



**HAL**  
open science

# Développement d'algorithmes pour l'estimation de vitesses de croissance au départ de séries chronologiques de longueurs racinaires

Pierre Faux

► **To cite this version:**

Pierre Faux. Développement d'algorithmes pour l'estimation de vitesses de croissance au départ de séries chronologiques de longueurs racinaires. Génétique des plantes. 2007. hal-04190224

**HAL Id: hal-04190224**

**<https://hal.inrae.fr/hal-04190224v1>**

Submitted on 29 Aug 2023

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

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

**Université Catholique de Louvain**  
Faculté d'Ingénierie Biologique, Agronomique et  
Environnementale

Développement d'algorithmes pour  
l'estimation de vitesses de croissance au  
départ de séries chronologiques de  
longueurs racinaires.

Mémoire présenté par  
**Pierre FAUX**  
en vue de l'obtention du grade  
de Bio-Ingénieur en Agronomie.

Promoteur: **Professeur Xavier DRAYE**

Année académique 2006 - 2007



Je tiens en premier lieu à exprimer ma profonde gratitude au Professeur Xavier Draye. Sa sympathie hors du commun, son extrême disponibilité, son inénarrable sens de l'humour sont autant de qualités qui, lorsqu'elles sont le fait d'un homme d'une telle science et d'un tel discernement dans les différentes situations auxquelles nous avons été confrontés, ne peuvent que combler les mémorisants qu'il promet.

J'adresse également mes plus sincères remerciements à Tristan Lavigne, qui a mis à profit la connaissance acquise durant plusieurs années de doctorat pour nous aider à réaliser le dispositif expérimental utilisé dans notre expérience principale et qui a été d'un conseil fort appréciable dans la conduite des cultures aéroniques. Tristan, merci également pour l'ambiance chaleureuse que tu as contribué à créer à plus d'une reprise.

Troisième membre de l'équipe, Jean-François fut, pour moi, le collègue de travail idéal. Merci, *Jef*, pour ton humour, ta sympathie et ton aide tout au long de l'année. Qu'il s'agisse du labo, du café, du Cercle Agro, des activités du Semeur ou de Michamps-la-Plage, ce sont autant de bons moments passés en ta compagnie que je retiendrai.

Quatrième membre de l'équipe, Pierre a pris une pause involontaire au beau milieu de cette année. Je tiens également à te remercier pour tous les bons moments passés au premier quadrimestre en ta compagnie, cher « *ami de la covariance* », et pour ton soutien au second quadrimestre et te souhaiter un rétablissement que j'espère prompt pour rapidement revenir « *faire parler les chiffres* »

Mes remerciements vont ensuite à tous les membres du laboratoire d'écophysiologie et d'amélioration végétale. Merci donc, à MM. Foucart et Ledent pour le regard porté sur nos cultures, à Aurélie et AnneMich pour leur aide lors des opérations de récolte, à Willy pour sa présence et son rangement, à Anne-Sophie, à Sophie, Audrey, Inès, Adelphine, Claire-Noëlle, M. Bertin, Charles, Khalid, Gustave, Gisèle, Isabelle, Pierre et Olivia pour leur soutien.

En ce qui concerne l'expertise scientifique, je tiens à remercier, outre MM. Foucart et Ledent, M. César de la Clinique des Plantes et M. Maraite, Professeur de phytopathologie.



Merci particulièrement à notre plus fidèle supporter, Milan '*Slobodan*' Urban. Ses visites, toujours inattendues, étaient l'occasion d'une détente en pleine expérience et ses photos ont été mises à profit dans la rédaction du présent mémoire.

Merci également à mes oncles et tantes pour l'hébergement lors de la rédaction du mémoire: Marie-Rose, Luc, Françoise, Michel et, bien sûr, Ernest.

Merci à Gilles, Jean-Philippe et Grégoire pour l'aide qu'ils ont bien consenti à m'offrir à certains passages exigeant en travail, qu'il s'agisse de couper des papier-filtres, de trouer des plaques de polystyrène ou de peser des racines.

Merci à mon père, ma mère, mes frères et mes soeurs, mes cousines et cousins d'avoir daigné me soutenir tout au long de la rédaction de ce mémoire, notamment en corrigeant certains premiers jets. Je remercie ici encore une fois Anne-Michelle pour ses conseils avisés et la qualité de ses corrections.

Merci enfin à tous les membres du Semeur avec qui j'ai passé une excellente dernière année et grâce à qui j'ai pu *combiner* mémoire et XIIIème Journée du Monde Rural. *Si ça compte*, ça sera encore mieux!

Et, pêle-mêle, merci aux concepteurs d'iMovie pour les moments de détente interrompant la douce mélancolie de la rédaction, au service de prêt du département COMU pour leur communication interne, aux habitants de Court-Saint-Etienne pour être des stéphanois, à la Hongrie pour son arrière-saison, à Suprême et Tina pour le bassin et les aplombs de Berline, à Julien Lepers, à Wikipedia etc ...

La caractérisation tri-dimensionnelle du système racinaire est une étape primordiale dans l'étude des fonctions de ce système. Alors qu'il est pour ainsi dire impossible de capturer l'architecture d'un système racinaire de manière satisfaisante, les outils de modélisation permettent d'en reconstituer une vue réaliste à partir de paramètres dynamiques décrivant son développement, en particulier les vitesses de croissances des différents types de racines d'une plante.

Notre travail vise à proposer des méthodes d'estimation de vitesses de croissance racinaire à partir d'une série chronologique de longueurs racinaires mesurées. Choisir ensuite une de ces méthodes d'estimation et l'appliquer à un ensemble de données constitue notre second objectif.

Nous avons mené une expérience de suivi (mesure de longueurs racinaires et observation de la croissance, au moyen d'un système code-barres) de plants d'orge (*Hordeum vulgare* ssp *spontaneum*) cultivés dans un dispositif aéronique. La mise au point de ce dispositif et le choix de la méthode de mesure ont requis plusieurs adaptations. Les longueurs de racines adventives (non identifiées distinctement entre elles) de plus de 600 plantes, de 149 génotypes différents issus de deux parents, à 5 dates espacées d'environ 4 jours ont ainsi été obtenues.

Les méthodes proposées se concentrent sur la mise en correspondance, sur la durée de l'expérience, des longueurs racinaires mesurées. De ce fait, elles permettent de retracer l'évolution d'une racine et d'en estimer sa vitesse de croissance. Le paradigme de chaque méthode est détaillé et chacune est implémentée en un algorithme écrit en langage MATLAB®.

La première et la seconde méthode reposent sur une approche combinatoire. La première évalue toutes les combinaisons de longueurs mesurées à deux dates de mesure naturellement possibles. La seconde évalue les combinaisons probables de longueurs mesurées à au moins trois dates successives. Ces méthodes imposent néanmoins que les racines retracées soient mesurées à chaque date, ce qui peut ne pas être le cas dans les données que nous avons générées.

Cette dernière raison nous pousse à proposer une troisième méthode. Celle-ci opère de manière inverse aux deux premières. D'abord, elle génère une première estimation des vitesses de croissance par maximisation de la proximité entre la croissance attendue d'une racine et les longueurs mesurées. Ensuite, elle retrace une racine en combinant de façon cohérente les longueurs mesurées maximisant cette proximité.

Nous appliquons finalement l'algorithme implémenté à partir de cette dernière méthode sur l'ensemble des longueurs de racines adventives. Ces résultats nous permettent d'estimer une vitesse moyenne de croissance des racines adventives pour chaque plante. Une analyse génétique (Dumasy, 2007) sur ce paramètre a révélé une région d'intérêt sur le chromosome IV de l'orge.

En outre, diverses voies d'amélioration de l'algorithme implémenté et d'application de la méthode retenue sont présentées.

*MOTS-CLÉS*

---

croissance racinaire – culture aéroponique – modélisation inverse – série chronologique –  
analyse des données

<b>Préambule</b> .....	<b>1</b>
<b>Chapitre I: Objectifs du mémoire</b> .....	<b>3</b>
<b>Chapitre II: Synthèse bibliographique</b> .....	<b>5</b>
II. 1. Le système racinaire .....	6
a. Physiologie du système racinaire .....	6
§ <i>Organisation d'une racine de Monocotylée</i> .....	6
§ <i>Le système racinaire</i> .....	7
b. Paramètres d'intérêt .....	9
c. Méthodes de mesure, d'observation et d'estimation des paramètres d'intérêt .....	11
§ <i>Mesure et observation</i> .....	11
§ <i>Estimation de paramètres</i> .....	11
II. 2. Modèles de simulation de la RSA et de croissance racinaire .....	14
a. Root Typ, modèle générique de simulation de la RSA .....	14
§ <i>Croissance axiale</i> .....	14
§ <i>Croissance radiale</i> .....	15
b. GRAAL, modèle fonctionnel et structurel .....	15
II. 3. L'orge .....	16
a. Généralités .....	16
b. Croissance jusqu'au stade 3 talles .....	16
II. 4. Modélisation inverse: principes et outils d'application .....	17
a. Principes de modélisation inverse .....	17
b. Outils d'application .....	19
<b>Chapitre III: Partie expérimentale</b> .....	<b>21</b>
III. 1. Principes de culture aéroponique .....	22
a. Principes de fonctionnement .....	22
b. Contrôle et réglages du dispositif .....	22
c. Conception d'un dispositif aéroponique .....	23
III. 2. Matériel végétal .....	24
III. 3. Conditions de culture .....	24
a. Environnement contrôlé .....	24
b. Solution nutritive .....	25
c. Germination .....	25
d. Repiquage .....	26
III. 4. Expériences préliminaires .....	27
a. Expérience 1: Test code-barres .....	27
§ <i>Objectifs</i> .....	27
§ <i>Matériel</i> .....	27
§ <i>Principe et méthodes de mesure au code-barres</i> .....	28
§ <i>Matériel végétal</i> .....	30
§ <i>Protocole de mesure et d'observation</i> .....	30
§ <i>Résultats</i> .....	31
b. Expérience 2: Test traitement par imagerie .....	32
§ <i>Objectifs</i> .....	32
§ <i>Matériel</i> .....	32
§ <i>Méthode de mesure par traitement d'images numérisées</i> .....	32
§ <i>Matériel végétal</i> .....	32
§ <i>Protocole de mesure</i> .....	32

§ Résultats .....	33
c. Expérience 3: Mise au point d'un bac de grande dimension .....	34
§ Objectifs .....	34
§ Matériel .....	34
§ Matériel végétal .....	36
§ Méthodes de mesure et d'observation .....	36
§ Protocole de mesure .....	36
§ Résultats .....	37
III. 5. Expérience principale .....	37
a. Conditions de culture .....	38
b. Plan expérimental .....	40
c. Matériel végétal .....	42
d. Types de mesures réalisées et protocoles de ces mesures .....	42
e. Méthode et algorithme de pré-traitement de l'information brute et d'encodage des différents protocoles de mesure .....	43
§ Méthode de lecture des fichiers de données .....	43
§ Description de la fonction de pré-traitement de l'information .....	45
<b>Chapitre IV: Correspondances sur deux dates par calcul combinatoire .....</b>	<b>47</b>
IV. 1. Introduction et hypothèses .....	48
§ Mise en situation .....	48
§ Hypothèses .....	50
IV. 2. Algorithme paradigmatique de correspondance .....	50
§ Schéma paradigmatique .....	50
§ Indicateurs implémentés .....	52
§ Pseudo-code de l'algorithme .....	54
IV. 3. Cas particuliers et avatars de l'algorithme .....	57
§ Cas particuliers .....	57
§ Avatars .....	58
IV. 4. Conclusions .....	58
<b>Chapitre V: Correspondances sur plus de deux dates par calcul combinatoire .....</b>	<b>59</b>
V. 1. Introduction et hypothèses .....	60
§ Mise en situation .....	60
§ Hypothèses .....	61
V. 2. Algorithme paradigmatique de correspondance .....	62
§ Schéma paradigmatique .....	62
§ Indicateur implémenté .....	65
§ Pseudo-code de l'algorithme .....	66
V. 3. Cas particuliers et avatars de l'algorithme .....	68
§ Cas particuliers .....	68
§ Avatars .....	68
V. 4. Evaluation et conclusions .....	70
<b>Chapitre VI: Correspondance sur plus de deux dates par optimisation des vitesses de croissance .....</b>	<b>71</b>
VI. 1. Introduction et hypothèses .....	72
§ Mise en situation .....	72
§ Hypothèses .....	75
VI. 2. Algorithme paradigmatique de correspondance .....	76
§ Schéma paradigmatique .....	76
§ Indicateur implémenté .....	80

§ <i>Pseudo-code de l'algorithme</i> .....	82
VI. 3. Cas particuliers et avatars de l'algorithme .....	84
§ <i>Approximation globale de l'optimum</i> .....	84
§ <i>Cas des ancres parasites</i> .....	84
§ <i>Cas des fausses redondances</i> .....	87
§ <i>Option graphique</i> .....	87
VI. 4. Conclusions et perspectives d'amélioration de l'algorithme .....	90
§ <i>Conclusions</i> .....	90
§ <i>Perspectives d'amélioration de l'algorithme</i> .....	94
<b>Chapitre VII: Conclusions générales</b> .....	<b>95</b>
<b>Glossaire</b> .....	<b>98</b>
<b>Références bibliographiques</b> .....	<b>99</b>
<b>Annexes électroniques</b> .....	<b>102</b>
<b>Annexes</b> .....	<b>104</b>



# Préambule

Développer un dispositif aéroponique de croissance pour l'orge cultivée (*Hordeum vulgare* ssp *spontaneum*) et une méthode de mesure et d'observation rapides et aisés de son système racinaire étaient les objectifs de départ des expériences préliminaires menées dès septembre 2006 par Jean-François Dumasy et moi-même, sous la houlette du Pr. Xavier Draye et de son doctorant, Tristan Lavigne. Suite à la réalisation de mesure et d'observations dans un dispositif mis au point lors de l'expérience principale de ce mémoire, cet objectif devait nous mener, Jean-François et moi-même, dans deux voies différentes, respectivement l'analyse génétique et la modélisation de paramètres dynamiques du système racinaire.

Si Jean-François a effectivement mené son objectif jusqu'à son terme, le mien a été redéfini du fait d'un problème de mise en correspondance de racines à travers le temps.

En effet, lors de notre expérience principale, nous suivons la croissance des racines séminales et adventives de plantes identifiables tout au long de l'expérience. Le problème réside dans le fait que les racines ne sont pas identifiables. Nous mesurons ainsi, à cinq dates espacées d'environ 4 jours, un ensemble de plusieurs racines adventives appartenant à une même plante. Cet ensemble est nommé, du fait de la configuration spatiale des racines, faisceau adventif.

Afin de pouvoir prédire leur vitesse de croissance, nous développons donc une méthode *a posteriori*, c'est-à-dire une méthode de modélisation inverse, qui permet, à partir d'un certain nombre d'hypothèses et des mesures effectuées, de mettre en correspondance les différentes racines appartenant à une même plante.

Cette méthode doit faire face à plusieurs contraintes. De date en date, une racine peut avoir disparu, une nouvelle racine peut être apparue et l'expérimentateur peut éronément mesurer une racine qui ne devait pas l'être. Ces différentes contraintes doivent être intégrées à la méthode.

Résoudre ce problème de mise en correspondance offre plusieurs perspectives. A commencer par l'analyse génétique de recherche de QTL impliqués dans le déterminisme du caractère *'vitesse de*



*croissance*'. Cette première perspective s'insère dans le mémoire de Jean-François Dumasy. Ensuite, la vitesse de croissance est un paramètre d'intérêt pour simuler une architecture racinaire. Enfin, ce problème de mise en correspondance est développé ici pour des mesures effectuées en un environnement contrôlé. Extrapoler son application à d'autres environnements dans lesquels la prise de mesure est plus difficile constitue également une perspective intéressante.

# Chapitre I

## *Objectifs du mémoire*

Proposer et choisir une méthode satisfaisante pour établir des correspondances entre racines observées et mesurées à plusieurs dates durant quatre semaines, en ayant seulement connaissance de la plante dont elles sont issues et des mesures réalisées, tel est l'objet de ce mémoire.

Dans le cadre d'une expérience de suivi de la croissance racinaire de 150 lignées d'orge (*Hordeum vulgare ssp spontaneum*), expérience réalisée dans un environnement contrôlé (dispositif de culture aeroponique) sur une durée de quatre semaines, nous obtenons dix sessions de mesure dans lesquelles chaque plante mesurée est identifiée. Cependant, les racines mesurées ne sont pas identifiées. Le seul critère de discrimination disponible est la mesure de la longueur de la racine.

