements en suivant les recommandations de tels modèles. En effet, l'ordinateur pourra tester des milliers de possibilités (culture de blé, d'orge sur tel parcelle, avec plus d'eau sur le nord de la parcelle que sur le reste par exemple) et fournir celle qui donne les meilleurs rendements. Les enjeux sont cruciaux. Les ressources viennent à manquer, alors que l'humanité n'a jamais eu autant besoin de cultiver pour se nourrir, pour produire des bio-carburants... Une agriculture intelligente, portée par la modélisation de la croissance des plantes apparaît donc comme une bonne alternative à l'agriculture intensive.

Si nous devons conseiller une équipe qui prendrait notre suite, nous lui dirions d'aller plus vite sur l'implémentation du modèle, pour pouvoir commencer le plus tôt possible la partie la plus intéressante, à savoir l'utilisation du modèle, sa manipulation sur des données réelles et les modifications par essai/erreur. Ce modèle peut être affiné et amélioré, et testé sur de nombreuses données, ou alors d'autres modèles peuvent être implémentés. En tout cas les possibilités ne manquent pas et les pistes intéressantes sont nombreuses.

Remerciements

Nous remercions Mme Le Chevalier et Mme Lopes, qui encadrent le projet de façon motivante et efficace. Leurs remarques et retours sur nos présentations ainsi que sur l'avancement de notre projet nous ont été très utiles pour prendre le recul nécessaire à la rédaction de ce rapport. Merci également à nos responsables d'ateliers Mme Lévy et M. Bertrand, qui nous ont permis de mieux appréhender l'aspect humain du projet, que ce soit pour présenter à l'oral ou communiquer au sein du groupe. Enfin, nous tenons tout particulièrement à remercier notre client Pierre Carmier, pour sa pédagogie, ses conseils avisés, sa disponibilité et tout le temps précieux qu'il nous a accordé.

Références

- [1] VisioDrones SA. Service d'imagerie par système de drones. URL : <http://www.visiodrones.com/produits/vegedrones/>.
- [2] Charlotte BAEY. « Modelling inter-individual variability in plant growth models and model selection for prediction ». Theses. Ecole Centrale Paris, fév. 2014. URL : <https://tel.archives-ouvertes.fr/tel-00985747>.
- [3] Pierre CARMIER. « LNAS model for wheat with soil coupling ». In : (nov. 2015).
- [4] Paul-Henry COURNÈDE et al. « Development and Evaluation of Plant Growth Models : Methodology and Implementation in the PYGMALION platform ». In : *Mathematical Modelling of Natural Phenomena* 8.4 (juin 2013), p. 112–130. URL : <https://hal-ecp.archives-ouvertes.fr/hal-00860902>.
- [5] Martine GUÉRIF et Dominique KING. *Agriculture de précision*. Éditions Quæ, 2007.
- [6] William G HOPKINS. *Physiologie végétale*. De Boeck Supérieur, 2003.
- [7] A JULLIEN et P HUET. *Agriculture de précision*. 2005. URL : <https://tice.agroparistech.fr/coursenligne/courses/SIAFEEAGRONOMIE9cea/document/machinisme/themes/documents/synthese.pdf>.
- [8] *Les émissions agricoles de particules dans l'air*. Agence de l'Environnement et de la Maîtrise de l'Energie. URL : http://www.chambres-agriculture.fr/fileadmin/user_upload/thematiques/Produire_durablement/Energies_biomasse/Emissions_agricoles_particules.pdf.
- [9] Przemyslaw PRUSINKIEWICZ et Aristid LINDENMAYER. *The algorithmic beauty of plants*. Springer Science & Business Media, 2012.
- [10] *Quatre gigantesques plans d'eau menacés de disparition*. 2014. URL : http://www.francetvinfo.fr/monde/environnement/en-images-quatre-gigantesques-plans-d-eau-menaces-de-disparition_708841.html.
- [11] *Quelle contribution de l'agriculture française à la réduction des émissions de gaz à effet de serre ?* 2013. URL : <http://institut.inra.fr/Missions/Eclairer-les-decisions/Etudes/Toutes-les-actualites/Etude-Reduction-des-GES-en-agriculture>.
- [12] Philippe de REFFYE, Marc JAEGER et Paul-Henry COURNÈDE. *Une histoire de la modélisation des plantes*. Avr. 2009. URL : https://interstices.info/jcms/c_38032/une-histoire-de-la-modelisation-des-plantes.
- [13] WIKIPEDIA. *Agriculture de précision* — *Wikipedia, The Free Encyclopedia*. URL : https://fr.wikipedia.org/wiki/Agriculture_de_pr%C3%A9cision.
- [14] WIKIPEDIA. *Population mondiale* — *Wikipedia, The Free Encyclopedia*. URL : https://fr.wikipedia.org/wiki/Population_mondiale.
- [15] WIKIPEDIA. *Stomate* — *Wikipedia, The Free Encyclopedia*. URL : <https://fr.wikipedia.org/wiki/Stomate>.
- [16] Philippe ZWAENEPOEL et Jean-Michel LE BARS. « L'agriculture de précision ». In : *Ingénieries-EAT* 12 (1997), p-67.

Annexes

A Physiologie des plantes

A.1 Généralités

Tout d'abord, présentons succinctement comment se développe et fonctionne une plante. La science qui décrit l'ensemble des mécanismes qui participent à l'édification d'un organisme vivant est appelée morphogénèse

Chez les plantes, la morphogénèse commence avec la germination de la graine et s'arrête à la mort de la plante. [2, p. 22]. Celle-ci diffère d'une plante à l'autre, et dépend également de l'environnement de la plante.

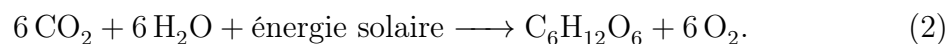
Ce sont les méristèmes³ qui permettent à la plante de se développer, en allongeant des organes déjà existants ou en créant de nouveaux organes. Lors de la germination de la graine, au stade embryonnaire, des méristèmes permettent déjà le développement de la plante (méristèmes racinaires, caulinaires). On distingue les méristèmes végétatifs, à l'origine des organes végétatifs (tige, feuilles, racines) et les méristèmes reproducteurs, à l'origine des fleurs. On distingue également les méristèmes primaires, qui permettent à la plante de croître en longueur (tiges et racines) et les méristèmes secondaires, qui permettent la croissance en épaisseur de la plante. La création de nouveaux organes est réalisée par alignement de nouvelles briques élémentaires (formée par les méristèmes), constituées :

- d'un noeud, auquel sont liés les feuilles
- d'un entrenoeud
- d'un bourgeon situé à la base du noeud, à l'aisselle des feuilles

Ces briques élémentaires sont appelées phytomères. Cette caractéristique permet de simplifier la modélisation de la croissance d'une plante.