Connaître, ou du moins, deviner, quelle est l'évolution de cette longueur pour une racine donnée offre certaines perspectives telles que l'étude de l'initiation de ces racines et des facteurs déterminant ce processus d'initiation. Une perspective à court terme, développée dans ce mémoire, est l'évaluation de la vitesse de croissance moyenne d'un ensemble de racines déterminées, c'est-à-dire connues ou devinées. Ce paramètre dynamique peut être soumis à une analyse génétique de recherche de Quantitative Trait Loci (QTL) ou être utilisé pour simuler la Root System Architecture (RSA).

Le problème qui nous concerne se situe donc au niveau de l'identification des racines au cours du temps. Le premier objectif de ce mémoire sera de proposer différentes voies de résolution du problème de correspondance. Au nombre de trois, ces voies diffèrent, en premier lieu, par les hypothèses qui les sous-tendent et, en second lieu, par les algorithmes de résolution implémentés.

Chacun de ces derniers sera testé sur l'ensemble des racines adventives mesurées durant l'expérience d'un nombre variable de plantes.

La pertinence des méthodes utilisées, des algorithmes implémentés et l'adéquation au problème des hypothèses posées détermineront le second objectif de ce mémoire, qui est le choix d'une voie de résolution parmi les différentes proposées.

Enfin, la méthode retenue sera testée sur l'ensemble de la population de plantes. Nous en obtiendrons les temps d'initiation et la vitesse de croissance moyenne des racines de l'ensemble des racines adventives des plantes de cette population. Le second paramètre, vitesse de croissance moyenne, fera l'objet d'une analyse génétique, réalisée par Jean-François Dumasy (BIR 23, Agronomie intégrée) et également présentée, entre autres résultats en septembre 2007.

La recherche présentée ici aura donc la structure suivante:

Premièrement, à travers une courte synthèse bibliographique qui fera l'objet du Chapitre II, nous esquisserons les connaissances actuelles en matière de morphogénèse racinaire, les techniques de modélisation inverse et les outils nécessaires à ces techniques, implémentés dans les recherches effectuées.

Deuxièmement, au Chapitre III, nous nous intéresserons à l'expérience sus-mentionnée, à sa conception et à sa mise en oeuvre, ainsi qu'aux dispositifs expérimentaux et résultats intrinsèques qui ont déterminé ces conception et mise en oeuvre. Cette partie couvrira l'acquisition de données brutes et le pré-traitement de celles-ci, effectué afin de faciliter leur utilisation dans la suite de la recherche.

Enfin, les chapitres IV à VI développeront les différentes voies de résolution, leur conception, leur mise en oeuvre et leur évaluation. Il importe de savoir que la chronologie de ces trois chapitres n'est pas fortuite. En effet, les méthodes présentées successivement dans chacun de ces chapitres s'articulent les unes aux autres. Ainsi, les conclusions de la première amèneront à la seconde et ainsi de suite.

Le chapitre VII présentera les conclusions générales de la recherche.

## Chapitre II

# Synthèse bibliographique

En raison de la localisation hétérogène des ressources nutritives dans le sol, l'architecture du système racinaire (RSA) déterminera le potentiel de cette plante à se développer et à profiter du milieu où elle vit. Cette configuration spatiale est donc un paramètre essentiel à considérer en amélioration des plantes. En plus de la morphologie racinaire, il est détaillé au premier point de ce chapitre.

L'écueil, en ce qui concerne la RSA, réside néanmoins en l'acquisition de données valides sur le système racinaire, acquisition difficile à réaliser sans destruction ni invasion de l'objet observé. Un modèle de simulation racinaire (*Root Typ*), développé au second point de ce chapitre, permet de simuler des données valides de RSA à partir de 21 paramètres d'entrée, dont le paramètre dynamique '*vitesse de croissance*'. La prédiction de ce paramètre est l'objet de ce mémoire.

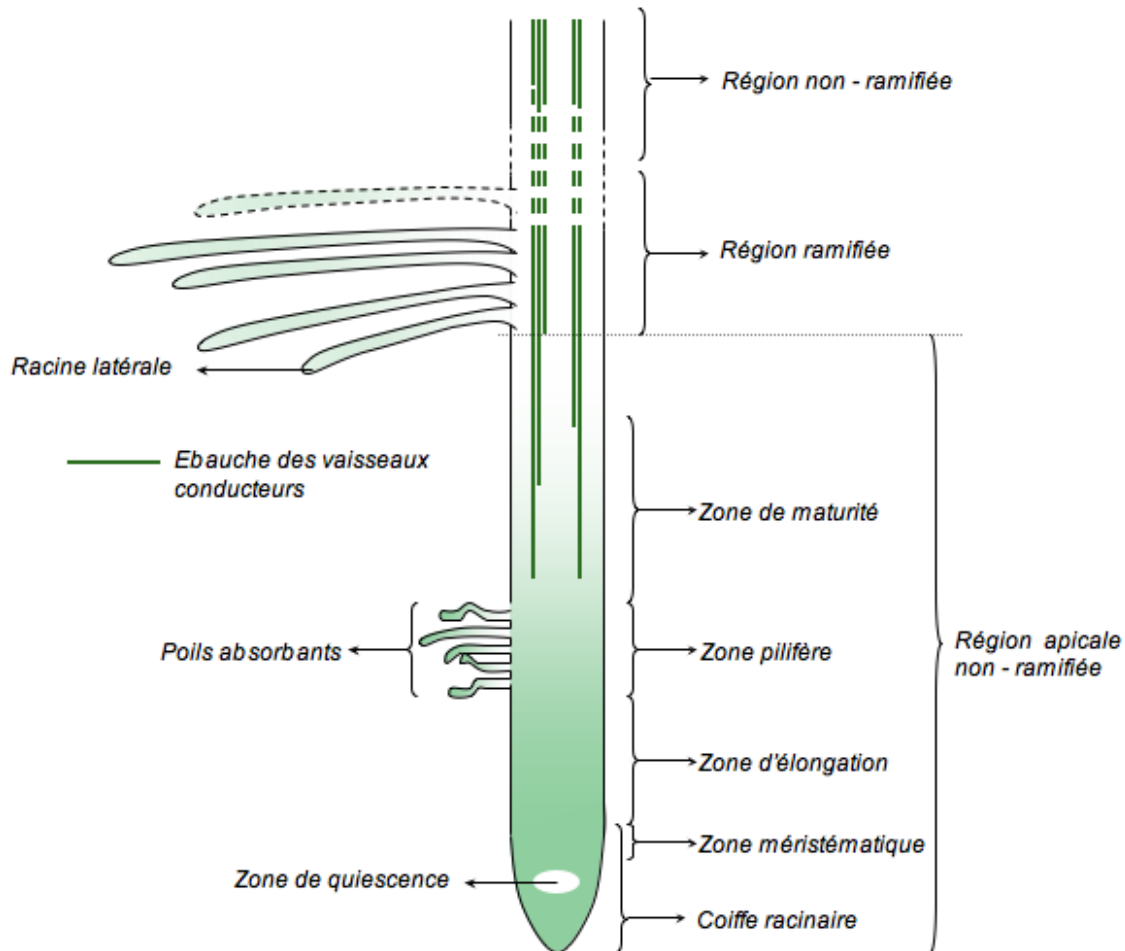
Les vitesses de croissance mesurées sont celles de racines de plants d'orge suivis durant quatre semaines de culture. Une esquisse des premiers stades de développement de cet plante est dressée au troisième point de ce chapitre

Cette prédiction se réalise par adaptation des données mesurées à une réalité attendue. Cette méthode est le *data fitting* (ou modélisation inverse). Les principes de cette méthode font l'objet du dernier point de ce chapitre.

## II. 1. Le système racinaire

### a. Physiologie du système racinaire

#### *Organisation d'une racine de Monocotylée*



**FIGURE II. 1.** : Schéma d'une racine de Monocotylée. D'après Draye (2005) et Prat et Rubinstein (2005).

Sur le schéma précédent, de bas en haut, on relève diverses zones constituant la racine d'une plante Monocotylée, distinguées sur base de leurs morphologies et de leurs fonctions

- La **coiffe racinaire** protège le méristème apical en empêchant les particules du milieu de heurter brutalement la racine en mouvement. Elle connaît un renouvellement constant de ces cellules internes, remplaçant les cellules externes progressivement desquamées; ce processus lui permet de maintenir une taille constante. Enfin, elle joue un rôle majeur dans l'orientation de la croissance de la racine en représentant le capteur de la gravité (Prat et Rubinstein, 2005).

- La **zone méristématique**, ou méristème apical de la racine, est le lieu de production des nouvelles cellules racinaires. C'est donc un lieu d'intenses mitoses. Les cellules les plus apicales

constituent la **zone de quiescence**, ainsi nommée en raison de leur activité de division quasi-nulle.

- La **zone d'élongation** est cette région non-ramifiée dans laquelle la croissance des cellules se réalise dans le sens principalement longitudinal. Cette croissance s'accompagne de phénomènes de différenciation (modifications du système vacuolaire) (Prat et Rubinstein, 2005).

- La **zone pilifère** est définie, comme son nom l'indique, par la présence de poils absorbants. Ces poils, produits nécessairement après élongation (Draye, 2005) réalisent l'essentiel de l'absorption de l'eau et des minéraux par les racines. Ils étendent la surface racinaire effective et, de ce fait, augmentent le contact sol-racine (Atkinson, 2000). Leur croissance est très rapide mais leur durée de vie est courte (Prat et Rubinstein, 2005).

- La **zone de maturité** du système vasculaire et de l'aérenchyme.

- La **région ramifiée** est bornée, en bas, par les racines latérales apparaissant ( voir ligne pointillée sur la figure II. 1.) et, en haut, par les racines latérales disparaissant. Cette région est, tout comme la précédente, également impliquée dans la maturation du système vasculaire, de l'aérenchyme, en plus de celle de l'épiderme et de la sénescence des poils absorbants.

- La **région non ramifiée** proche du collet est le lieu de sénescence des racines latérales. On n'y observe pas de néoformations de racines latérales (Draye, 2005). Cette affirmation n'est valable que pour les racines de Monocotylées;

On distingue, suivant leur origine, deux types de racines chez les Monocotylées, les racines séminales et les racines adventives. La description d'une racine faite ci-dessus s'applique, pour ces plantes, aussi bien à un type qu'à l'autre.

Le premier type regroupe les premières racines apparues, provenant de la graine germée. Morphologiquement, elles diffèrent du second type par leur diamètre plus faible.

Le second type regroupe les racines provenant de la tige de la plante Monocotylée. Elles apparaissent à la base de cette plante, au niveau des entre-noeuds. Leur émission s'arrête au moment de la floraison.

### *Le système racinaire*

**Comment définir le système racinaire?** Selon Pagès et Kervella (1990), le système racinaire est un ensemble d'axes, chaque axe étant généré par un méristème apical. Entre ces axes, une hiérarchie d'ordre est définie : les axes d'ordre 1 sont ceux qui sont directement connectés à la tige, ceux d'ordre deux sont connectés aux premiers et ainsi de suite.

Dans le cas des Monocotylées, le système racinaire est fasciculé; c'est-à-dire composé de nombreuses racines, de même diamètre, croissant en faisceau selon un gravitropisme positif . Il est

possible de distinguer plusieurs types de systèmes racinaires sur base de leur structure (diamètre, nombre de pôles xylémiens et durée de vie) et de leur évolution (de l'orthotropie à la plagiotropie); cette distinction a été établie par Kahn (1977). Cependant, dans le cadre de ce travail, la seule distinction pertinente concernant la composition du système racinaire portera sur l'origine des racines.

En fait de systèmes racinaires distincts, nous parlerons alors de :

- faisceau séminal, pour désigner l'ensemble des racines séminales présentes sur une plante à un moment donné de sa croissance (voir Figure II. 2);
- faisceau adventif, pour désigner l'ensemble des racines adventives (provenant des entre-nœuds) présentes sur la plante à un moment donné de sa croissance.

Cette dichotomie importe car ces deux types de racines ont un développement différent. Or, l'objet de ce travail, à travers la prédiction de vitesses de croissance, porte sur le développement racinaire.

**Comment se développe le système racinaire?** La dynamique de formation de la structure tridimensionnelle du système racinaire est régie par trois processus (Pagès et Kervella, 1990):

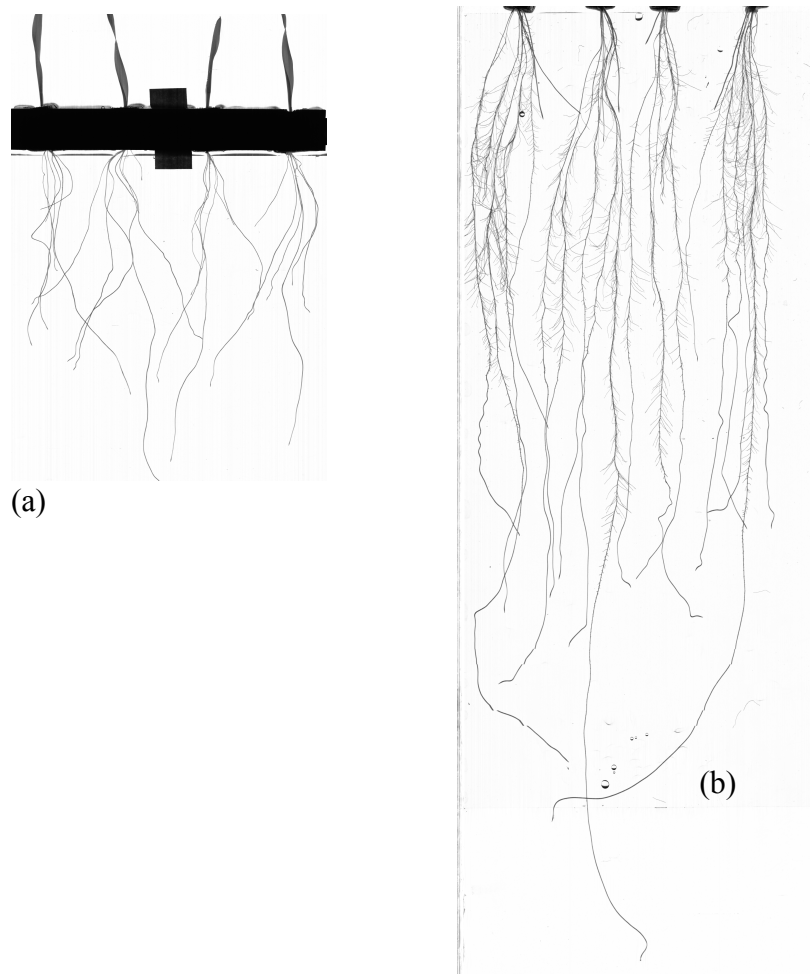
- l'émission des nouveaux axes racinaires est réalisée par la semence (racine séminale) ou par les entre-nœuds (racine adventive), cette deuxième possibilité étant liée notamment à la phyllotaxie ;
- la croissance en longueur, qui s'effectue dans une certaine direction et peut être nulle (mort du méristème) ou négative (nécrose) et la croissance en diamètre, cette seconde option ne nous concernant pas, dans le cas de Monocotylées ;
- la ramification, qui est le processus d'émission d'axes latéraux sur axes existants).

L'aspect dynamique du développement du système racinaire s'obtient par intégration sur une période donnée de discrétisation de ces différents processus.

Le système racinaire de jeunes plants d'orge (variété SCRI B83) est illustré par la Figure II. 2. regroupant deux photographies, à 7 jours d'intervalle, des systèmes racinaires des même plantes.

Sur cette figure, on remarque l'apparition de racines latérales sur les racines séminales, qui sont les uniques racines visibles en (a). On dénombre également plus de racines émises en (b) qu'en (a). Parmi celles-ci, on peut voir les premières racines adventives.

Ces photographies illustrent en outre deux problèmes : celui de la reconnaissance des racines (adventives ou séminales?) et celui de la mise en correspondance de racines à travers le temps.



**FIGURE II. 2.** : *Systèmes racinaires (plantes de variété SCRI B83) photographiés à 7 jours d'intervalle, le 17 juin 2004 (a) et le 24 juin 2004 (b).* Photo: Xavier Draye.

#### b. Paramètres d'intérêt

*« L'acquisition des ressources du sol par les systèmes racinaires de la plante est un sujet d'intérêt considérable en agriculture et en écologie. »* (Lynch, 1995).

Savoir quels sont les paramètres intéressants à observer dans le système racinaire implique de connaître les principaux rôles de ce système. On relève habituellement 4 fonctions majeures du système racinaire (Lejeune et Auquier, 2001).

Premièrement, le système racinaire assure l'ancrage de l'organisme végétal dans le sol ou tout autre environnement dans lequel il se développe.

Deuxièmement, l'absorption des nutriments et de l'eau nécessaire au métabolisme litho-autotrophique est réalisée au niveau des poils absorbants.

Troisièmement, le système racinaire est également impliqué dans la conduction de l'eau et des solutés vers la tige.



Quatrièmement, les tissus racinaires peuvent servir au stockage de substances riches en énergie.

Les trois premières fonctions sont directement influencées par l'architecture du système racinaire (RSA, *Root System Architecture*), laquelle réfère à sa configuration spatiale dans le sol, c'est-à-dire le déploiement géométrique explicite des axes racinaires (Lynch, 1995).

L'hétérogénéité de la distribution des ressources en eau et en nutriments dans le sol explique l'importance que dénote la configuration spatiale des racines. Cette RSA déterminera l'aptitude de la plante à se procurer les ressources du terrain et donc sa productivité. Manipuler la RSA en vue de maximiser cette aptitude permettrait de réduire l'effets des stress abiotiques sur le rendement des cultures d'intérêt (de Dorlodot *et al.*, 2007).

Lynch (1995) abonde en ce sens en remarquant que l'importance de la RSA dans la productivité des plantes est déterminée par le fait que de nombreuses ressources du sol sont irrégulièrement distribuées, ou localement manquantes, de telle sorte que le déploiement du système racinaire dans le sol déterminera, en grande part, l'aptitude d'une plante à exploiter ces ressources;

L'étude de l'architecture racinaire inclut la topologie (c'est-à-dire la façon dont la ramification s'opère entre axes racinaires) et la distribution racinaires (la présence de racines suivant un gradient positionnel) (Lynch, 1995). Donc, le nombre de racines dans le faisceau, leur longueur, la ramification des axes (approchée par la densité de ramification) et le moment d'initiation des axes sont des paramètres d'intérêts dans la connaissance de l'architecture racinaire et *a fortiori* dans l'introspection des processus et fonctions dont elle dépend et dans le déterminisme de ces processus et fonctions.

La connaissance de ces paramètres permet, en outre, de les intégrer dans l'élaboration de modèles de croissance racinaires. Ainsi, Tsegaye *et al.* (1995) développe une procédure d'acquisition de paramètres d'entrée, à partir de données expérimentales, pour le modèle ROOTMAP de prédiction de la croissance racinaire. Ce modèle comporte, entre autres paramètres, le nombre de racines, la densité de ramification le long d'un axe et le taux d'élongation racinaire.

Pagès et Kervella (1990), dans la description d'un modèle dynamique tri-dimensionnel de la structure du système racinaire, font intervenir trois processus pour faire évoluer cette structure simulée au cours du temps: l'émission de nouveaux axes racinaires à partir des tiges ou de la semence, la croissance et la ramification des axes existants (strictement acropète).

Enfin, la connaissance de ces paramètres permet également une étude du déterminisme génétique des processus qui les sous-tendent (Dumasy, 2007).

### c.Méthodes de mesure, d'observation et d'estimation des paramètres d'intérêt

#### *Mesure et observation*

Le point d'achoppement principal des méthodes de mesure et d'observation des paramètres sus-mentionnés demeure la nécessité de réaliser des mesures ni destructives ni invasives, c'est-à-dire des méthodes permettant des mesures répétées du même objet (mesure non-destructive) et n'interférant pas avec la physiologie de l'objet (mesure non-invasive).

On peut relever trois classes de méthodes d'observation du système racinaire, répondant à cette nécessité:

- 1) les méthodes basées sur des interfaces transparentes (rhizotrons, minirhizotrons);
- 2) les méthodes basées sur des systèmes de culture artificiels (systèmes hydroponiques, systèmes comprenant des plaques d'agar, boîtes de Pétri avec agar);
- 3) les méthodes d'imagerie par résonance magnétique ou la tomographie par rayon X (tomographie assistée par ordinateur).

La méthode que nous développerons relève de la seconde catégorie. Il s'agit d'un système de culture en aéroponie. Ce système et le dispositif mis en place seront développés dans le Chapitre III. Partie expérimentale.

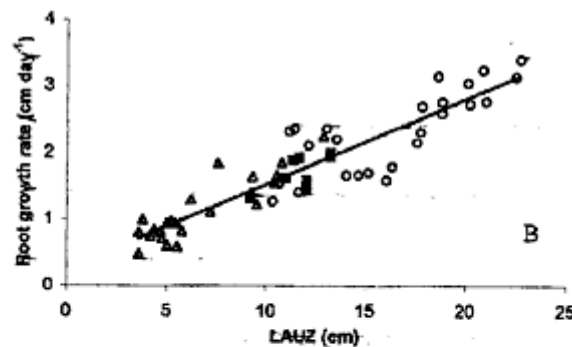
#### *Estimation de paramètres*

Il est possible d'étendre le champ d'investigation de paramètres racinaires au-delà de ces méthodes en effectuant, non pas une mesure d'un certain paramètre, mais bien une estimation de ce paramètre.

La vitesse de croissance d'un axe racinaire de premier ordre peut ainsi être estimée sur base de la longueur de la zone non-ramifiée (en anglais, Length of Apical Unbranched Zone, LAUZ). Comme le montre le graphique de la page suivante (Pagès *et al.*, 2001), il existe une corrélation entre ces deux paramètres. La pente de la droite de régression peut être interprétée comme le laps de temps entre l'initiation d'un primodium de racine latérale et son émergence sur l'axe d'ordre supérieur, tandis que l'intercept représente la distance entre le voisinage de l'apex et le point où le premier primodium est initié.

Cette corrélation (voir page suivante, Figure II. 3.) s'explique par le processus de production de nouvelles cellules au sein de la zone méristématique (voir Figure II. 1). Pellerin et Tabourel (1995) ont mis en évidence que l'apparition des racines latérales survient après un certain temps, peu sensible à l'environnement pour de nombreuses espèces (3-4 jours), durant lequel on n'observe pas de

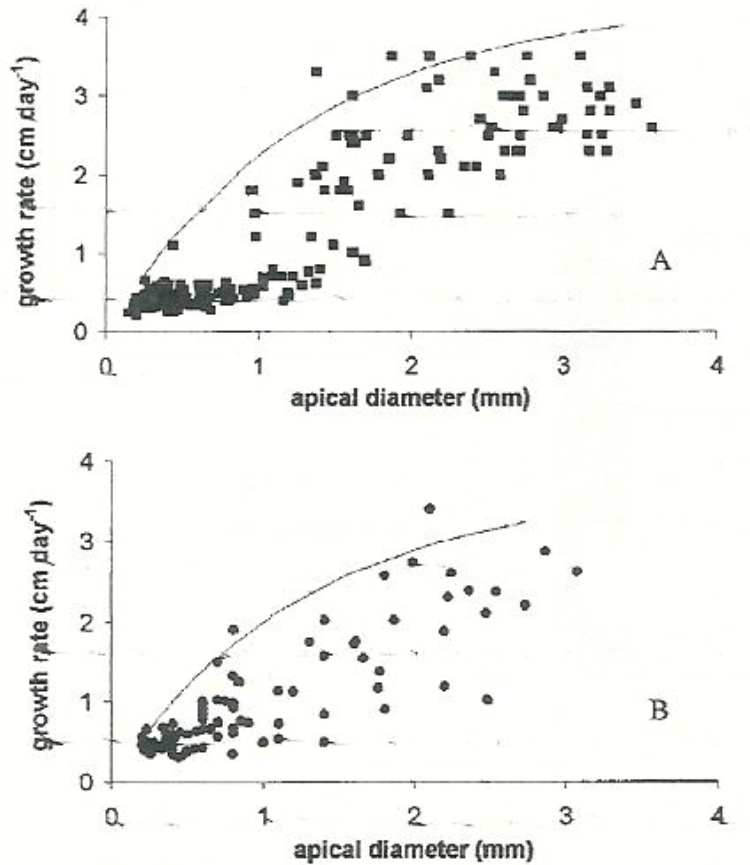
ramification à l'extrémité de la racine. La longueur de cette zone non-ramifiée (LAUZ), rapportée à ce laps de temps, indique donc la vitesse à laquelle la racine a progressé depuis que la zone méristématique a produit les cellules initiant les ramifications les plus proches de l'apex.



**FIGURE II. 3.:** Mise en évidence de la relation entre vitesse de croissance racinaire et longueur de la région apicale non-ramifiée, pour différents types de racines (primaire, latérale d'ordre 1 et adventives) cultivées en rhizotron. Illustration: Pagès *et al.*, 2001.

Pagès *et al.* (2001) en conclut que cette corrélation peut être un outil intéressant pour une étude complète du modèle de racine relié aux conditions du sol. Deux limitations s'imposent néanmoins à cette utilisation: la nécessité d'avoir des ramifications sur l'axe concerné (qui exclut donc les racines jeunes de moins de 4-5 jours) et la limitation de cette évaluation à une période de 3-4 jours précédents les mesures, car la relation porte sur la vitesse de croissance durant les jours écoulés depuis l'initiation du dernier primordium.

Dans la même publication, Pagès *et al.* (2001) met également en évidence une corrélation entre le diamètre apical d'un axe racinaire et sa vitesse de croissance (Figure II. 4.). Il observe ainsi que le diamètre apical semble indiquer une vitesse d'élongation potentielle tandis que la croissance finale est limitée par d'autres facteurs. Cependant, la vitesse peut varier d'un facteur 1 à 4 pour un même diamètre apical mesuré. Cette observation tend à confirmer l'hypothèse selon laquelle le diamètre apical donne une indication de la vitesse potentielle d'élongation de l'axe racinaire.



**FIGURE II. 4.:** *Mise en évidence de la relation entre diamètre apical et vitesse de croissance d'une racine. La vitesse de croissance est mesurée en rhizotron (A) ou estimée sur base de la LAUZ en champ (b). Illustration : Pagès et al., 2001.*

Ce premier sous-chapitre clarifie certaines notions (RSA, systèmes adventif et séminal, LAUZ, ...) et remarque l'intérêt agronomique de la connaissance de ces notions. Lesquelles sont utilisées et comment le sont-elles? Le second sous-chapitre, en développant des modèles de simulation racinaire, permettant d'utiliser au mieux l'information mesurée, répondra à ces questions.

## II. 2. Modèles de simulation de la RSA et de croissance racinaire

### a. Root Typ, modèle générique de simulation de la RSA

*Root Typ* est un modèle générique adapté pour la représentation du développement architectural de systèmes racinaires de différentes plantes croissant dans des sols variés (Pagès *et al.*, 2004).

Le modèle admet 21 paramètres se référant aux processus de croissance axiale (dont la vitesse de croissance initiale), de ramification, de réitération (forme alternative de ramification), de transition (changement de stade morphogénétique), de croissance radiale et de déperissement ainsi qu'au type de racine concerné.

Ce modèle travaille sur une plante individuelle et son système racinaire en interaction avec le sol, lequel peut réguler l'élongation racinaire, la direction de croissance, la densité de ramification. Le système racinaire s'étend et se ramifie par un ensemble d'apex racinaires de stades variés. Chacune de ces pointes étend son axe de développement (croissance axiale) et développe des axes latéraux (ramification).

La simulation racinaire s'effectue sur une durée fixée par l'utilisateur (pas de temps: 1 jour) et est codée en un fichier de noeuds, représentant à chaque temps des positions successives des apex et des points de ramifications. Ces noeuds contiennent de l'information concernant la localisation 3D, la position et la filiation dans l'architecture racinaire (axe d'origine, noeud précédent et noeud suivant) et plusieurs attributs structuraux (type de pointe racinaire, age, diamètre). L'incrémentation de ce fichier de noeuds se fait à chaque pas de temps.

#### *Croissance axiale*

Dans le modèle *Root Typ*, la croissance axiale est caractérisée par une valeur scalaire d'élongation et une valeur vectorielle de direction de croissance. Dans notre cas, nous nous en tiendrons au seul processus d'élongation.

Durant un pas de temps, l'élongation d'une racine est calculée à partir d'une valeur potentielle spécifique ( $GR_{pot}$ ), multipliée par un coefficient d'influence réductive du sol dépendant de la couche de sol dans laquelle la racine se trouve ( $GR_{soil}$ ). Le taux d'élongation potentielle est calculé, par pas de temps  $dt$ , avec l'équation (2. 1.):

$$GR_{pot} = GR_{init} \cdot e^{\frac{-GR_{mit}}{L_{max}} \cdot t} \cdot dt \quad (2. 1.)$$

dans laquelle  $t$  est le temps écoulé depuis l'émergence,  $L_{max}$  est la longueur maximale de la

racine  $GR_{init}$  le taux de croissance initial (c-à-d taux de croissance à  $t = 0$ ). Ce dernier taux est un paramètre défini pour chaque type de racine tandis que le paramètre  $L_{max}$  est tiré au hasard dans une distribution normale paramétrisée pour chaque type de racine (Pagès *et al.*, 2004).

### *Croissance radiale*

Dans Root Typ, la surface d'une section de racine à une position donnée procède de l'accumulation de sections élémentaires de racines située distalement par rapport à cette position. Cela signifie que chaque racine contribue à la croissance radiale de la racine sur laquelle est connectée. Cette contribution est basée sur la section propre de la racine et sur un coefficient de croissance radiale  $RI$  (unique pour l'ensemble du système racinaire). Ce dernier coefficient est proche de 1 pour les dicotylées et proche de 0 pour les monocotylées, car celles-ci font montre de peu ou pas du tout de croissance radiale (Pagès *et al.*, 2004).

Le modèle Root Typ est apte à simuler diverses RSA dans des conditions de culture ou des conditions plus ou moins naturelles. Il peut contribuer de ce fait à caractériser des différences génotypiques (Pagès *et al.*, 2004). Donc, à partir de 21 paramètres, on peut réaliser des inférences statistiques plus fiables que celles qui auraient été menées à partir d'un très grand nombre de mesures de longueur, de LAUZ ou de diamètre.

### b. GRAAL, modèle fonctionnel et structurel (Drouet *et al.*, 2007)

Le modèle GRAAL (pour GR*o*wth, Arch*it*ecture, AL*l*ocation) a été développé pour simuler et analyser les interactions entre des processus morphogénétiques et la répartition des assimilats durant le développement végétatif de plantes individuelles. Ce modèle associe d'autres modèles, de morphogénèse et de simulation de la croissance des parties de la plantes en fonction de la disponibilité en assimilats.

En ce qui concerne la croissance effective d'un organe donné, GRAAL propose une règle de proportionnalité de cette croissance effective à la croissance potentielle et à la disponibilité en assimilats. Cette disponibilité est définie par le ratio entre l'offre (carbone provenant de la graine ou produit par la photosynthèse) et la demande en assimilats. Ce est identique pour l'ensemble des organes à un moment donné de la simulation.

Dans ce modèle, l'élongation racinaire effective, établie en volume de racine, dépend de la disponibilité en carbone.

Dans la conception des algorithmes de mise en correspondance de racines faisant l'objet de ce mémoire, nous prendrons comme hypothèse, entre autres, la règle de proportionnalité croissance – disponibilité énoncée ci-dessus. En outre, nous supposerons que le volume racinaire est fortement corrélé à la longueur racinaire, en raison du diamètre constant des racines d'orge; cette seconde supposition permet de faire le lien également énoncé ci-dessus entre élancement racinaire et disponibilité en carbone.

## II. 3. L'orge

### a. Généralités.

L'orge de printemps (*Hordeum vulgare* L.) est une céréale annuelle, à inflorescence en épi barbu à 2 rangs. Elle est semée en champ dès la fin février et récoltée à surmaturité, lorsque la tige portant l'épi se courbe (*Source* : [www.inra.fr](http://www.inra.fr)).