A.2 Photosynthèse

Sans doute que la particularité la plus intéressante des plantes, et qui justifie le mieux leur étude est leur capacité à synthétiser de la matière organique (des glucides...) à partir de CO_2 , d'eau et d'énergie solaire. C'est la célèbre photosynthèse qui permet ainsi à la plante de transformer de la matière *minérale* en matière *organique*, selon l'équation



Tous les éléments de la plante participent à ce processus de photosynthèse :

- les racines puisent dans le sol l'eau et les sels minéraux nécessaires

3. Le méristème est un tissu formé de cellules indifférenciées (embryonnaires) qui se divisent activement et permettent la croissance.

- les feuilles captent l'énergie solaire et le dioxyde de carbone (grâce aux cellules chlorophylliennes et aux stomates ⁴)

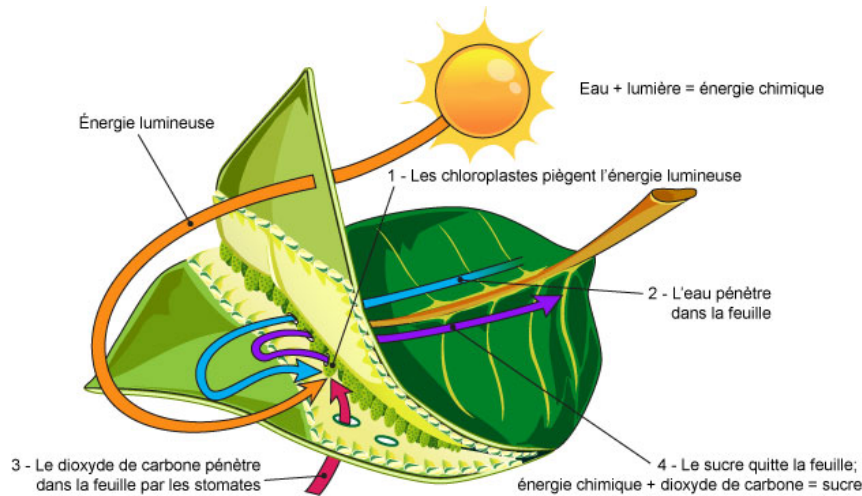


FIGURE 9 – Schéma de la photosynthèse illustrant les différents mécanismes qui permettent à celle-ci d'avoir lieu.

Les sucres ainsi formés apportent l'énergie nécessaire au fonctionnement de la plante et assurent son développement en permettant la synthèse de cellulose qui est l'élément constitutif principal des plantes. On identifie ainsi les éléments qui agissent sur la croissance de la plante :

- la lumière
- l'eau
- le dioxyde de carbone
- la température : car la température agit sur l'ouverture des stomates et donc sur le flux des échanges gazeux
- l'azote, qui permet à la plante de construire les acides aminés nécessaires à l'élaboration des protéines
- d'autres minéraux, comme le potassium qui favorise les transferts au sein de la plante

Tous les organes de la plante s'unissent donc pour aboutir à la production de biomasse. Cette biomasse est ensuite distribuée aux organes pour permettre leur développement.

Parce que ce mécanisme permet de synthétiser de la matière minérale en matière organique et de capter du CO₂, gaz à effet de serre notoire, il est crucial de comprendre ce mécanisme de photosynthèse. C'est pourquoi il est, au même titre que la croissance des plantes l'objet de nombreuses recherches (avec comme application : créer de l'électricité propre en dissociant H₂O en Oxygène et en Hydrogène, capter du CO₂...

4. Orifice de petite taille situé sur les feuilles qui permet les échanges gazeux entre l'air et la plante.

B Histoire de la modélisation de la croissance des plantes

Nous présentons dans cette partie tout d’abord l’origine de la botanique et de l’agronomie. Les premiers modèles de croissance des plantes seront ensuite décrits, puis on expliquera les avancées potentielles apparues avec l’avènement de l’informatique. Finalement, un modèle particulièrement intéressant, le modèle GreenLab, sera présenté.

Cette section reprend la démarche suivie dans l’article *Une histoire de la modélisation des plantes*, COURNÈDE et al., 2009 [12].

B.1 Débuts de la botanique et de l’agronomie

L’étude des plantes a très tôt été un domaine privilégié du savoir humain. En effet, les plantes sont un élément majeur des écosystèmes dans lesquels l’homme évolue. Elles sont source de nourriture, de remèdes, de médicaments, de matériaux, d’esthétique. Enfin, elles sont un objet scientifique d’intérêt qui a très tôt aiguïlé le sens de l’observation, l’esprit d’analyse, de synthèse, de déduction des hommes. La connaissance des plantes s’est accrue lors de l’histoire des hommes, qui ont développé la cueillette, l’agriculture, l’usage des plantes médicinales. La connaissance et l’inventaire des variétés de plante ont ainsi été des enjeux majeurs car ils permettaient la connaissance de nouveaux remèdes et étaient sources de nourritures et matériaux. L’homme a ainsi cherché à regrouper, croiser, faire croître et conserver les espèces qui lui étaient utiles.

La botanique, issue de l’étude de l’anatomie des plantes, est une science très ancienne. En témoignent les traités de classification de plantes, comme ceux d’Aristote (vers -300), ou encore l’inventaire et la description de centaines de plantes médicinales par Dioscoride (1er siècle), ainsi que les traités chinois qui inventorient les espèces utiles à l’agriculture et à la médecine traditionnelle avec de premiers efforts de classification. Efforts de classification qui se poursuivront vraiment en Europe à partir du XVII^{ème} siècle avec les premières distinctions par famille, par genre, par espèce, par structure de graine (*Les éléments de Botanique* par Joseph Pitton de Tournefort en 1656 – 1708, *Systema naturae* en 1735 et *Philosophia botanica* en 1751 par Linné et les travaux de la famille de Jussieu pendant le XVIII^{ème} siècle). Ces classifications ne sont pas objectives, elles sont le fruit d’un raisonnement *empirique*.

En parallèle, l’agronomie se développe au XVII^{ème} siècle en Europe et s’intéresse au processus de croissance et développement des plantes. Des travaux d’abord très pratiques sont réalisés sur des méthodes agricoles (labour, ensemencement, taille, greffes...), notamment ceux de Jean-Baptiste de la Quintinie et d’Olivier de Serres.

Au XIX^{ème}, les processus biologiques commencent à être étudiés de façon plus précise, en particulier la provenance du carbone, de l’azote, de l’oxygène et de l’eau dans la plante. On s’intéresse également aux problématiques de nutrition et au rôle de organes, comme en témoignent les ouvrages *Recherches chimiques sur la végétation* de

Théodore de Saussure en 1804. Quelques années plus tard, on découvre la respiration, la photosynthèse (voir équation 2 à la page 22).

La physiologie, science qui étudie le fonctionnement des plantes, se sépare alors de la botanique qui se contente de les classer.

B.2 Les premiers modèles

La modélisation mathématique précise (qui va au-delà de la simple description qualitative et fournit une description quantitative avec des capacités prédictives) n'arrive pas tout de suite en biologie. Le développement de la biologie n'a pas suivi le même schéma que celui de nombreuses autres sciences comme la physique, où l'observation a permis de tirer des concepts quantitatifs au niveau macroscopique (loi de Mariotte par exemple) avant de les expliquer par des lois qui s'appliquent au niveau microscopique (Boltzmann). De même pour la mécanique, l'optique, l'électricité avec des applications qui n'ont pas eu à attendre la compréhension au niveau atomique. La biologie végétale par contre a paradoxalement été mieux comprise au niveau cellulaire et microscopique sans que des lois précises macroscopiques en soient tirées.

Trois types de modèles vont se développer et vont changer cela : les modèles de l'architecture botanique, les modèles de production en agronomie et les modèles géométriques en informatique. Ainsi la convergence de ces trois modèles initialement séparés va permettre récemment les débuts de la modélisation précise de la croissance des plantes à la fin du XX^{ème} siècle.

L'architecture botanique va considérer la structure des plantes non plus comme une description statique issue de la classification traditionnelle mais comme le résultat de l'organogénèse des méristèmes, la cinétique de mise en place des axes feuillés, en se basant sur une combinatoire des modes de croissance, de ramification et de floraison. (Francis Hallé et Roelof Oldeman).

En parallèle, l'agronomie s'est attaquée à la prédiction de la production surfacique de biomasse. Les modèles hollandais comme celui de De Witt (1970) en sont les précurseurs. On ne considère plus la plante en elle-même mais la surface foliaire au mètre carré LAI⁵ et la production végétale par mètre carré. Les organes ne sont plus considérés individuellement mais par compartiments. A chaque compartiment est allouée une certaine quantité de la biomasse créée en fonction de sa force de puits. La force de puits d'un organe est proportionnelle à la quantité de biomasse qui sera allouée à cet organe.[6, p.229–231]

Les agronomes ont ainsi montré que la production de biomasse est proportionnelle au LAI, ainsi qu'à l'énergie utile de la lumière incidente : PAR⁶, à la lumière interceptée : I et à un facteur d'efficacité énergétique : LUE⁷. On se reporte à la loi

5. LAI : Leaf Area Index. Cela correspond au ratio entre la surface totale supérieure des feuilles vertes et la surface de sol sur laquelle se développe la culture.

6. PAR : Photosynthetically Active Radiation.

7. LUE : Light Use Efficiency.

de Beer-Lambert pour trouver la quantité de lumière interceptée, qu'on note I :

$$I = 1 - e^{-k \cdot \text{LAI}}$$

Ce qui permet ensuite de trouver la production de biomasse Q en déterminant le LUE et en mesurant la PAR.

$$Q = \text{LUE} \times \text{PAR} \times I$$

Dernier concept empirique développé, celui de temps thermique. Le temps thermique correspond à l'accumulation de températures dépassant un certain seuil :

$$\tau^{(n+1)} = \tau^{(n)} + \max[0, \underline{T^{(n)}} - T_c],$$

où $\underline{T^{(n)}}$ est l'écart de température constaté et T_c est le seuil de variation de température qu'on impose.

En effet, si on modélise la croissance de la plante en fonction du temps, cette croissance est très irrégulière et se fait par à coup. Mais si l'on considère le temps thermique on peut trouver une relation quasi-linéaire.

C Mise en place des modèles de croissance des plantes

Les modèles mathématiques de modélisation de la croissance des plantes sont généralement caractérisés par un grand nombre de processus en interaction, un grand nombre de paramètres et une acquisition coûteuse des données expérimentales. Nous présentons des éléments de bonnes pratiques de modélisation, afin de donner un aperçu global des différentes étapes de modélisation dans le cadre de la croissance des plantes. Le modèle LNAS (celui sur lequel on travaillera en premier dans le cadre du projet enjeu) est présenté comme illustration de ces méthodes, ici appliqué à la betterave, et montre comment il peut être paramétré, évalué et appliqué à la prédiction des rendements, et ce à travers des données expérimentales réelles. Ce modèle a d'intéressante capacité de prédiction lorsqu'il est couplé à de bonnes méthodes d'acquisition de données.

C.1 Caractéristiques propres aux modèles de croissance de plantes

- **Une complexité importante** au niveau des processus et des paramètres
- **Une paramétrisation difficile** à cause de cette complexité ainsi que du coût important des données.
- **Un besoin croissant** en techniques sophistiquées en informatique, mathématiques et statistiques.
- **Une importante diversité** de modèles existants sans benchmarking entre les différentes approches (lacune de méta-études statistiques)

Les solutions mathématiques et statistiques classiques et généralistes ne sont pas immédiatement adaptées à la modélisation des plantes et nécessitent un travail d'adaptation important pour prendre en compte ses spécificités.

D'un autre côté les logiciels de modélisations spécialisés, bien que performant, ne prennent pas assez en compte l'aspect paramétrisation et évaluation statistiques.

Les logiciels développés récemment dans ce domaine offrent des solutions intéressantes mais elles sont peu compatibles entre elles et limitent donc la comparaison, le benchmarking des modèles.

L'objectif est donc à la fois bien de situer les bonnes pratiques de modélisation, et de proposer une implémentation pratique, notamment à travers la plateforme Pygmalion en C++ qui fournit un template générique de modèle, des structures de données, des méthodes, des classes et framework appropriés, ainsi qu'une méthodologie statistique.

C.2 Bonnes pratiques en modélisation

- **Analyse du modèle :** Etude des comportements généraux du modèle, au niveau théorique et numérique par des simulations. Pour déterminer les données nécessaires à la paramétrisation, et faire une analyse de sensibilité pour

repérer les paramètres importants.