L'orge a un système racinaire double: le système séminal (embryonnaire, originaire de la semence) et le système adventif (émergeant depuis les noeuds). Le premier système pénètre dans le sol jusqu'à une profondeur de 15 à 30 cm et demeure actif longtemps après l'initiation du système racinaire adventif. Ce dernier atteint des profondeurs de l'ordre de 60 à 100 cm, bien que la majorité des racines se trouvent dans les 25 premiers centimètres du sol. Le système racinaire total de l'orge représente 8,7% de la biomasse de la plante entière. Par ailleurs, la force de pénétration dans le sol de l'orge est moindre que n'importe quelle autre céréale, ce qui oblige une profondeur d'ensemencement de 6 à 7 cm maximum (*Source*: [agrifish.jrc.it](http://agrifish.jrc.it)). Une illustration de systèmes racinaires d'orge est disponible en Figure II. 2.

Céréale pauvre en protéines, elle entre dans la composition de nourriture pour bétail et est surtout mise à profit, en alimentation humaine, par le procédé de maltage, germination des graines suivie d'une torréfaction, qui est à la base de la production brassicole.

### b. Croissance jusqu'au stade 3 talles

La température minimale pour la germination est de 1°C à 3 °C. Après humidification de la graine, la radicule apparaît. Le coléoptile, première structure aérienne, apparaît ensuite.

Les radicules apparaissant en premier lieu sont des racines séminales. Chez l'orge, on dénombre jusqu'à 8 racines séminales, apparaissant toutes dans les deux premières semaines.

Les racines émises par les noeuds de la tiges sont les racines adventives.

La première feuille apparaît une fois que le coléoptile a fini sa croissance (jusqu'à 7,6 cm maximum). Une nouvelle feuille apparaît ensuite tout les 3-5 jours, en fonction des conditions et de la

variété (Anderson *et al.*, 2002). Le stade 2-feuilles est atteint ainsi 12 jours après l'émergence.

Le tallage débute 15 jours après l'émergence, c'est à dire lorsque la plante a environ trois feuilles. Un plus grand nombre de talle est formé lorsque les températures sont basses, la population petite et la teneur en azote élevée.

En ce qui concerne la croissance racinaire, May *et al.* (1965) observe que les axes racinaires d'ordres 1, 2 et 3 croissent tous linéairement (vitesse de croissance constante) dans un environnement homogène. La vitesse de croissance décroît significativement cependant de l'ordre 1 à l'ordre 3.

## II. 4. Modélisation inverse: principes et outils d'application

### a. Principes de modélisation inverse

La modélisation inverse est la réciproque de la modélisation prédictive (ou simulation). Elle consiste à reconstruire les conditions initiales et les paramètres du modèle à partir de données observées. Les buts de la modélisation inverse sont donc de nous renseigner sur les composantes souvent mal connues que sont les paramètres du modèle et d'exploiter avantageusement l'information contenue dans les observations disponibles (Vézina et Pahlow, 2003).

Tarantola (1987) divise la procédure scientifique pour l'étude d'un système physique en trois étapes:

- i) paramétrisation du système: découverte d'un set minimal de paramètres du modèle dont les valeurs décrivent complètement le système, d'un certain point de vue;
- ii) *Forward modeling* (modélisation directe): découverte des lois physiques autorisant, pour des valeurs données de paramètres du modèle, de réaliser des prédictions de mesure sur certains paramètres observables;
- iii) *inverse modeling* (modélisation inverse): utilisation des résultats finaux de certaines mesures des paramètres observables pour inférer les valeurs finales des paramètres du modèle.

Dans les sections suivantes, nous passerons en détail chaque étape.

### *1. Paramétrisation*

Les paramètres utilisés dans la modélisation sont ceux définis par l'utilisateur dans la relation de modélisation directe entre une variable dépendante  $y^*$  et un set de variables explicatives  $\mathbf{x}$  (Lambot *et al.*, 2004):

$$y^*(\mathbf{x}) = y^*(\mathbf{x}; \mathbf{p}) \quad (2. 2)$$

Dans cette relation, le vecteur  $\mathbf{p}$  est un jeu de  $m$  paramètres. Ce vecteur de paramètres



correspond au vecteur  $\mathbf{x}$  de variables explicatives.

## 2. Modélisation directe

La modélisation directe repose, comme sus-mentionné, sur la connaissance des lois physique inféodant les observations réalisées.

Dans le cas de correspondance à travers le temps de racines non-identifiées, le modèle direct supputé sera la croissance linéaire de l'ensemble des racines d'un faisceau, ce qui impose un coefficient de croissance, identiques pour chacune ou pas, toujours supérieur ou égal à 0.

Afin de pouvoir inverser le modèle, il incombe de respecter plusieurs conditions (Lambot *et al.*, 2004):

- Conditions d'existence de la solution pour le modèle direct : le modèle direct permet d'obtenir une solution.
- Condition d'unicité de la solution: la solution existante est unique, c'est-à-dire qu'un même jeu de valeurs pour les paramètres du modèle ne renvoie qu'une seule solution.
- Condition portant sur le caractère identifiable des paramètres: les paramètres sont tous identifiables. Si tel n'est pas le cas, deux jeux de valeurs de paramètres peuvent mener à une même solution; un paramètre supplémentaire, c'est-à-dire une variable dotée d'un pouvoir explicatif, n'aurait pas été inclus dans la paramétrisation du modèle.
- Condition portant sur la sensibilité du problème: la réponse du modèle est totalement dépendante des paramètres du modèles. En d'autres termes, la dérivation de cette réponse par rapport aux vecteur  $\mathbf{x}$ , en tout point du domaine de celui-ci, doit toujours être différente de 0, c'est-à-dire, si  $D \subseteq \mathbb{R}^m$  est le domaine de  $\mathbf{p}$ , alors :

$$\forall \mathbf{p} \in D, \quad \frac{\partial(y^*(\mathbf{x}; \mathbf{p}))}{\partial(\mathbf{x})} \neq 0 \quad (2. 3.)$$

## 3. Modélisation inverse

La modélisation inverse fait partie de la famille des techniques d'hybridation entre mesures et modèles dont le but commun est de définir au mieux l'état d'un système en utilisant le maximum d'informations (Pison, 2005).

En modélisation inverse, on parle de (Pison, 2005):

- Valeur de départ des paramètres, ou valeurs *a priori*, ce sont ces valeurs que la méthode

d'inversion optimise pour obtenir les valeurs optimisées.

- Dimension du problème inverse, ce terme représente le nombre de paramètres à inverser.
- Vecteur d'état, ce qui correspond à un groupe de valeur prises par chacun des paramètres à inverser.
- Contraintes, ce sont les valeurs observées que l'on veut faire approcher par le modèle.
- Opérateur d'observation, ce qui relie le vecteur d'état à l'espace des observations pour comparer les valeurs de contraintes aux prévisions. Ce terme correspond donc au modèle qui permet de calculer les valeurs simulées aux mêmes points que les observations.

### b. Outils d'applications

Cette seconde section détaille les bases théoriques de plusieurs outils utilisés dans la conception de méthodes de modélisation inverses et leur implémentation au problème nous concernant.

*Notions de calcul combinatoire* (Bogaert, 2003)

Dans la première méthode proposée (voir Ch. IV), le dénombrement des situations possibles repose sur un arrangement combinatoire. Deux notions théoriques sont à la base de cet arrangement, celles de permutation et d'arrangement sans répétitions.

On appelle nombre de permutations  $P_n$  le nombre de manières qu'il y a de ranger  $n$  objets distincts dans un certain ordre. Une permutation donnée est le résultat de la démarche:

- (1) on choisit parmi les  $n$  objets celui qui prendra la position n°1;
- (2) on choisit parmi les  $(n-1)$  objets restants celui qui prendra la position n°2;
- (3) on continue ainsi de suite jusqu'à ce que tous les objets aient été placés.

La formule de récurrence  $P_n = n.P_{n-1}$  peut s'écrire

$$P_n = n! \quad (2. 4.)$$

La notion d'arrangement sans répétitions procède d'une démarche de récursion analogue à la précédente. En effet, dans la démarche ci-dessus, la récurrence se poursuit non pas jusqu'à ce que tous les objets aient été placés, mais jusqu'à ce que  $k$  objets, parmi  $n$  objets distincts, soient placés.

Le nombre d'arrangements sans répétitions  $A_n^k$  est donc le nombre de manières qu'il y a de ranger dans un certain ordre  $k$  objets pris parmi  $n$  objets distincts, pour autant que le tirage s'effectue sans remise (un objet ne peut être choisi qu'une seule fois).

Le paramètre  $k$  a un condition limite, il doit être inférieur ou égal à  $n$ . Dans le second cas ( $k=n$ ), le nombre d'arrangements sans répétition égale le nombre de permutations de  $n$  objets.

*Algorithme basé sur la méthode simplex de Nelder et Mead.*

Le logiciel de traitement de données Matlab propose une fonction d'optimisation d'un vecteur de paramètres d'entrées d'une fonction objectif. Cette fonction, nommée *fminsearch*, procède par la recherche d'un minimum local de la fonction objectif en utilisant la méthode simplex de Nelder et Mead.

Cette méthode (Lagarias *et al.*, 1998) tente de minimiser une fonction non-linéaire de  $n$  variables réelles sans procédés de dérivation (implicite ou explicite). La méthode Nelder-Mead relève de la classe des méthodes de recherche directe, dont une sous-classe, incluant cette méthode, maintient à chaque pas un simplexe non-dégénéré, c'est-à-dire une figure géométrique en  $n$  dimensions, de volume non-nul, qui correspond au polyèdre convexe de  $n$  dimensions, donc à  $n+1$  sommets.

Computationnée, la méthode Nelder-Mead définit un algorithme itératif dont la propriété de convergence sur le minimum de la fonction à optimiser a été démontrée si cette fonction est de dimension 1 (Lagarias *et al.*, 1998).

L'utilisation de la fonction *fminsearch* sera détaillé dans le Chapitre VI., lors de son emploi.

# *Chapitre III*

## *Partie expérimentale*

Ce troisième chapitre développe l'ensemble des expériences menées durant l'année académique 2006-2007 conjointement avec Jean-François Dumasy, ainsi que la méthode de pré-traitement des données brutes et l'algorithme qui est conçu afin de pourvoir à cette tâche.

Cette partie expérimentale regroupe donc l'ensemble des étapes qui ont permis de générer les données d'observation et de mesures racinaires pouvant être traitées par les algorithmes implémentés dans les trois chapitres suivants (Chapitres IV., V. et VI.).

Ce chapitre comprend cinq parties. La première s'intéresse à la culture aéroponique, la seconde au matériel végétal utilisé pour nos expériences, la troisième aux conditions de culture communes à toutes les expériences et les deux dernières aux expériences proprement dites. La quatrième partie détaille les expériences dites préliminaires, effectuées dans le courant du premier quadrimestre de cette année académique, ayant pour objectif la réalisation, au second quadrimestre, de l'expérience principale, laquelle est détaillée dans la cinquième partie.

### III.1. Principes de culture aéroponique

#### a. Principes de fonctionnement

Dans un système de culture aéroponique, un support maintient la plante cultivée au-dessus d'un atomiseur qui envoie de façon périodique de la solution nutritive. L'énergie nécessaire à cette brumisation est fournie par une pompe et son atomisation est rendue possible par la présence du brumisateur. Schématiquement, un système de culture aéroponique se présente comme ceci (Figure III. 1.).

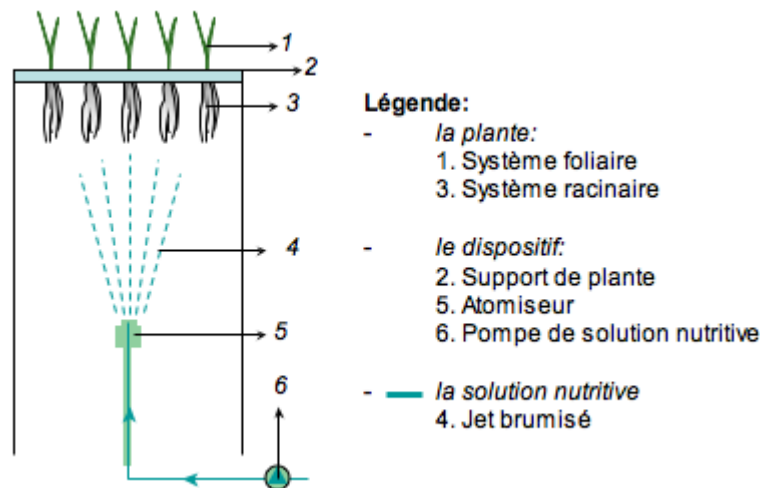


FIGURE III. 1.: Schéma d'un système de culture aéroponique.

Ce système présente des avantages certains pour l'observation et la mesure de racines. En effet, dans un tel dispositif, il ne faut pas manipuler les plantes, ou du moins très peu comparativement à d'autres systèmes de culture (hydroponie, rhizotron). En outre, en aéroponie, très peu de noeuds se forment entre racines, ce qui est plus fréquent en hydroponie.

En dehors des périodes de brumisation, aucune contrainte physique n'est appliquée à la surface des racines. Cela rend la culture fortement différente d'une culture menée en champ car soumise à une contrainte physique de façon non permanente, mais permet néanmoins l'observation et la mesure des systèmes racinaires.

En somme, observer et mesurer la croissance racinaire dans un tel dispositif n'est ni destructif (matériel végétal non détruit par l'intervention expérimentale), ni invasif (l'intervention expérimentale n'interfère pas physiquement avec l'objet d'expérience) (de Dorlodot *et al.*, 2005).

#### b. Contrôle et réglages du dispositif

Plusieurs degrés de liberté (ou points de contrôle), que l'on peut faire varier pour régler les conditions de culture dans le dispositif, peuvent être relevés.

Premièrement, les points de contrôles relatifs à l'aspersion. Le **type d'asperseur** est un

premier paramètre; il définira en effet le diamètre des gouttelettes brumisées. Le type de **pompe** utilisée, le rythme et la durée de la **période d'aspersion** en sont d'autres; ils seront déterminant dans l'approvisionnement de la plante en éléments nutritifs. Le **placement de l'asperseur** par rapport au support de la plante peut interférer avec la croissance du système racinaire. Enfin, **la pression de la solution nutritive** à la sortie de la pompe déterminera sa capacité à atteindre l'ensemble des racines.

Ensuite, les points de contrôle relatifs au matériel inerte (support, pompe et tuyauterie), qu'il s'agisse du **matériau** (bois, PVC, ...) et de **l'agencement** de ce matériel. Ce second point mérite explication. Il désigne le placement des uns par rapport aux autres de l'ensemble des segments du circuits de solution nutritive. L'insertion, par exemple, dans le circuit de pompage d'un élément de tuyauterie tel qu'un clapet anti-retour peut occasionner une perte de charge. D'autre part, il convient de ne pas mettre en contact avec la solution nutritive ou le matériel végétal des matériaux pouvant interférer avec ces derniers (matériel corrodé par la solution ou matériel toxique pour les plantes).

S'en suit enfin la **composition de la solution nutritive**, qui constitue également un point de contrôle du dispositif.

#### c. Conception d'un dispositif aéroponique.

Nous avons travaillé dans trois types de dispositifs aéroponiques, toujours situés en environnement contrôlé (Serres PATO et ECAV de la Faculté AGRO de l'UCL). Les deux premiers types furent préalablement conçus et construits respectivement par M. T. Lavigne et par Mme. S. de Dorlodot. Le troisième type a été est une réalisation commune, conçue et construite par MM. Draye, Dumasy, Lavigne et moi-même. Cette réalisation se base sur les acquis des précédents dispositifs cités ci-dessus.

Sans rentrer dans le détail de ce dernier dispositif (développé ultérieurement, section III. 4. c.), il convient cependant de citer quelques caractéristiques communes à tous les dispositifs aéroponiques utilisés.

En premier lieu, une pompe telle qu'utilisée est constituée d'un tuyau en PVC rigide, d'une arrivée et d'une sortie de solution nutritive et d'une arrivée d'air comprimé. Périodiquement, la pompe effectue deux mouvements. Le premier est le remplissage du tuyau par la solution nutritive et le second est la détente d'un volume d'air comprimé dans le tuyau, chassant la solution nutritive vers la sortie. Collectant la solution nutritive préalablement aspersée, l'arrivée de la solution est munie d'un clapet anti-retour afin d'empêcher son retour par ce circuit. Quant à elle, la sortie conduit aux asperseurs placés en parallèle. L'élément moteur de la pompe est donc la détente de l'air comprimé dans le tuyau ; la pression à l'entrée de ce dernier déterminera donc la pression de la solution nutritive à la sortie de la pompe.