- **Identification du modèle :** Confronter les modèles à des données expérimentales. Identification de la structure : pour identifier dans la famille du modèle la plus intéressante. Identification des paramètres : pour identifier la valeur des paramètres pour la structure choisie.
- **Évaluation du modèle :** Vérifier qualitativement (comportement générale et aptitude de simulation) et quantitativement (comparaison aux données réelles) si le modèle atteint ses objectifs, c'est-à-dire vérifier la correspondance aux données actuelles (goodness of fit), tester ses capacités prédictives et évaluer son niveau d'incertitude.

Ces étapes ne sont pas linéaires : des allers retours sont nécessaires pour ajuster finement le modèle.

D Quelques modèles de la croissance des plantes

D.1 Informatique et modèles géométriques

L'arrivée des ordinateurs a révolutionné les méthodes de simulation ainsi que de modélisation des systèmes et l'étude des plantes en a bien sûr profité. Les ordinateurs ont fait leur apparition presque en même temps que les modèles agronomes et botaniques. Ainsi, très vite, ils ont été vus comme un moyen de simuler la structure géométrique complexe des plantes avec le développement d'arbres combinatoires, binaires et fractals. Mais ces structures purement géométriques et trop rigides ne simulent pas encore bien le développement complexe des plantes. Les travaux d'Aristide Lindenmayer aboutissent à une grammaire au formalisme puissant, grammaire générative basée sur le principe de réécriture [9].

Qu'est-ce qu'un L-Système ? Un L-Système est noté :

$$\{V, S, \omega, P\}$$

Une grammaire formelle qui comprend :

- Un ensemble alphabet **V** : Ensemble de variable du L-Système
- Un ensemble de constantes **S** : Ensemble de constantes servant notamment lors du dessin géométrique.
- Un axiome de départ ω : État initial.
- Un ensemble de règles **P** : Règles de production.

Un exemple simple : Algues de Lindenmayer

$$\left\{ \begin{array}{l} V = \{A, B\} \\ S = \{\} \\ \omega = A \\ P = (A \rightarrow AB) \wedge (B \rightarrow A) \end{array} \right.$$

Avec le résultat sur 6 générations :

- $n = 0$, A
- $n = 1$, AB
- $n = 2$, AB A
- $n = 3$, AB A AB
- $n = 4$, AB A AB AB A
- $n = 5$, AB A AB AB A AB A AB
- $n = 6$, AB A AB AB A AB A AB AB A AB AB A

Une suite de symboles générée par L-Système peut être prise en entrée par un programme informatique qui s'en servira pour générer une structure géométrique, le plus simple étant d'utiliser un turtle en 2D voire en 3D, ou encore dans un langage orienté objet avec des pointeurs on peut générer une chaîne cellulaire qui évolue. Les symboles dans V sont les parties des plantes dessinées et les symboles dans S donnent des informations sur la façon dont elles sont dessinées, leur orientation par exemple.

Ces modèles de L-système conviennent bien à la simulation des structures des plantes, elles marchent d'autant mieux combinées à la notion de temps thermique qui ordonne la dynamique de croissance et permettent d'aboutir in fine à une architecture fidèle. On constate cependant qu'elles n'intègrent pas la notion de production de biomasse et si elles permettent de prédire une structure finale aussi fidèle que possible, elles ne permettent pas de prédire le rythme de production. Les éco-physiologistes se sont alors efforcés d'intégrer des mécanismes de plus en plus fins, avec une simulation locale et géométrique de la photosynthèse, des échanges entre organes par un système de transport-résistance avec la notion de force de puits. de organes puits qui puisent la biomasse produite par les organes sources etc... Ces systèmes complexes qui permettent enfin une simulation fine au niveau individuel ne sont pourtant pas adaptés à la simulation et encore moins la modélisation d'un

grand ensemble de plantes et ceux pour deux raisons :

1. *Le coût en ressources de calcul* : il croît avec le temps de la simulation et le nombre d'individus considérés, encore plus si l'on doit considérer les interactions entre les individus.
2. *La difficulté à paramétrer* : dû au grand nombre de paramètre notamment par rapport aux données que l'on peut raisonnablement récolter.

D.2 Le modèle GreenLab : entre modèle individuel et modèle de production

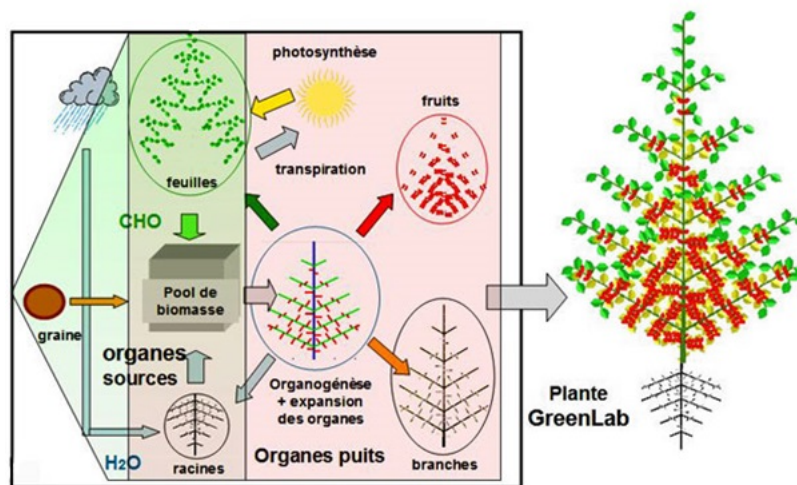


FIGURE 10 – Schéma type d'un modèle GreenLab.

Le modèle GreenLab propose une alternative. Il reprend des simplifications utiles du modèle agronome au niveau de la production : la prise en compte de l'architecture individuelle de la plante n'est pas utile au niveau d'un champ, on considère plutôt la distribution des différents types d'organes (densité de feuilles...). Autrement dit, les aspects géométriques peuvent être négligés mais pas la composition quantitative des structures. On fait alors l'hypothèse d'un *pool de biomasse* commun dans lequel les organes vont piocher, et on ne considère que la photosynthèse nette, ie les proportions de glucides qui servent effectivement à la construction de matière sèche des organes.

Au niveau de l'allocation, cette approche décrit précisément des compartiments d'organes se comportant de façon similaire, ce qui rend l'allocation de biomasse plus facile à modéliser entre organes sources et organes puits et permet de se passer d'une description géométrique ou même topologique exhaustive. Les organes sont générés par cohortes (groupes générés en même temps) de même type grâce à des équations de production des méristèmes selon leur âge physiologique, et des lois de probabilités qui déterminent la croissance, la sénescence et le branching. Et comme tous les organes d'une même cohorte d'organe du même type ont le même état, on peut opérer facilement une factorisation en sous-structures qui rend les calculs plus léger, ainsi le temps de calcul n'est plus proportionnel au nombre d'organe mais

seulement à l'âge de la plante. En particulier en multipliant le nombre d'organe de chaque cohorte par la force de puit correspondante et en additionnant le tout on obtient la demande de la plante.

Ensuite l'augmentation δq de biomasse de chaque organe est obtenu avec la formule suivante :

$$\delta q = \phi \cdot Q/D$$

Avec ϕ la force de puit, Q la pool de biomasse global et D la demande totale de la plante.

La masse des organes est la somme cumulée de l'augmentation des biomasses, on peut donc obtenir rapidement les dimensions (longueur, diamètre, surface) des organes en utilisant de simples lois géométriques (beaucoup plus simple que celles utilisées lorsque la géométrie est prise en compte dans la production).

Puis pour que le cycle se répète, la biomasse des compartiments s'obtient en sommant les cohortes de même organes, en particulier on peut obtenir la masse du feuillage puis la surface foliaire ce qui permettra de déterminer la production de biomasse au prochain cycle. La boucle est bouclée.

Pour résumer, ce modèle est un modèle dynamique de croissance des plantes qui marche par rétroaction entre croissance (production de biomasse) et développement (allocation quantitative et architecture). Le calcul de la production ne prend que peu en compte l'architecture de la plante mais seulement l'équation global de production et les relations sources-puits, ce qui permet des économies de calcul intéressantes et n'empêche pas dans un second temps de générer des structures géométriques fidèles issus d'un modèle de production simplifié mais robuste. Ainsi la plasticité des plantes est très bien représenté par ce modèle et on peut rendre compte de différents phénotypes d'une même espèce dans deux environnement très différents.



FIGURE 11 – Exemple de plantes générées grâce à GreenLab.

Cela permet notamment la visualisation en images de synthèse très fidèles et complexes de plantes dont la croissance a été modélisé sans prendre en compte le détail

géométrique de cette même structure, donc avec un temps de calcul très intéressant. L'augmentation de biomasse de chaque organe a déjà été simplement déterminée grâce aux équations précédentes, dont on déduit également la forme et le volume, de simples règles géométriques positionnent ces organes dans l'architecture selon l'empilement des entre-nœuds autour d'un axe, la phyllotaxie, l'angle de branchement, la courbure des axes, comme on peut le visualiser en Figure 11.

Comme montré dans la Figure 12, on peut même simuler des paysages entiers grâce à ces méthodes, avec le raffinement des paysages fonctionnels qui prennent en compte raffinement les interactions plantes-environnement, la distribution des ressources hydriques et des radiations lumineuses entre plantes qui sont en compétition.

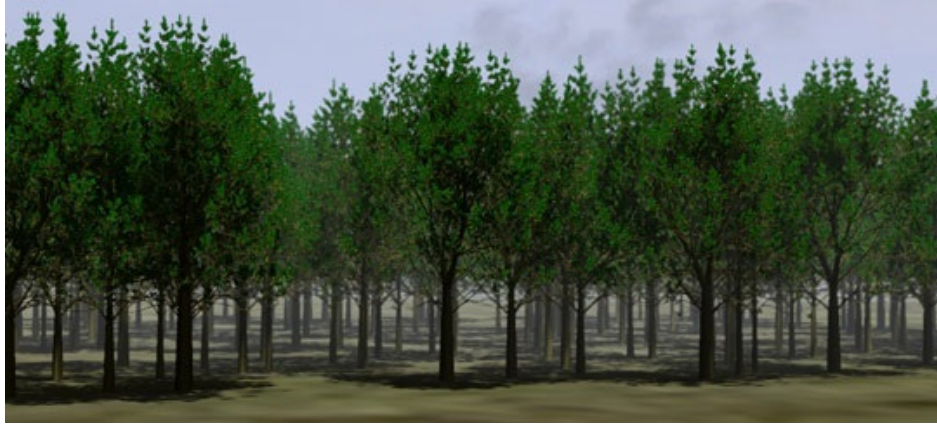


FIGURE 12 – Paysage fonctionnel généré grâce à GreenLab

D.3 Modèle LNAS appliqué à la betterave

Le modèle LNAS (Long Normal Allocation and Senescence) appliqué à la betterave. C'est un modèle innovant et suffisamment simple pour illustrer tous les enjeux de la thèse. Il est plutôt robuste étant donné sa simplicité, et le faible nombre de données et paramètre requis le rend pratique à l'utilisation. Il traite la production de biomasse au niveau compartemental et peut être considéré comme une simplification du modèle Greenlab qui décrit les mêmes processus au niveau des organes. C'est un modèle Markovien en temps discret.

+Production de Biomasse :

$$Q(t) = (\mu \cdot \text{PAR}(t)(1 - e^{-\lambda Q_g(t)}) \cdot (1 + \eta_Q(t))$$

+Allocation au feuillage :

$$Q_f(t+1) = Q_f(t) + \gamma(t) \cdot Q(t)$$

+Allocation à la racine :

$$Q_r(t+1) = Q_r(t) + (1 - \gamma(t)) \cdot Q(t)$$

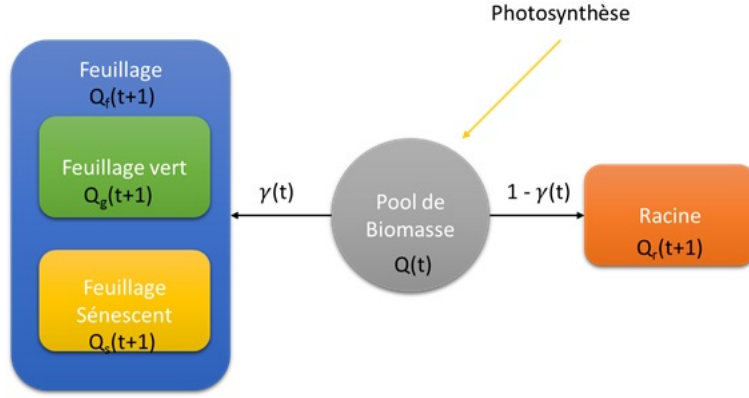


FIGURE 13 – Schéma du modèle LNAS appliqué à la betterave

+Fonction d'allocation :

$$\gamma(t) = (\gamma_0 + (\gamma_f - \gamma_0) \cdot G_a(\tau(t))) \cdot (1 + \eta_\gamma(t))$$