En second lieu, les plantes sont maintenues au-dessus des asperseurs au moyens de bandes de polystyrène extrudé (panneaux Floormate), larges d'environ 7 cm, hautes de 3 cm et de longueur variable. Ces bandes (voir Figure III. 3.) sont trouées en différents points régulièrement espacés sur la longueur pour permettre l'insertion d'embouts creux de polyuréthane dans lesquels sont placés les plantules germées. Un papier filtre permet de fixer ces dernières dans les embouts creux de polyuréthane.

### III.2 . Matériel végétal

Les variétés (génotypes) différentes utilisées sont au nombre de 151.

156 croisements ont été réalisés par le Scottish Crop Research Institute (SCRI) entre la variété Derkado (phénotype *prostré nain*) et la lignée B83-12/21/5 du SCRI (phénotype *érigé semi-nain*). Sur des individus de la F1 de ce croisement, des gamètes ont été prélevés et diploïdisés afin d'obtenir des individus F2 totalement homozygotes. Cette population de ségrégation a été créée, génotypée et phénotypée (sur le caractère nain et la résistance à un mildiou) en vue d'étudier les associations phénotype-génotype pour les caractères de rendement et de tolérance aux stress salins (Ellis et *al.*, 2002).

De cette population, nous utilisons, dans nos expériences, 149 variétés.

Outre celles-ci, les génotypes parentaux (ci-après dénommés DERK et B83) entrent également dans le cadre de nos expérimentations.

Pour disposer d'un nombre suffisant de graines, nous avons préalablement multiplié ces lignées. Elles ont été semées en serre-tunnel sur des parcelles de 100 cm \* 40 cm en mai 2006, sont arrivées à maturité vers le 20 juillet 2006 et ont été récoltées dans la semaine du 24 au 28 juillet 2006.

Nous ferons germer, dans nos expériences, des graines provenant de cette récolte.

### III. 3. Conditions de culture

#### a. Environnement contrôlé

L'ensemble des manipulations et expériences ont été menées dans les serres de phytotechnie tempérée (serres ECAV) et de phytopathologie (serres PATO) de notre Faculté. Les conditions de température, d'humidité et de luminosité ont été réglées sur un climat tempéré et maintenues à ce climat, en admettant des variations journalières de lumière et de température. Ces deux paramètres ont en outre été relevés durant l'expérience principale et sont présentés en Figure III.12 et Figure III. 11.

En outre, afin d'éviter les contaminations pathogènes, les composants du système de culture susceptibles de souiller la solution nutritive ont été désinfectés.

Nous traitons les pompes, tuyaux et bacs aéroniques par circulation pendant ½ heure à 1 heure (en fonction de la taille du dispositif) d'une solution 0,05 M d'ammoniaque. Ce traitement est suivi par plusieurs rinçages successifs à l'eau déminéralisée (de Dorlodot, 2005).

Quant au matériel de surface tels que les plaques et bacs de germination, les bandes de repiquage, les asperseurs plastiques et les pots de manutention, il est traité par une solution d'éthanol 95%.

Les embouts de polyuréthane sont désinfectés différemment: trempage durant ½ heure dans une solution de NaOH et durant ½ heure dans une solution d'éthanol 95%.

Nous assainissons les asperseurs inoxydables avec un acide fort, suivi d'un traitement avec une base forte, afin de dissoudre les sels précipités dans ce matériel.

Enfin, les pots contenant les solutions-mères (voir section IV. 3. b.) sont placés durant 1 heure dans l'étuve à 100 °C.

#### b. Solution nutritive

Nous avons utilisé, au cours des différentes expériences menées, la solution nutritive de Hewitt. L'utilisation de cette solution nous a été recommandée par le SCRI (Scottish Crop Research Institute), qui a développé les lignées d'orge.

La préparation de la solution nutritive est reprise en Annexe III. 1. Elle comprend deux phases. La première est l'élaboration de six solutions-mères et la seconde est la dilution de six volumes identiques de ces six solutions-mères dans un volume de 20 litres d'eau déminéralisée, auquel on a préalablement ajouté une quantité donnée de silicium sous la forme de métasilicate de sodium.

Deux concentrations ont été d'usage dans les expériences.

Pour la concentration *full*, on dilue un volume de 20 ml de chaque solution-mère dans 20 litres d'eau déminéralisée auquel volume a été ajouté 2,26 grammes de métasilicate de sodium. Pour la concentration *half*, on utilise la moitié des quantités de solutions-mères et de métasilicate pour un volume de 20 litres d'eau.

#### c. Germination

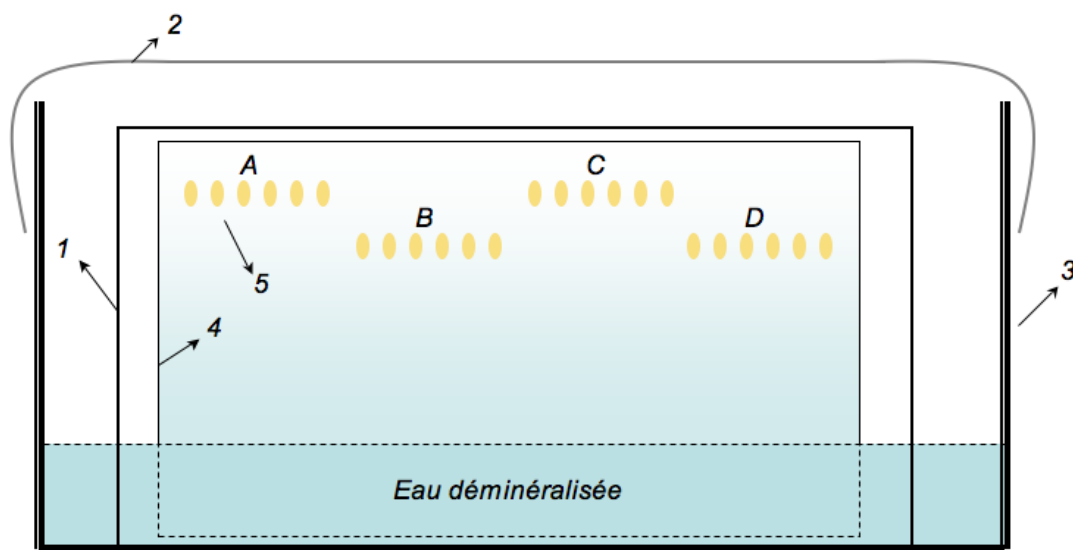
La germination s'effectue sur un papier buvard posé sur une plaque de verre disposée verticalement et partiellement immergée dans de l'eau déminéralisée (voir Figure III. 2., page suivante). Afin de maintenir le grain d'orge correctement polarisé sur le papier buvard au-dessus du



niveau de l'eau, une goutte d'agarose à 4% en masse, chauffée à 70 °C, permet de coller le grain dans une orientation correcte. Ce dispositif de germination est maintenu dans l'obscurité durant 3 jours. Après ces trois jours, on observe l'apparition de 2 ou plusieurs racines et, dans la majorité des cas, du coléoptile. Les plantules sont alors considérées comme prêtes pour le repiquage dans un bac aéroponique.

Afin de prévenir toute contamination pathogène, les grains d'orge font l'objet d'une désinfection à l'éthanol préalablement à l'étape de germination.

La figure suivante schématise une plaque de germination contenant 24 graines de 4 génotypes différents.



**Légende:**

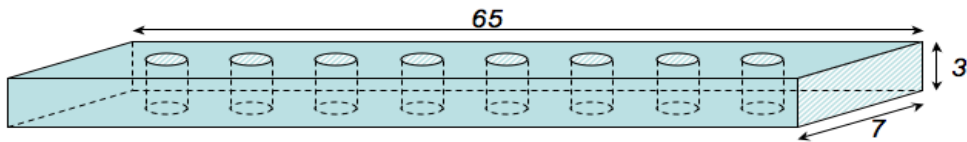
- |                    |                                 |
|--------------------|---------------------------------|
| 1. Plaque de verre | 4. Papier buvard                |
| 2. Bâche noire     | 5. Grains d'orge                |
| 3. Bac à eau       | A, B, C, D génotypes différents |

**FIGURE III. 2. :** Schéma d'une plaque de germination de 24 graines de 4 génotypes différents A, B, C et D. Dans ce dispositif, on n'observe aucun effet de la hauteur d'une graine sur le pouvoir germinatif.

**d. Repiquage**

Les plantules sont repiquées dans des embouts creux de polyuréthane et y sont maintenues au moyen d'un morceau de papier filtre. Ces embouts sont alors placés dans une bande de polystyrène, elle-même disposée au-dessus du bac aéroponique. L'espacement entre plantes est de l'ordre de 7 cm à chaque intervalle.

La Figure III. 3. détaille une bande de repiquage telle qu'utilisée pour l'expérience principale. La Figure III. 4. (page suivante) montre les plantules 2 jours après le repiquage, également lors de l'expérience principale.



**FIGURE III. 3.** : Schéma d'une bande de repiquage telle qu'employée pour l'expérience principale (proportions non respectées, mesures en cm).



**FIGURE III. 4.:** Photographie des plantules dans le bac aéroponique deux jours après le repiquage (expérience principale). Photo: M. Urban.

### III. 4. Expériences préliminaires

#### a. Expérience 1: Test code-barres

##### *Objectifs*

Le but de cette première expérience est d'évaluer une méthode de mesure des racines se basant sur la lecture d'informations code-barres, afin de déterminer quelle méthode de mesure sera retenue pour l'expérience principale. En outre, elle nous permet de nous familiariser avec un dispositif de culture aéroponique.

##### *Matériel*

2 bacs aéroponiques, de dimensions 80 cm (longueur) \* 40 cm (largeur) \* 82 cm (hauteur), solution nutritive HEWITT *half concentration* (20 litres, renouvelée par 5 jours), 2 bars de pression, 6 asperseurs plastiques GARDENAprac bac, durée de la période d'aspersion : 15 secondes par minute.

##### *Principes et méthodes de mesure au code-barres*

### (1) Stockage de l'information

Les codes-barres stockent une information ponctuelle sous forme d'une série de barres. Un lecteur (ou scanner) code-barres effectue une lecture automatisée de cette information par reconnaissance optique. Cette automatisation est mise à profit pour nos expériences. En effet, l'information contenue peut être de plusieurs types. Nous en utilisons quatre: des longueurs, des nombres, des états observés et des identificateurs. Par ailleurs, l'information peut être de nature différente en fonction de l'objet observé ou mesuré.

Nos codes-barres sont tous composés de deux parties. La première, appelée '*étiquette*', réfère à la nature de la mesure donc, à l'objet mesuré. Il peut s'agir d'une racine (L), d'un talle (T), d'un identifiant (ID), ... La seconde partie, appelée '*donnée*', est une valeur mesurée d'un certain type. Il peut s'agir d'une longueur racinaire (0015 mm, par exemple), d'un nombre de talles (stade 3 talles, par exemple), d'un état de santé d'une racine (état 3 sur une échelle de 1 à 5) ou d'un identificateur de plante (plante numéro 0514, par exemple). Le tableau suivant reprend l'ensemble des sigles code-barres utilisés et leur signification.

TABLEAU III. 1.: *Sigles des étiquettes et types de données mesurées.*

<b>Étiquette</b>	<b>Type de donnée</b>	<b>Information correspondante</b>
F	Etat (1 à 5)	Etat pathologique des feuilles. De 1 à 5 en fonction de l'altération visible de la feuille (taches jaunes et brunes, chloroses, ...).
R	Etat (1 à 5)	Etat pathologique des racines. De 1 à 5 en fonction de l'altération visible de la racine (importance des zones tachées de noir).
T	Nombre (1 à 3)	Nombre de talles sur le maître-brin
L	Longueur (0 à 1000 mm)	Longueur d'une racine ou de la zone apicale non-ramifiée d'une racine. De 0 à 1000 millimètres.
H	Longueur (0 à 1000 mm)	Hauteur d'une partie aérienne. De 0 à 1000 millimètres.
ID	Identificateur (0 à 700)	Numéro d'une plante. De la plante numéro 1 à la plante numéro 700.

Les codes-barres d'étiquette F, R et T sont imprimés sur une feuille disponible à l'expérimentateur. Les codes barres d'étiquette ID sont imprimés en languettes disposées de part et d'autre de la plante identifiée (visibles sur la Figure III. 4.). Enfin, les codes-barres de type *longueur* (étiquette H et L) font l'objet d'une disposition particulière : une règle de mesure graduée en codes-barres, expliquée au paragraphe suivant.

Une règle de mesure peut être conçue comme étant une suite d'objets (les graduations, en

millimètres par exemple) rangés régulièrement dans un certain ordre. Chacun de ces objets ont une valeur (0, 5, 10, ... millimètres) et une nature en fonction de l'objet mesuré (une longueur ou une hauteur). Cette double information est stockée en un code-barres. En outre, si l'on dispose de tels codes-barres sur une règle aux positions définies par les valeurs de ces codes-barres, on obtient une règle de mesure automatisée. L'information lue à une certaine position est révélatrice de la distance de cette position à un niveau de référence.

Afin de mesurer des longueurs racinaires, nous développons une telle règle code-barres, graduée de 0 à 1000 mm par pas de 5 mm, qui, une fois posée verticalement sur un support de référence, nous permettra, par un simple '*clic*' sur le lecteur code-barres, d'enregistrer la mesure de la longueur de la racine.

### (2) Méthode de mesure

Les observations et mesures sont organisées en une séquence chronologique de codes-barres. Cette séquence peut être divisée en sous-séquences, chacune relative à une plante.

Toute observation ou mesure portant sur une plante commence toujours par l'identification de cette plante. Ensuite, les mesures et les observations de différents types se succèdent dans la sous-séquence. L'ordre de cette succession est particulièrement important pour pouvoir traiter la sous-séquence par après. Un protocole de succession de mesure est donc défini par l'expérimentateur.

La séquence se poursuit de la sorte avec d'autres identificateurs, donc d'autres sous-séquences, portant soit sur la même plante, soit sur une autre.

Une séquence typique est donc telle que la séquence en gras ci-dessous.

**ID0001** : Identification de la plante numéro 1.

**L0045** : Mesure d'une racine longue de 45 mm sur la plante n°1.

**H0130** : Mesure d'une feuille haute de 130 mm sur la plante n°1.

...

**ID0002** : Identification de la plante n° 2.

**L0150** : Mesure d'une racine longue de 150 mm sur la plante n°2.

...

Cette séquence est finalement enregistrée dans un tableur Excel®.

### (3) Lecture du fichier de donnée et conversion des données

Plusieurs méthodes peuvent être conçues pour lire le fichier de donnée et le convertir en un tableau, plus facile de lecture qu'une séquence chronologique. Ces méthodes doivent bien évidemment intégrer le protocole de succession de mesure défini par l'expérimentateur pour élaborer la séquence d'observation et de mesure.

Nous avons développé deux méthodes. Toutes deux se basent sur la valeur d'étiquette (c-à-d la

nature) d'un code-barres. Celle que nous avons finalement retenue est présentée à la section III. 5. e. *Pré-traitement de l'information.*



**FIGURE III. 5.:** *Prise de mesure avec un lecteur et une règle code-barres.* Photo: M. Urban.

La photographie (Figure III. 5.) ci-dessus illustre la méthode présentée ci-précédemment. On y voit la règle code-barres placée à un niveau de référence (sur un rail métallique) et on distingue, en arrière-plan, les racines faisant l'objet de mesures.

#### *Matériel végétal*

6\*7 plants d'orge de variété B83 sont mis en culture dans le premier bac et 6\*7 plants d'orge, génotypes tirés au hasard dans le second bac.

#### *Protocole de mesure et d'observation*

Le protocole comporte différentes étapes de manipulation.

- Une bande de 7 plantes est retirée du bac aéronomique et suspendue sur un support de telle sorte que l'ensemble du système racinaire soit visible;
- la bande est maintenue durant un maximum de 15 minutes sur ce support;
- les mesures nécessaires sont effectuées au moyen d'une règle code-barres;
- la bande est replacée dans le bac aéronomique.

Le support utilisé est soit un cadre métallique adapté à la bande de façon à ce que les racines soient disposées verticalement à l'air libre, soit un bac à eau dans lequel les racines sont disposées horizontalement.