+Fonction de sénescence :

$$Q_s(t) = G_s(\tau(t) - \tau_{sen})Q_f(t)$$

+Part de la biomasse produite qui arrive aux feuilles vertes :

$$Q_g(t) = Q_f(t) - Q_s(t)$$

- $Q(t)$: Production de biomasse au jour t par unité de surface.
- μ : Efficacité énergétique
- $1 - e^{-\lambda Q_g(t)}$: Fraction des radiations interceptées
- $PAR(t)$: quantité de radiations photosynthétiquement actives par unité de surface.
- λ : paramètre
- $Q_g(t)$: masse total des feuilles vertes au jour t
- $Q_f(t)$: masse total du feuillage au jour t
- $Q_s(t)$: masse total du feuillage sénescant au jour t
- γ_0, γ_f : paramètres
- $\eta(t)$: variables aléatoires normales
- $G(t)$: variables aléatoires lognormales
- $\tau(t)$: temps thermique au jour t (température cumulée depuis l'émergence de la plante)
- τ_{sen} : temps thermique où la sénescence commence.
- Les paramètres des variables aléatoires font parti des paramètres du modèle.

E Tableaux des variables du modèle informatique

Modèle mathématique	Implémentation
Q_r	root_biomass
Q_s	stem_biomass
Q_l	green_leaf_biomass
Q_g	grain_biomass
Q_y	yellow_leaf_biomass
θ	soil_humidity
τ	thermal_time
R	soil_contained_water
E_s	soil_water_evaporated
T_p	water_transpired
z_r	root_horizon
SSI	stomatal_stress_index
TSI	thermal_stress_index
TSI_{\uparrow}	thermal_stress_index
TSI_{\downarrow}	soil_thermal_stress_index
q	produced_biomass
LAI	leaf_area_index
E_{spot}	soil_req_evaporation
T_{ppot}	req_transpiration
E_{smax}	soil_max_evaporation
T_{pmax}	max_transpiration
T_{\downarrow}	soil_temperature
α_g	alpha_g
α_s	alpha_s
α_r	alpha_r
α_l	alpha_l
β_s	beta_s
β_l	beta_l
γ_l	gamma_l
γ_y	gamma_y

TABLE 1 – Variables d'état du modèle mathématique et leurs noms dans l'implémentation.

Modèle mathématique	Implémentation
T	vec_ext_temperature
PAR	vec_par
W	vec_water_input
ET0	vec_et0

TABLE 2 – Variables de *contrôle* du modèle mathématique et leurs noms dans l'implémentation.

Modèle mathématique	Implémentation
t_c	t_c
t_{opt}	t_opt
μ_g	mu_g
σ_g	sigma_g
η_s	eta_s
η_l	eta_l
τ_l	tau_l
μ_l	mu_l
σ_l	sigma_l
τ_y	tau_y
μ_y	mu_y
σ_y	sigma_y
RUE	rue
λ	lambda
ρ_l	rho_l
K_c	k_c
K_s	k_s
θ_{max}	theta_max
θ_{min}	theta_min
z_s	z_s
z_m	z_m
ρ_r	rho_r
$t_{\downarrow c}$	t_soil_c
$t_{\downarrow opt}$	t_soil_opt

TABLE 3 – Paramètres du modèle mathématique et leurs noms dans l'implémentation.

F Planning

Date	
11/01 – 16/01	<ul style="list-style-type: none">- Discussion après l'étude documentaire- Rendez-vous avec le client : discuter sur la soutenance d'étude documentaire et fixer le planning après l'étude documentaire- Installation de GitLab et de Julia
'18/01 – 30/01	<ul style="list-style-type: none">- Prise en main de GitLab et Julia- Création d'une nouvelle branche « sbt11 » sur la plateforme de GitLab- Lecture de modèles généraux existants et de documentations sur l'implémentation avec Julia
01/02 – 13/02	<ul style="list-style-type: none">- Assimilation du modèle LNAS appliqué au blé- 1^{ère} implémentation du modèle LNAS en Julia sur la plateforme GitLab- Contacts avec Pierre Carmier pour répondre à nos interrogations
15/02 – 03/03	<ul style="list-style-type: none">- Résolution de certains problèmes rencontrés pendant l'implémentation de code- Amélioration du code pour le rendre plus lisible, avec des commentaires
14/03 – soutenance	<ul style="list-style-type: none">- Contacts avec le client : questions/réponses sur les problèmes rencontrés- Utilisation du code sur des données expérimentales- Rédaction du rapport et préparation de la soutenance
Après la soutenance	<ul style="list-style-type: none">- Analyse de sensibilité du modèle pour déterminer l'influence des différents paramètres sur les résultats du modèle.- ...

FIGURE 14 – Planning de l'organisation des travaux au cours du semestre.

G Article modèle LNAS wheat

Le modèle se trouve dans les pages suivantes.

LNAS model for wheat with soil coupling

Pierre Carmier

Equipe Digiplante, Laboratoire MAS, CentraleSupélec, France

(Dated: November 18, 2015)

The main idea behind this model is to introduce allocation functions which describe how the daily produced biomass is allocated to the different organs of the plant. It is a rather simple model in which the plant can be seen as a black box which converts environmental resources (provided on a stochastic basis) into biomass. With respect to sugarbeet, wheat represents more of a challenge to modelize due to the larger number of organs involved in the growth process (see Fig. 1). In the following, we shall provide a detailed yet simple model which describes the coupling of shoot and root processes and enforces the balance of water and biomass at all times.