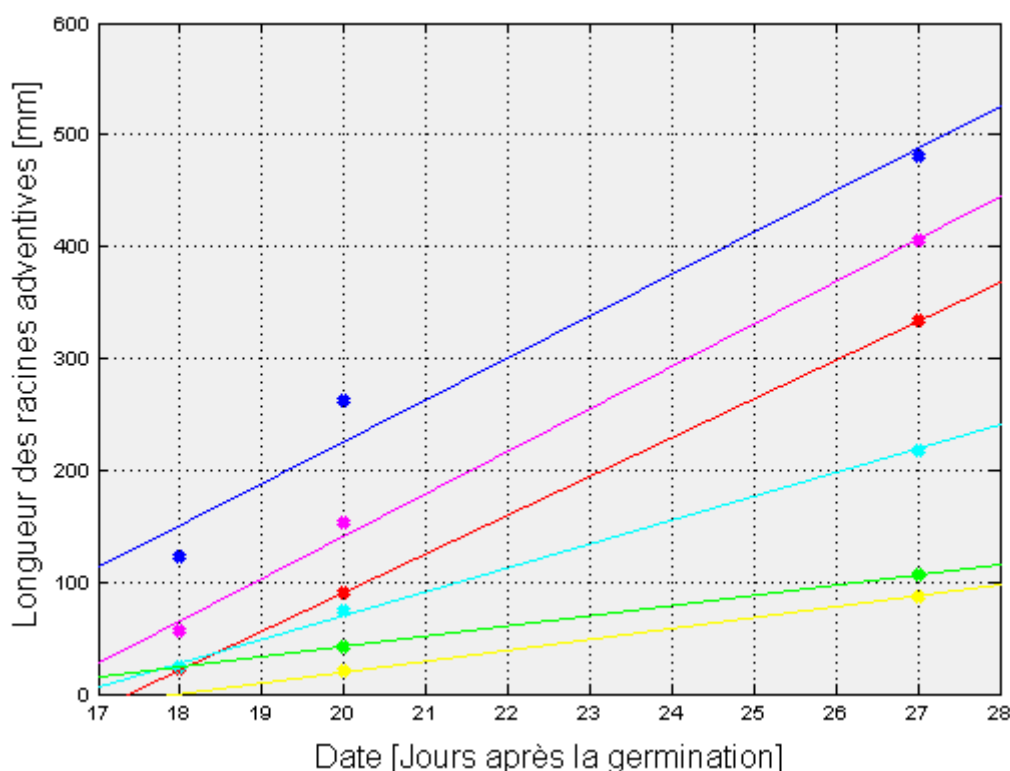
## Résultats

On constate sur les feuilles une carence minérale (probablement magnésienne), affectant l'ensemble de la population mise en culture, sans toutefois compromettre la croissance des plantes.

Le support 'bac à eau' ne permet pas de faire des manipulations propres. En effet, les racines s'entremêlent irrémédiablement dans le bac, obligeant ainsi l'expérimentateur de le toucher pour les désentrelacer. Un tel contact peut conduire à la mort de la racine s'il est effectué sur l'apex.

Le support 'cadre métallique' offre la possibilité de mesurer aisément les longueurs racinaires, les racines se disposant naturellement bien sous l'effet de la gravité. Cependant, il est difficile d'éviter un dessèchement durant la prise de mesure. Un dispositif dans lequel le système racinaire resterait en contact avec un environnement humide est vraiment préférable. Néanmoins, la prise de mesure avec le lecteur code-barres est aisée, rapide et offre une précision suffisante.

Enfin, cette expérience nous apprend à différencier les types de racines (adventives et séminales) sur base de leur morphologie et de leur moment d'apparition. Nous mesurons également la vitesse de croissance de racines adventives (moyenne sur plusieurs plante pour chaque adventive classée en fonction de son âge). Ces vitesses varient de 38,01 mm/jour à 9,64 mm/jour en fonction de l'âge de la racine, comme le montre le schéma ci-dessous.



**FIGURE III. 6.:** Evolution de la longueur de racines adventives (moyenne sur plusieurs plantes). Pour chaque plante, les racines sont classées suivant leur âge et la moyenne est calculée pour chaque classe d'âge. Sur ce schéma, on retrouve la première adventive (en bleu), la seconde adventive (en rose), la troisième adventive (en rouge), la quatrième adventive (en cyan), la cinquième adventive (en vert) et la sixième adventive (en jaune).

## b. Expérience 2: Test traitement par imagerie

### *Objectifs*

Cette seconde expérience évalue la prise de mesure au moyen de traitement d'images numérisées. Cette méthode sera comparée à la précédente pour choisir laquelle sera mise en oeuvre et utilisée lors de l'expérience principale.

Cette expérience a un second objectif : valider la relation vitesse de croissance – LAUZ (*Length of Apical Unbranched Zone*) développée dans la synthèse bibliographique (Chapitre II. 1. c.).

### *Matériel*

Bac aéronique 120 cm (longueur) \* 80 cm (largeur) \* 100 cm (hauteur), solution nutritive HEWITT-half et HEWITT-full (40 litres, renouvelée en fonction de la croissance observée), 4,2 bars de pression, 12 asperseurs inoxydables, durée de la période d'aspersion: 15 secondes par minute.

Le diamètre des gouttelettes brumisées avec les asperseurs inoxydables est inférieur à celui des gouttelettes obtenues avec les asperseurs GARDENA.

### *Méthode de mesure par traitement d'images numérisées*

Un scanner modifié, muni d'un bac à eau, capture une image à haute résolution d'un système racinaire placé horizontalement dans ce bac. L'image obtenue est ensuite traitée au moyen du plug-in SmartRoot (conçu et développé par X. Draye, non publié) développé pour le logiciel ImageJ (disponible sous licence freeware à [rsb.info.nih.gov/ij](http://rsb.info.nih.gov/ij)). Avec ce plug-in, nous mesurons les longueurs de racines et les longueurs de la zone apicale non-ramifiée.

### *Matériel végétal*

63 plants de 6 géotypes différents et des deux lignées parents sont mis ensemble en culture.

### *Protocole de mesure*

- La plante est extraite du bac aéronique et son système racinaire placé dans le bac à eau du scanner;
- le système racinaire est numérisé à haute résolution et l'image est enregistré sur un ordinateur (la Figure IV. 1. présente un fragment d'une telle image);
- l'image est chargée dans ImageJ et les longueurs sont mesurées par SmartRoot;
- les longueurs mesurées sont identifiées et présentées dans un tableau.

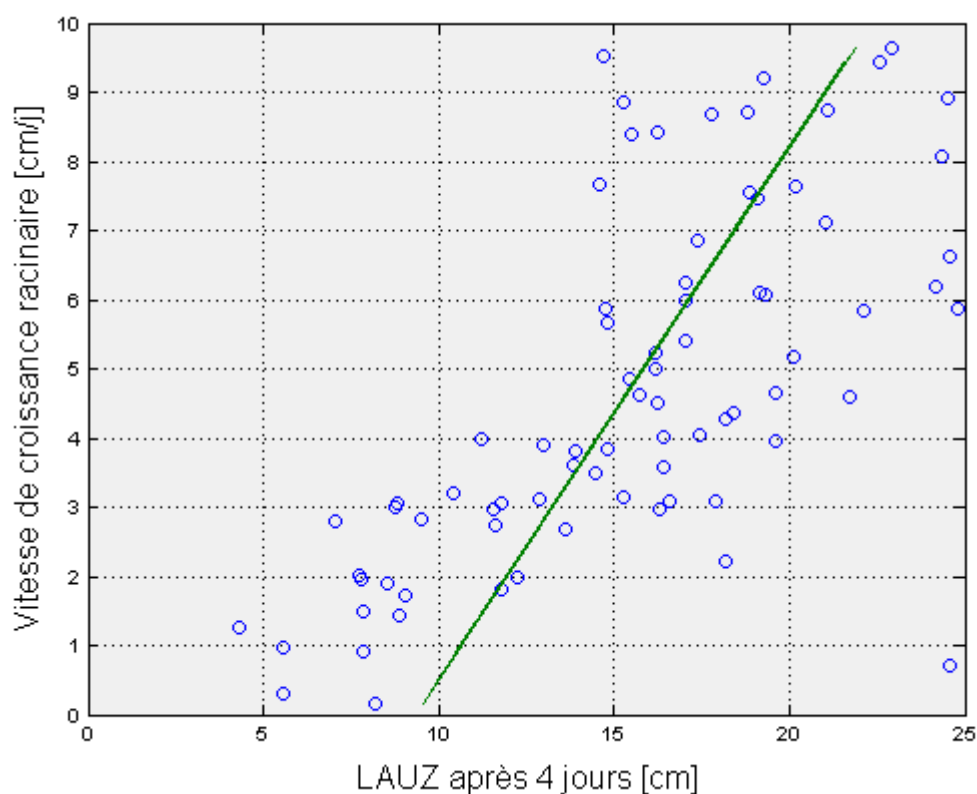


## Résultats

Les deux étapes de numérisation et traitement d'image sont trop exigeantes en temps et nous amènent à exclure l'option imagerie pour l'expérience principale. Néanmoins, nous obtenons des mesures de longueur de racine sur deux dates espacées de 4 jours et de LAUZ sur une date. Ces mesures visent à valider la relation vitesse de croissance – LAUZ. La vitesse de croissance s'obtient, pour une racine mesurée à deux dates, par la différence de longueur rapportée au temps espaçant les deux dates (4 jours, durée choisie sur base des hypothèses présentées en Chapitre II. 1. c.).

Cependant, ce second objectif bute sur un obstacle inattendu: la correspondance de racines séminales d'une même plante entre deux sessions espacées de 4 jours. En effet, s'il est possible de se baser sur des ressemblances entre deux photos, à deux temps différents, d'un même système racinaire, il n'en demeure pas moins malaisé de réaliser cette opération pour chacune des 5-8 racines que contient le système. Nous avons alors tenté de résoudre le problème en faisant correspondre les racines deux à deux suivant un minimum croissant des différences entre la longueur attendue, basée sur la vitesse de croissance moyenne de la première expérience, et la longueur observée après 4 jours.

Les résultats obtenus par cette mise en correspondance sont néanmoins trop peu fiables et trop peu précis pour pouvoir valider la corrélation vitesse de croissance – LAUZ. Ces résultats sont présentés sur le graphique de la page suivante.



**FIGURE III. 7.** : Relation entre la vitesse de croissance et la longueur de la zone apicale non-ramifiée mesurée 4 jours après la première mesure de longueur de racines séminales.



### c. Expérience 3: Mise au point d'un bac de grande dimension

#### *Objectifs*

L'objectif de cette troisième expérience est d'évaluer la possibilité de réaliser une culture aéroponique d'orge dans un bac de grande dimension. Il s'agit donc de voir si la culture telle qu'elle fût menée dans les expériences précédentes est envisageable dans de plus grandes dimensions et d'estimer le temps que requiert les mesures et l'observation des plantes dans une culture de cette taille.

#### *Matériel*

Un nouveau bac aéroponique, de plus grande dimension est construit. En voici une description complète.

De 3,6 mètres de long sur 1,9 m de large, il est profond de 115 cm (voir Figures III. 8. et III. 9., aux pages suivantes). Sur une armature en gouttières inoxydables reposent des pans verticaux constitués d'une double épaisseur (3 cm + 2 cm) de panneaux de polystyrène extrudé bleu (panneaux Floormate). Le bac se partage dans la longueur en deux compartiments, large chacun de 70 cm. Dans cet espace prennent place deux portes de 50 cm de large sur 100 cm de haut (porte visible à la Figure III. 5.), qui laissent entrevoir chacun de ces compartiments et permettent à deux expérimentateurs d'observer et mesurer les systèmes racinaires d'une bande de racines défilant dans le réduit prévu à cet effet à hauteur de la porte.

Légèrement surélevé d'un côté, le fond du bac est préalablement constitué d'une bâche de plastique alimentaire. Cependant, cette bâche sera ultérieurement remplacée par des panneaux de bois marin joints avec une colle étanche (FixAll). Un évacuateur sis au côté bas de chaque compartiment permet l'évacuation de la solution vers la pompe. Le dessus se compose de bandes non solidaires longues de 65 cm sur une largeur de 7 cm reposant sur des rails plastiques collés en haut des parois intérieures de chacun des compartiments. Chacune de ces bandes est trouée huit fois sur un diamètre de 1 cm afin de placer ainsi 8 plantes espacées de 7 cm par bande. Les bandes sont disposées en blocs: 2 blocs d'un côté (23 et 20 bandes) et 2 blocs de l'autre (22 et 20 bandes), permettant ainsi de mettre en culture un maximum de 680 plantes. La raison de ces blocs est anodine; un asperseur étant situé 10 cm plus bas que d'autres à la suite d'une réparation, nous avons choisi de ne pas disposer de plantes dans le rayon d'aspersion de cet élément de la rampe.

Dans chacun des compartiments, on retrouve une rame de 20 asperseurs. La tête (inoxydable) de ces asperseurs culmine à 40 cm du fond du bac, c'est-à-dire à 75 cm en-dessous du collet des plantes repiquées. Les rampes d'asperseurs sont reliées au conduit de sortie de pompe.

La pompe se trouve à l'extérieur et est constituée de divers éléments.

Tuyau en PVC rigide de 22 cm de diamètre extérieur sur 207 cm de long, le corps de pompe

reçoit pendant la période de repos la solution venant des évacuateurs par un tuyau d'admission muni d'un clapet anti-retour. Pendant la période d'aspersion, le corps de pompe est mis sous pression à l'aide d'air comprimé, forçant la sortie de la solution par l'unique conduit vers les asperseurs (le retour vers l'évacuateur est bloqué par le clapet anti-retour). La mise sous pression s'opère de manière progressive. En première phase, l'air comprimé à 7 bars est envoyé à un débit faible jusqu'à atteindre une pression de 3 bars dans le corps de pompe. En seconde phase, le débit est augmenté, la pression atteint 7 bars et est maintenue à ce niveau pour la durée d'aspersion.

Un deuxième dispositif extérieur est relié au bac aéroponique: il s'agit d'une pompe péristaltique qui assure l'apport d'un volume donné de solution nutritive fraîche au volume déjà présent dans le bac et retire le même volume de solution nutritive présente dans le bac (pour peu que le niveau de celle-ci soit suffisant). Par ce mécanisme, la solution nutritive est renouvelée en continu dans le bac d'expérience. Le rythme d'apport initial est de 216 ml par heure.

Ci-dessous se trouve une vue de profil du bac aéroponique au niveau d'un doublon d'asperseurs. Une vue de haut de l'ensemble du dispositif expérimental est reprise à la page suivante. Dans chacune de ces figures, les proportions ne sont pas respectées.

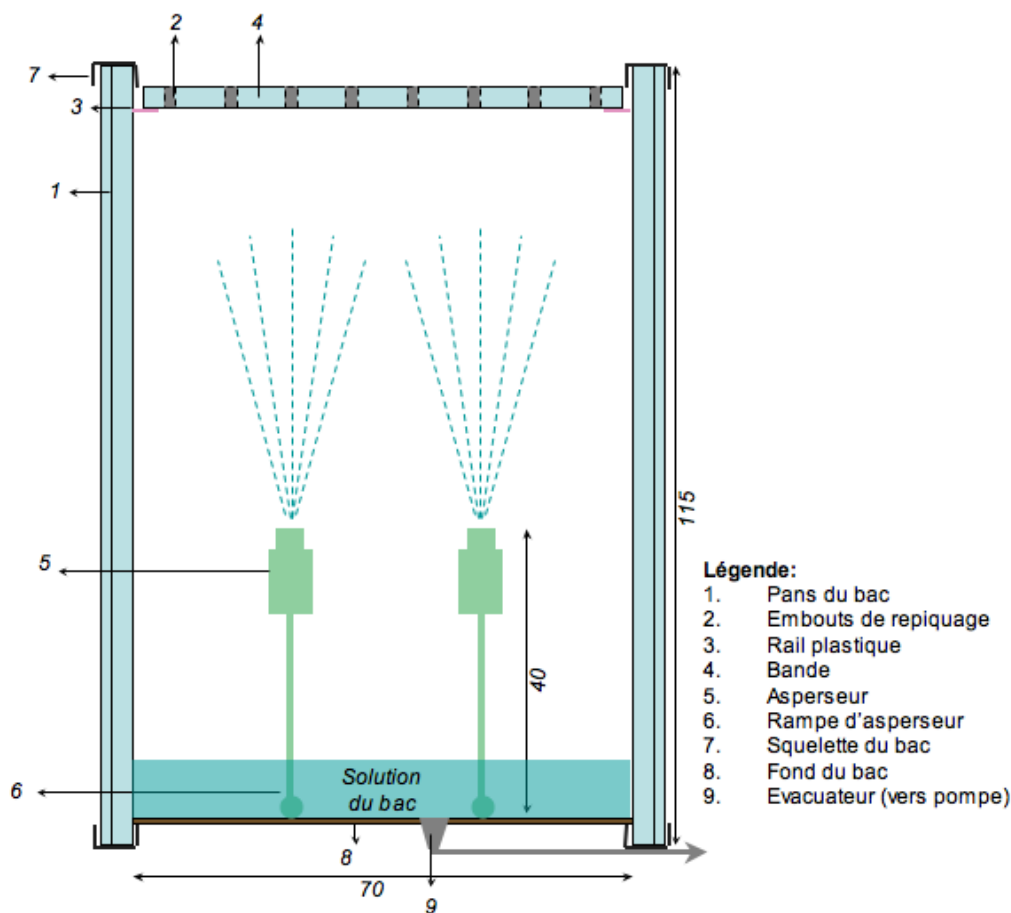


FIGURE III. 8.: *Vue de profil du bac aéroponique au niveau d'un doublon d'asperseurs.*



face d'une porte (voir Figure III. 5.). Chaque plante est identifiée et les longueurs de ses racines sont mesurées.