A. Shoot part

The compartmental circulation of biomass for organ $o = \{\text{grain, stem, root, green leaf}\}$ obeys the following equation:

$$Q_o^{(n+1)} = (1 - \beta_o^{(n)} - \gamma_o^{(n)})Q_o^{(n)} + \alpha_o^{(n)}q^{(n)}, \quad (1)$$

with α , β , γ respectively allocation, remobilisation and senescence functions. The case of yellow leaves is special, since it thrives on the senescence of green leaves and receives no allocation from the biomass pool: $Q_y^{(n+1)} = (1 - \gamma_y^{(n)})Q_y^{(n)} + \gamma_l^{(n)}Q_l^{(n)}$. The various functions responsible for biomass circulation depend on the thermal time $\tau(t) = \tau_d^{-1} \int_0^t dt' \max[0, T(t') - T_c]$, such that

$$\tau^{(n+1)} = \tau^{(n)} + \max[0, \underline{T}^{(n)} - T_c] \quad (2)$$

where T_c is a cultural-dependent base temperature. For example, assuming that allocation to grain only starts after a characteristic time τ_g and is asymptotically exhaustive, one can parametrize the corresponding allocation function as a cumulative log-normal distribution:

$$\alpha_g^{(n)} = F_{\log \mathcal{N}(\mu_g, \sigma_g)}(\tau^{(n)} - \tau_g) = \frac{1}{2} \left(1 + \operatorname{erf} \left[\frac{1}{\sigma_g \sqrt{2}} \log \left(\frac{\tau^{(n)} - \tau_g}{\tau_{1/2} - \tau_g} \right) \right] \right), \quad (3)$$

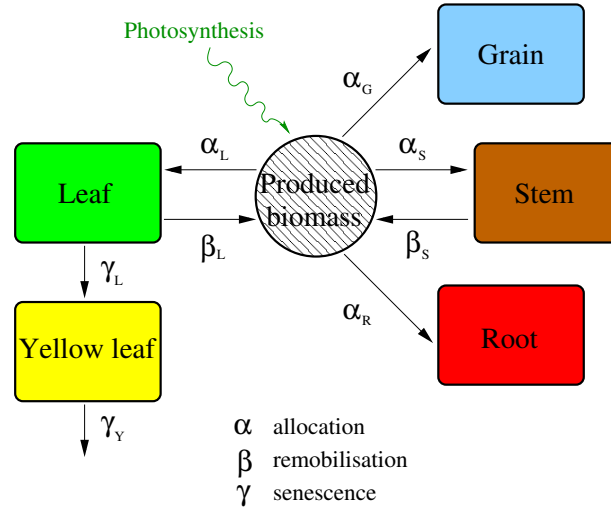


Figure 1: Circulation of biomass inside the plant modeled as different organ compartments, in terms of allocation, remobilisation and senescence functions.

with $\mu_g = \log(\tau_{1/2} - \tau_g)$ and σ_g the mean and variance of the distribution, $\tau_{1/2}$ the time at which $\alpha_g = 1/2$, and $\text{erf}(x) = (2/\sqrt{\pi}) \int_0^x dt e^{-t^2}$ the so-called error function. In first approximation, other allocation functions can be taken as complementary to α_g : $\alpha_{s,l,r}^{(n)} = \alpha_{s,l,r}^{(0)}(1 - \alpha_g^{(n)})$, with $\alpha_s^{(0)} + \alpha_l^{(0)} + \alpha_r^{(0)} = 1$. Remobilisation and senescence functions can likewise be parametrized by lognormal cumulative distributions: $\beta_o = \eta_o F_{\log \mathcal{N}(\mu_l, \sigma_l)}(\tau - \tau_l)$, where η_o is the fraction of biomass asymptotically involved in the process, $\gamma_l = (1 - \eta_l) F_{\log \mathcal{N}(\mu_l, \sigma_l)}(\tau - \tau_l)$ and $\gamma_y = F_{\log \mathcal{N}(\mu_y, \sigma_y)}(\tau - \tau_y)$. Here τ_l and τ_y respectively mark the beginning of the remobilisation/senescence phase and the fall of yellow leaves.

The pool of allocated biomass is itself given as the sum of remobilized biomass and of biomass produced by photosynthesis:

$$q^{(n)} = \text{RUE} \min[\text{SSI}^{(n)}, \text{TSI}_{\uparrow}^{(n)}] \text{PAR}^{(n)} (1 - e^{-\lambda \text{LAI}^{(n)}}) + \sum_o \beta_o^{(n)} Q_o^{(n)}, \quad (4)$$

with RUE the radiation use efficiency, SSI and TSI_{\uparrow} stress indices (defined in section C), PAR the photosynthetically active radiation and $\lambda = \sin \phi$, where ϕ is the average angle between a leaf and the incoming light. The leaf area index $\text{LAI}^{(n)} = S_l^{(n)}/S = \rho_l Q_l^{(n)}$ appears here as a geometrical factor inside the Beer-Lambert law for light interception: $(1 - e^{-\lambda \text{LAI}})$ represents the soil coverage probability for a density LAI of leaves randomly distributed in space (but oriented with angle ϕ with respect to incoming light).

B. Root and soil part

The water balance equation reads

$$R^{(n+1)} = R^{(n)} + \underline{W}^{(n)} - \text{Es}^{(n)} - \text{Tp}^{(n)} - d^{(n)}, \quad (5)$$

where W is the daily water input coming from precipitations and irrigation, Es is the water lost by evaporation from the soil, Tp is the water transpired by the plant and d is a drain function which takes into account the fact that the soil can only withhold a certain quantity of water which is bounded by the field capacity R_{\max} . Starting from a continuous description, the soil water content can be expressed as

$$R^{(n)} = \int_0^{z_M} dz \theta^{(n)}(z) \quad (6)$$

with θ the depth-dependent soil humidity and z_M the soil depth. The field capacity is simply $R_{\max} = \int_0^{z_M} dz \theta_{\max}(z)$. Actually, due to the inability of the plant to extract the water from the soil (using capillary forces) below a certain soil humidity threshold θ_{\min} referred to as the permanent wilting point, one can define the transpirable water content in the soil as $\int_0^{z_M} dz \max[0, \theta^{(n)}(z) - \theta_{\min}(z)]$. Daily evaporation and transpiration requirements, given the environmental conditions, are expressed as

$$\text{Espot}^{(n)} = K_s \underline{\text{ET0}}^{(n)} e^{-\lambda \text{LAI}^{(n)}}, \quad (7)$$

$$\text{Tppot}^{(n)} = K_c \underline{\text{ET0}}^{(n)} (1 - e^{-\lambda \text{LAI}^{(n)}}). \quad (8)$$