### *Résultats*

Les mécanismes d'aspersion fonctionnent sans problème.

En ce qui concerne les plantes cultivées, deux problèmes de culture sont relevés. Le premier est une carence minérale (probablement magnésienne), telle qu'observée précédemment (Expérience préliminaire 1.). Elle survient après quelques jours de culture. Le second problème est plus important; il concerne l'infestation et la colonisation du matériel d'expérience par une multitude de pathogènes. La solution nutritive présente dans le bac dégage alors une forte odeur, imprègne les parois d'un film d'aspect huileux et mousse en surface.

Les causes envisagées de ce second problème sont l'utilisation d'une bâche de plastique alimentaire dans le fond du bac et une mauvaise désinfection des graines d'orge. La bâche de plastique alimentaire est le seul matériau différent des autres matériaux utilisés dans les expériences précédentes. Après avoir répété plusieurs expériences de culture dans les bacs aéroponiques utilisés précédemment (expérience 1) avec ou sans cette bâche et avec ou sans désinfection des graines, il apparaît que la bâche de plastique alimentaire fournit un refuge aux pathogènes issus de graines non-désinfectées. C'est pour cette raison que nous choisissons de remplacer la bâche par des panneaux de bois marin.

En ce qui concerne la prise de mesure, il s'avère très aisé de déplacer et maintenir les plantes mesurées dans un milieu humide (donc pas de dessèchement) et la mesure est rapide. Nous estimons à 3 minutes le temps nécessaire pour prendre correctement les mesures de l'ensemble des systèmes racinaires d'une bande. Cette durée est considérablement plus courte que celles des expériences précédentes.

En conclusion, ce bac d'expérience sera réutilisé dans l'expérience principale.

## III. 5. Expérience principale

Cette expérience principale est ainsi dénommée parce qu'elle nous permet de générer les données utiles à nos développements dans les Chapitres IV à VI. Le détail de cette expérience reprend, dans l'ordre, les conditions (serre, température, lumière, ...) dans lesquelles l'orge a été cultivée durant cette expérience, le plan expérimental de disposition des plantes suivi durant cette expérience, le matériel végétal cultivé, les types de mesures réalisées et leurs protocoles et, enfin, le pré-traitement de l'information générée afin d'en permettre une lecture claire et rapide.

Rappelons préalablement que le matériel utilisé est identique à celui de la dernière expérience préliminaire. Les rythmes d'aspersion sont également identiques. Cependant, le rythme d'apport de solution nutritive fraîche et de vidange de solution défraîchie différera au long de l'expérience en fonction du besoin des plantes et de l'évolution du pH.

#### a. Conditions de culture

La culture est menée dans le bac aéroponique développé pour l'expérience préliminaire 3 et situé dans la serre PATO 1 (serres de phytopathologie). Elle s'est déroulée sur plusieurs semaines, du 19 février 2007 (date de repiquage) à la semaine du lundi 26 mars 2007. Durant cette semaine, la récolte des plantes a été échelonnée sur plusieurs jours (du mercredi 28 mars au samedi 31 mars).

Plusieurs paramètres environnementaux ont été mesurés tout au long de l'expérience principale. Ces paramètres sont :

- le pH : mesuré en unité de pH au moyen d'une électrode à pH;
- la conductivité : mesurée en  $\mu\text{S}\cdot\text{cm}$  au moyen d'une électrode;
- le PAR (photosynthetically active radiation) : mesuré en  $\mu\text{moles}\cdot\text{cm}^{-2}\cdot\text{s}^{-1}$  au moyen de 3 capteurs disposés deux à la base des plantes et un au sommet de la canopée et reliés à un appareil d'enregistrement, le PARDataLogger;
- la température: mesurée à l'extérieur du bac aéroponique en deux endroits: à la base des plantes (capteur *Tynitag*) et en-dessous de cette base (sonde dans le PARDataLogger).

L'appareil d'enregistrement PARDataLogger effectue des mesures de température et de PAR toutes les 15 minutes.

Le pH a fait l'objet d'ajustements successifs. Nous avons, en effet, observé de grandes variations de pH (pH variant de 5,8 à 8,3). Vraisemblablement, ces variations sont dues à l'augmentation progressive d'anions hydroxydes ( $\text{OH}^-$ ) dans la solution du bac. Cette augmentation résulterait du fait que l'azote présent dans la solution nutritive est en majeure partie sous forme nitrate et en moindre partie sous forme ammonium (voir Annexe III. 1.). Le prélèvement des ions  $\text{NO}_3^-$  par les plantes serait donc la cause d'une grande augmentation d'ions hydroxydes, rejetés par les plantes, dans la solution du bac.

Les irrégularités observées dans la courbe d'évolution du pH (Figure III. 10.) sont dues aux ajustements successifs du pH par retrait de solution du bac défraîchie et apport de solution nutritive fraîche (pH de cette solution apportée: 5,75). En raison d'un besoin accru en nutriments, le rythme d'apport de la pompe péristaltique est augmenté à 338,41 ml par heure, à partir du 15/03/2007.

La conductivité de la solution a également contribué à apprécier la nécessité d'un apport de solution fraîche; de la nouvelle solution devant être apportée si la conductivité était trop basse.

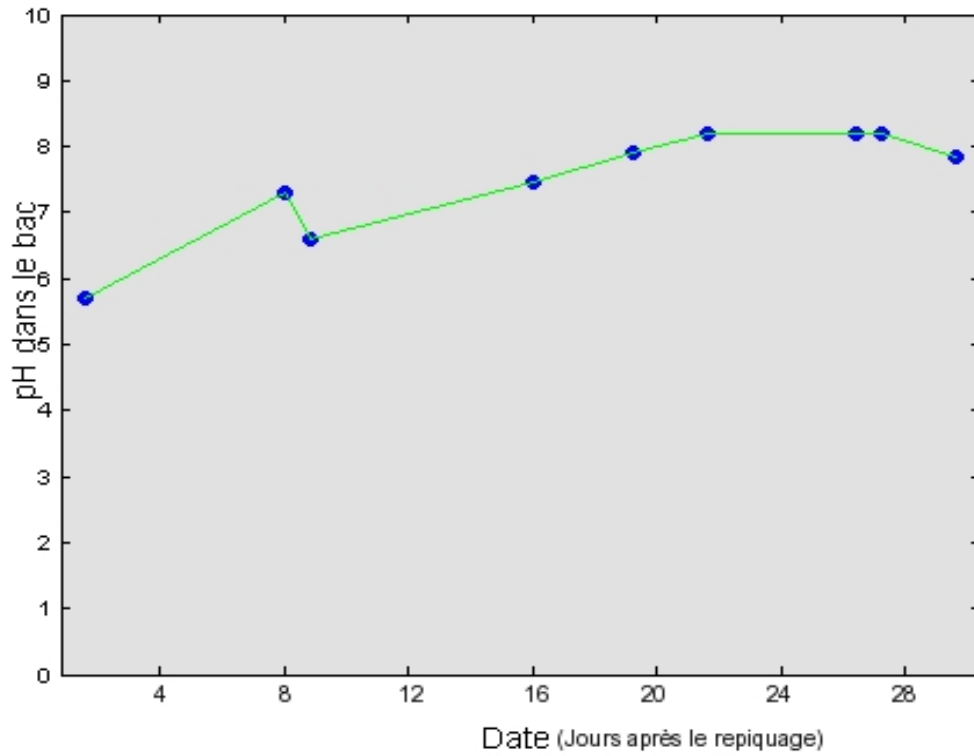


FIGURE III. 10.: Evolution du pH au long de l'expérience.

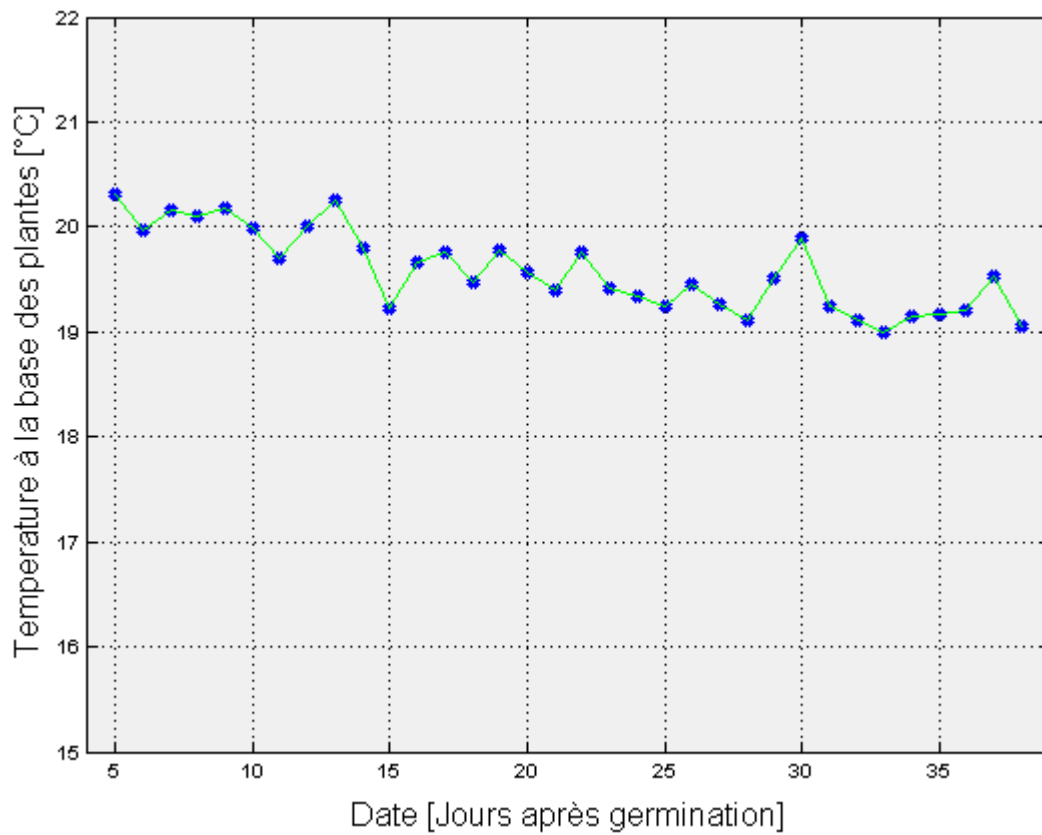
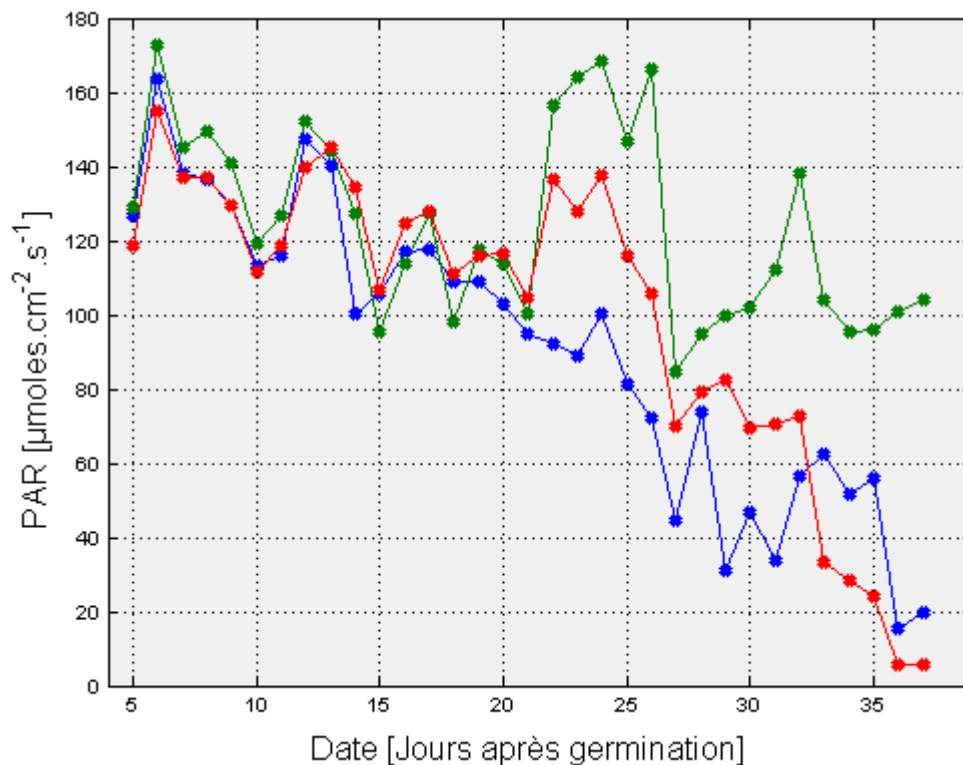


FIGURE III. 11.: Evolution (moyennes journalières) de la température enregistrée par le PARDataLogger en-dessous de la base des plantes du 21/02/2007 au 23/03/2007.

La température (Figure III. 11.) est relevée en-dessous de la base des plantes du 21 février (soit deux jours après le repiquage) au 27 mars. Elle montre des variations journalières et une tendance linéaire décroissante. Cette tendance reste mineure; on observe, en effet, des écarts de l'ordre du degré celsius entre le début et la fin de l'expérience.

Le rayonnement photosynthétiquement actif (PAR, *photosynthetically active radiation*) est mesuré en deux points, du 21 février au 23 mars. Il tend à diminuer sur la durée de l'expérience. Cette diminution est néanmoins plus marquée pour les capteurs situés à la base des plantes (en bleu et rouge, sur la Figure III. 12.) que pour le capteur situé au niveau de la canopée (vert).

Les moyennes journalières des données mesurées en  $\mu\text{moles.cm}^{-2}.\text{s}^{-1}$  sont reprises sur le graphique ci-dessous.



**FIGURE III. 12.:** Evolution (moyennes journalières) du rayonnement photosynthétiquement actif enregistré par le PARDataLogger à la base des plantes (bleu et rouge) et au niveau de la canopée (vert), du 21/02/2007 au 23/03/2007.

### b. Plan expérimental

Nous avons suivi un plan d'expérience respectant les caractéristiques suivantes:

- i. 4 plantes sont repiquées par génotype (8 pour les lignées parentales).
- ii. Les 4 plantes de chaque génotype se retrouvent sur la même bande, aux positions 1, 3, 5, 7 (ou, inversement, 2, 4, 6, 8). La Figure III. 12. (page suivante) illustre cette position en

alternance.