Here ET0 is the potential evapotranspiration evaluated using the Penman-Monteith formula for grass, and $K_s(K_c)$ are cultural coefficients (adapting the formula for wheat). On the other hand, daily maximal evaporation and transpiration, given the available water resources, are expressed as

$$\text{Esmax}^{(n)} = \int_0^{z_M} dz \left(\min[\theta^{(n)}(z), \theta_{\min}(z)] e^{-z/z_S} + \frac{\max[0, \theta^{(n)}(z) - \theta_{\min}(z)] e^{-z/z_S}}{\max[1, e^{-z/z_S} + e^{-z/z_r^{(n)}}]} \right), \quad (9)$$

$$\text{Tpmax}^{(n)} = \int_0^{z_M} dz \frac{\max[0, \theta^{(n)}(z) - \theta_{\min}(z)] e^{-z/z_r^{(n)}}}{\max[1, e^{-z/z_S} + e^{-z/z_r^{(n)}}]}, \quad (10)$$

with z_S the surface layer of the soil and $z_r^{(n)} = \rho_r Q_r^{(n)}$ the root horizon at day n . Water partition in the soil can then be described sequentially at each time step:

1) Update the water content after input and possible drainage: $R^{(n)} = \min[R^{(n)} + \underline{W}^{(n)}, R_{\max}]$. In other words, this amounts to finding $z_W^{(n)}$ such that $\underline{W}^{(n)} = \int_0^{z_W^{(n)}} dz (\theta_{\max}(z) - \theta^{(n)}(z))$, and update $\theta^{(n)}(z) = \theta_{\max}(z)$ for $z \leq \min[z_W^{(n)}, z_M]$.

2) Update the water content after evaporation and transpiration: $R^{(n+1)} = R^{(n)} - \text{Es}^{(n)} - \text{Tp}^{(n)}$, with

$$\text{Es}^{(n)} = \min [\text{Esmax}^{(n)}, \text{Espot}^{(n)}] , \quad (11)$$

$$\text{Tp}^{(n)} = \min [\text{Tpmx}^{(n)}, \text{Tppot}^{(n)}] . \quad (12)$$

In other words, this amounts to finding $z_{\text{Es}}^{(n)}$ and $z_{\text{Tp}}^{(n)}$ such that $\text{Es}^{(n)} = \int_0^{z_{\text{Es}}^{(n)}} dz \dots$ and $\text{Tp}^{(n)} = \int_0^{z_{\text{Tp}}^{(n)}} dz \dots$, and update $\theta^{(n)}(z)$ accordingly.

C. Effect of hydric and thermal stresses

Hydric stress can be accounted for using the stomatal stress index

$$\text{SSI}^{(n)} = \frac{\text{Tp}^{(n)}}{\text{Tppot}^{(n)}} = \min[1, \frac{\text{Tpmx}^{(n)}}{\text{Tppot}^{(n)}}] , \quad (13)$$

which describes the falling of the transpirable water reserve in the soil below the daily requirement. The main impact of hydric stress will be to reduce the daily produced biomass in Eq. (4). Recalling that root horizon is proportional to root biomass, such that $z_r^{(n+1)} = z_r^{(n)} + \rho_r \alpha_r^{(n)} q^{(n)}$, hydric stress could also favor root growth by having an impact on allocation functions, for example taking

$$\frac{\alpha_r^{(n)}}{1 - \alpha_g^{(n)}} = 1 - (1 - \alpha_r^{(0)}) \text{SSI}^{(n)} , \quad \frac{\alpha_{l,s}^{(n)}}{1 - \alpha_g^{(n)}} = \alpha_{l,s}^{(0)} \text{SSI}^{(n)} . \quad (14)$$

The effect of thermal stress on root growth can furthermore be incorporated as $\alpha_r^{(n)} \rightarrow \text{TSI}_{\downarrow}^{(n)} \alpha_r^{(n)}$, with the soil thermal stress index

$$\text{TSI}_{\downarrow}^{(n)} = \chi_{[0,1]} \left(\frac{T_{\downarrow}^{(n)} - T_{\downarrow c}}{T_{\downarrow \text{opt}} - T_{\downarrow c}} \right) , \quad \text{with the characteristic function } \chi_{[a,b]}(x) = \begin{cases} x & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases} . \quad (15)$$

Soil temperature T_{\downarrow} is actually depth-dependent, but we only need its value at root horizon. Its evolution proceeds as $T_{\downarrow}^{(n+1)} = T_{\downarrow}^{(n)} + (\underline{T}^{(n+1)} - \underline{T}^{(n)}) e^{-z_r^{(n)}/z_S}$. Note that $\alpha_l^{(n)}$ and $\alpha_s^{(n)}$ will also be impacted by $\text{TSI}_{\downarrow}^{(n)}$, but with an opposite trend, since biomass conservation ensures that biomass not allocated to the roots must necessarily be allocated to other organs: thus, $\alpha_{l,s}^{(n)} = \alpha_{l,s}^{(0)} (1 - \alpha_g^{(n)}) [\text{SSI}^{(n)} \text{TSI}_{\downarrow}^{(n)} + (1 - \text{TSI}_{\downarrow}^{(n)}) / (1 - \alpha_r^{(0)})]$. Finally, the influence of thermal stress above ground on the daily produced biomass must also be taken care of using:

$$\text{TSI}_{\uparrow}^{(n)} = \chi_{[0,1]} \left(\frac{\underline{T}^{(n)} - T_c}{T_{\text{opt}} - T_c} \right) . \quad (16)$$

D. List of variables and parameters

Main state variables: $\mathbf{X} = \{Q_r, Q_s, Q_l, Q_y, Q_g, \tau, \theta, R, \text{Es}, \text{Tp}, z_r, \text{SSI}, \text{TSI}_{\uparrow(\downarrow)}\}$

Auxiliary state variables: $\{q, \text{LAI}, \text{Espot}, \text{Tppot}, \text{Esmx}, \text{Tpmx}, d, T_{\downarrow}\}$

Environmental variables: $\mathbf{E} = \{T, \text{PAR}, W, \text{ET0}\}$

Shoot parameters: $\mathbf{P}_{\uparrow} = \{T_c, T_{\text{opt}}, \tau_g, \mu_g, \sigma_g, \eta_s, \eta_l, \tau_l, \mu_l, \sigma_l, \tau_y, \mu_y, \sigma_y, \text{RUE}, \lambda, \rho_l, K_c(K_s)\}$

Soil and root parameters: $\mathbf{P}_{\downarrow} = \{\theta_{\max}(\min), z_S, z_M, \rho_r, T_{\downarrow c}, T_{\downarrow \text{opt}}\}$

E. Initialisation

$$\theta^{(0)}(z) = \theta_{\max}(z), \tau^{(0)} = 0, \alpha_{s,r,l}^{(0)} = 1/3, Q_{r,s,l}^{(0)} = ?$$