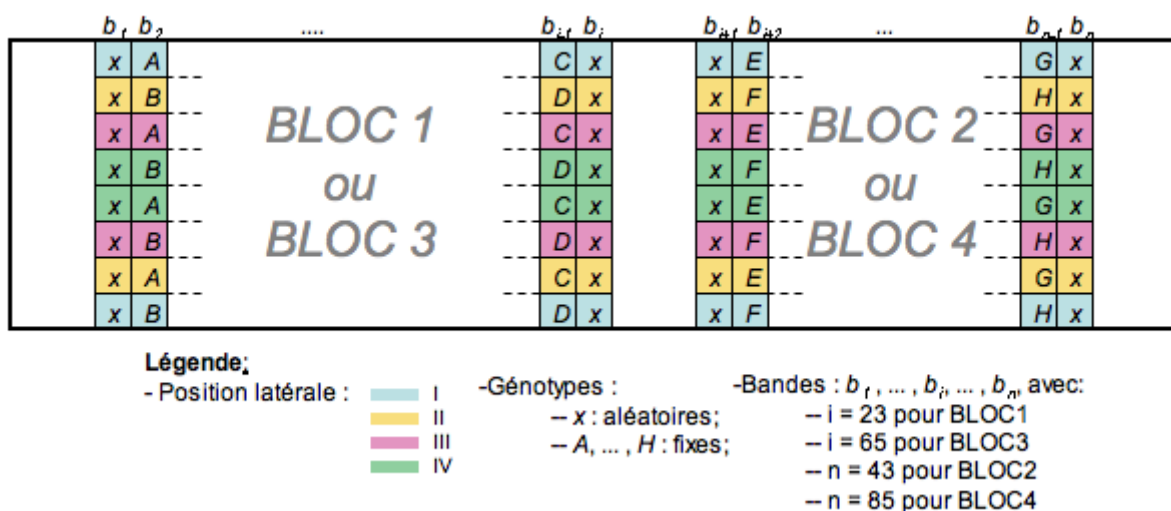
- iii. Deux génotypes sont présents par bande.
- iv. Le placement des génotypes sur les différentes bandes est tiré aléatoirement.
- v. Pour des raisons évoquées plus haut (III. 4. c. §*Matériel*), le plan est divisé en 4 blocs.

Chacun de ces 4 blocs est bordé, de part et d'autre, par une bande sur laquelle se trouvent 8 plantes, de génotypes différents, tirés aléatoirement.

Chacune des autres caractéristiques s'expliquent par le souci de randomiser au mieux, le cas échéant, des effets environnementaux aléatoires. Ainsi, les caractéristiques ii et iii permettraient de contrôler l'effet bande (85 niveaux) et position sur la bande (4 niveaux).

D'éventuelles erreurs durant les opérations de récolte (contamination par une parcelle voisine, de génotype différent), de germination (interversions entre graines de génotypes différents) et de repiquage (interversions entre graines germées de génotypes différents) pourraient être détectées, sur base de la succession des génotypes durant ces différentes étapes. La caractéristique iv, en imposant une disposition aléatoire des génotypes, met un terme à cette succession. De ce fait, elle rend possible une telle évaluation, au moyen d'outils statistiques d'appréciation de la dépendance entre valeurs d'une variable aléatoire en fonction de la distance qui séparent ces valeurs (semi-variogrammes).

Enfin, la dernière caractéristique, en plaçant une bande à chaque bord de bloc crée un environnement identique pour toutes les bandes situées à l'intérieur du bloc, qu'elles se trouvent à proximité du bord ou non.



**FIGURE III. 13.:** Schéma du plan expérimental détaillant les facteurs fixes (position latérale, génotype et bande). Ce schéma porte sur un compartiment du bac aéroponique illustré en Figure IV. 8.



### c. Matériel végétal

Dans cette expérience, les 149 génotypes cités précédemment interviennent, à raison chacun de 4 plantes par génotypes. Afin d'obtenir, après germination, effectivement ces 4 plants par génotype, nous avons donc fait germer 6 grains, choisis sur leur état apparent, par génotype.

A ces 596 plantes s'ajoutent 8 bandes de 8 plantes de génotypes choisis aléatoirement parmi les 149 génotypes. Ces bandes bordent chacun des 4 blocs, afin de se prémunir contre d'éventuels effets de bord. Placées en germination dans la journée du 16 février 2007, ces plantes ont été repiquées dans la journée du 19 février 2007.

Enfin, à ces 660 plantes, se rajoutent 8 plantes de chacun des génotypes parentaux, DERKADO et B83. Remarquons finalement que ces 16 dernières plantes sont plus jeunes de 4 jours par rapport au reste de la population, c'est-à-dire mises en germination le 20 février 2007 et repiquées le 23 février 2007.

### d. Types de mesures réalisées et protocoles de ces mesures

Les mesures sont réalisées par deux expérimentateurs placés devant chaque porte du bac aéronomique. Chaque expérimentateur est muni d'un lecteur code-barre, d'une règle graduée en code-barres et des informations disponibles en Tableau III. 1. La méthode et les principes de mesure sont tels qu'énoncés en section 4. a. de ce chapitre.

Les types de mesures diffèrent à chaque date et peuvent également être différents en fonction de l'expérimentateur. Cependant, les informations reprises dans une sous-séquence (portion de la séquence enregistrée comprise entre deux identificateurs) restent de même types, quel que soit l'expérimentateur.

Voici les objets, dans l'ordre chronologique, des dix sessions de mesure effectuées durant l'expérience.

(1) Mesure de la longueur des racines séminales et de la hauteur des feuilles du maître-brin, le 21/02/2007.

(2) Mesure de la longueur des racines séminales et de la hauteur des feuilles du maître-brin, le 23/02/2007.

(3) Mesure de la longueur des racines séminales, de la LAUZ de la plus longue racine séminale et de la hauteur des feuilles du maître-brin, le 26/02/2007.

(4) Mesure de la longueur des racines adventives, de la longueur de la racine séminale la plus longue, de la LAUZ de la plus longue racine séminale, de la hauteur des feuilles du maître-brin et appréciation de l'état pathologique des feuilles et racines, le 02/03/2007.

(5) Mesure de la longueur des racines adventives, de la longueur de la plus longue

séminale, de la LAUZ de la plus longue séminale et de la hauteur de la dernière et de l'avant-dernière feuille apparue sur le maître-brin et chacun des talles, le 08/03/2007.

(6) Mesure de la longueur des racines adventives, de la longueur de la plus longue racine séminale, de la LAUZ de la racine adventive et de la racine séminale les plus longues, le 12/03/2007.

(7) Mesure de la hauteur des dernières et avant-dernières feuilles apparues sur le maître-brin, sur les talles primaires et sur certaines structures indéterminées, le 14/03/2007.

(8) Mesure de la longueur des racines adventives et de la LAUZ de la racine adventive la plus longue, le 16/03/2007.

(9) Mesure de la longueur des racines adventives et de la LAUZ d'une racine adventive, le 20/03/2007.

(10) Mesure de la longueur des racines adventives, le 23/03/2007.

Dans tous les cas présentés ci-dessus, la longueur de la zone non-ramifiée (LAUZ) s'obtient par soustraction de la longueur de la zone ramifiée (*length of branched zone*, LBZ) à la longueur de la racine. Ce qui est mesuré et enregistré dans la séquence de mesure, c'est donc la LBZ et non pas la LAUZ.

En Annexe III. 2., sont reprises chacune de ces sessions de mesure avec une séquence typique de mesure et les commentaires sur le protocole mené.

e. Méthode et algorithme de pré-traitement de l'information brute et d'encodage des différents protocoles de mesure.

*Méthode de lecture des fichiers de données*

Comme expliqué précédemment (section 4. c. de ce chapitre), l'information, stockée dans les tableurs (fichiers de données) est structurée en une séquence de lecture. Chaque élément de cette séquence est constituée de l'information contenue dans un code-barres et peut être divisée en deux parties. La première, partie alphabétique, réfère à la nature de l'objet mesuré (racine, feuille, ..), c'est l'étiquette. La seconde, partie alphanumérique, est une valeur que peut prendre un certain type de mesure, c'est la donnée proprement dite.

La succession des éléments dans la séquence enregistrée est préalablement définie par le protocole de mesure de chaque session de mesure. Pour réorganiser la séquence en un tableau ordonné, dans lequel chaque ligne reprend l'ensemble des observations et mesures faites sur une plante, il faut donc développer une méthode qui intègre le protocole de mesure. La méthode conçue doit pouvoir être implémentée en un algorithme générique, c'est-à-dire qui soit identique quel que soit le protocole de mesure spécifié.

Cependant, au sein d'une séquence concernant une même plante, des éléments de nature et de type identiques peuvent concerner différents objets de mesure. Cela arrive souvent, notamment lorsqu'à une même date, on mesure pour une même plante des longueurs racinaires ('L0250' par exemple) et des longueurs de la zone non-ramifiée ('L0150' par exemple). Nous aurions pu prévoir deux étiquettes différentes pour ces deux objets de nature forcément différente. Cependant, nous avons préféré travailler avec un nombre plus restreint de code-barres. Les différentes acceptions que prend une étiquette sont intégrées au protocole de mesure. Ainsi, par exemple, la longueur d'une racine est lue après une première lecture de l'identificateur et la LAUZ d'une racine est lue après une seconde lecture de l'identificateur.

Certaines étiquettes ont donc, tout au long de l'expérience, des acceptions différentes (Tableau III. 2.). Dans ce tableau, *Lg* désigne la longueur totale d'une racine (*LBZ* + *LAUZ*), *LBZ* la longueur de la zone ramifiée, *sem* une racine séminale, *adv* une racine adventive, *PATO* un état pathologique, *STRF* une structure foliaire et *Ht* une hauteur de feuille. Enfin, *FLAG* désigne une observation de type signal. Ce dernier type n'est pas repris dans l'énumération des types de mesure et d'observation parce qu'il n'offre qu'une observation ponctuelle, sans valeur. Par exemple, le signal ID0693 signifie que l'élément précédent, dans la séquence, est erroné.

**TABLEAU III. 2.:** *Acceptions établies à chacune des étiquettes employées durant l'expérience.*

<b>Etiquette</b>	<b>Acception 1</b>	<b>Acception 2</b>	<b>Acception 3</b>	<b>Acception 4</b>	<b>Acception 5</b>
<b>L</b>	Lg sem	LBZ sem	Lg adv	LBZ adv	L0005 = FLAG
<b>R</b>	PATO racine				
<b>F</b>	PATO feuille	F1 = STRF	F2 = STRF	F3 = STRF	F4 = STRF
<b>T</b>	Talle				
<b>ID</b>	Identificateur	ID0693 = FLAG			
<b>H</b>	Ht feuille				

Lors de la lecture du fichier, le choix des acceptions se fait sur base d'un certain nombre de règles. Ces règles, déduites des protocoles de mesures, sont (les abréviations sont reprises du Tableau III. 2.): .

- si *adv* et *sem* apparaissent dans un même protocole, il n'y a qu'une seule *sem* et c'est la plus longue;
- si *LBZ sem* et *LBZadv* apparaissent dans un même protocole, il n'y en a qu'une seule de chacune et leurs longueurs les discriminent;
- avec n'importe quel autre étiquette elles puissent se trouver, *LBZ* et *Lg* n'appartiennent jamais au même ID;

– si L est accepté comme *FLAG*, il a une valeur toujours égale à 0005, donc, nettement inférieures à ce que peuvent prendre les autres racines comme valeurs de L.

Ces règles nous permettent donc de réduire considérablement le nombre de paramètres nécessaires pour coder l'ensemble des protocoles de mesure. En effet, il ne sera plus nécessaire que de connaître l'acception correspondante à l'étiquette L. Cette acception est codée par deux paramètres. Le premier a pour valeur 1, 2 ou 3 en fonction des racines mesurées, respectivement séminales, adventives et les deux ensemble. Le second paramètre admet deux valeurs 1 ou 0, selon que le protocole considéré contient des signaux étiquetté L ou n'en contient pas.

Ce codage permet de déduire l'ensemble du protocole de mesure d'une session donnée.

Ces paramètres seront donc les paramètres d'entrée de l'algorithme implémenté sous la forme d'une fonction MATLAB®, disponible sur le CD attaché à ce mémoire.

### *Description de la fonction de pré-traitement de l'information*

#### *(1) Généralités.*

Nommée *rootmat.m*, cette fonction fonctionne dans l'environnement MATLAB. Elle s'accompagne de plusieurs fonctions subordonnées. Ces fonctions subordonnées sont: *transcode.m*, *metainfo.m*, *feuille.m*, *pato.m* et *racine.m*.

Dans les développements qui suivent, sera appelé « jeu de données » un ensemble d'observations et/ou mesures relevant d'un même identificateur.

#### *(2) Paramètres d'entrée.*

Les paramètres d'entrées codant les protocoles sont dénommés *code* et *codeFLAG*.

Le tableur contenant l'information subit une modification préalable, facilement réalisée dans l'environnement Excel®. Afin d'effectuer sa lecture par *rootmat*, il convient de séparer l'unique colonne en six colonnes, contenant chacune un caractère. Ce tableur transformé doit ensuite être enregistré sous un format texte séparé par tabulation.

#### *(3) Outputs.*

Après utilisation sur une session de mesure, *rootmat* renvoie 15 outputs.

Les 3 premiers outputs résument en une matrice, sa légende et les positions de sa légende, les données contenues dans le tableur de base. Nous avons ainsi la matrice, suivie d'une variable texte contenant les codes de chacune des informations contenues dans cette matrice et d'un vecteur reprenant les positions de ces informations en terme de colonnes de la matrice.

Les 12 derniers inputs correspondent aux douze objets différents ayant été observés et ou

mesurés sur une seule plante durant l'expérience. Il s'agit, dans l'ordre de sortie de *rootmat*:

- de la longueur et de la LAUZ de séminales et d'adventives;
- de la hauteur des feuilles du maître-brin, des talles 1 à 3, des structures indéterminées de types 1, 2 et 3-4 (voir Annexe IV. 1.: Session 7);
- de l'état pathologique des feuilles et racines.

#### (4) Pseudo-code de la fonction *rootmat*

Cette section développe les lignes majeures de fonctionnement de l'algorithme.

Étape 1: Lecture du fichier de données. Les informations, retranscrites au moyen de la fonction *transcode*, sont stockées en deux vecteurs jumeaux E et D, étiquette et donnée correspondante.

Étape 2: Élimination des FLAGS d'observation erronée (ID0693) et suppression des lignes correspondantes.

Étape 3: Évaluation du contenu de la séquence et des sous-séquences (portions de séquence séparées par deux identificateurs) en terme de nature des mesures effectuées, évaluation réalisée par la fonction *metainfo*, qui renvoie une matrice dont chaque ligne correspond à un jeu et chaque colonne à une étiquette. Chaque valeur de cette matrice donne donc le nombre d'observations ou de mesures relevant d'une certaine étiquette, pour une sous-séquence donnée.

Étape 4: Sur base de cette matrice, élimination des sous-séquences vides. Ensuite, si certains identificateurs sont présents en triples exemplaires, le déroulement de la fonction est interrompu pour permettre à l'expérimentateur de résoudre ce problème. Il n'y a jamais plus de deux sous-séquences par identificateur dans les protocoles de mesure.

Étape 5: Pour le premier identificateur relevé dans la liste, la fonction procède au traitement des sous-séquences relatives à cet identificateur. Par comparaison entre ces sous-séquences et sur base des informations relatives disponibles dans la matrice émanant de *metainfo*, la fonction détermine alors de quel objet relève chacune des observations et/ou mesures effectuées. Ces informations sont alors retranscrites dans la matrice correspondant à leur objet par les fonctions *pato*, *feuille* et *racine*.

Étape 6: La fonction réitère l'étape 5 jusqu'au dernier identificateur relevé dans la liste.

Étape 7: L'ensemble des observations et mesures sont concaténées en une seule matrice. La légende et l'indice de colonne relatif à cette légende sont également établis.

## *Chapitre IV*

# *Correspondances sur deux dates par calcul combinatoire*

Ce chapitre s'intéresse au problème posé par la correspondance de racines appartenant à une même plante et dont l'ensemble du faisceau est mesuré à deux dates successives. Ce problème a été rencontré pour la première fois suite à l'expérience préliminaire 2 (Chapitre III. 4. b.), menée en novembre 2006.

Pour rappel, l'objectif de cette expérience était de valider la relation linéaire entre vitesse de croissance d'un axe racinaire sur une période donnée et longueur de la zone non-ramifiée de cet axe après cette période. Les mesures étaient alors effectuées par ordinateur sur le plug-in Smart Root. La mise en évidence de cette relation impliquait que l'on puisse déterminer sans ambages à quelle racine d'un faisceau d'une plante donnée correspondait une racine du même faisceau mesurée un certain temps après la première.

La méthode conçue et développée dans ce chapitre fait appel au calcul combinatoire et son implémentation peut se baser sur deux indicateurs différents.

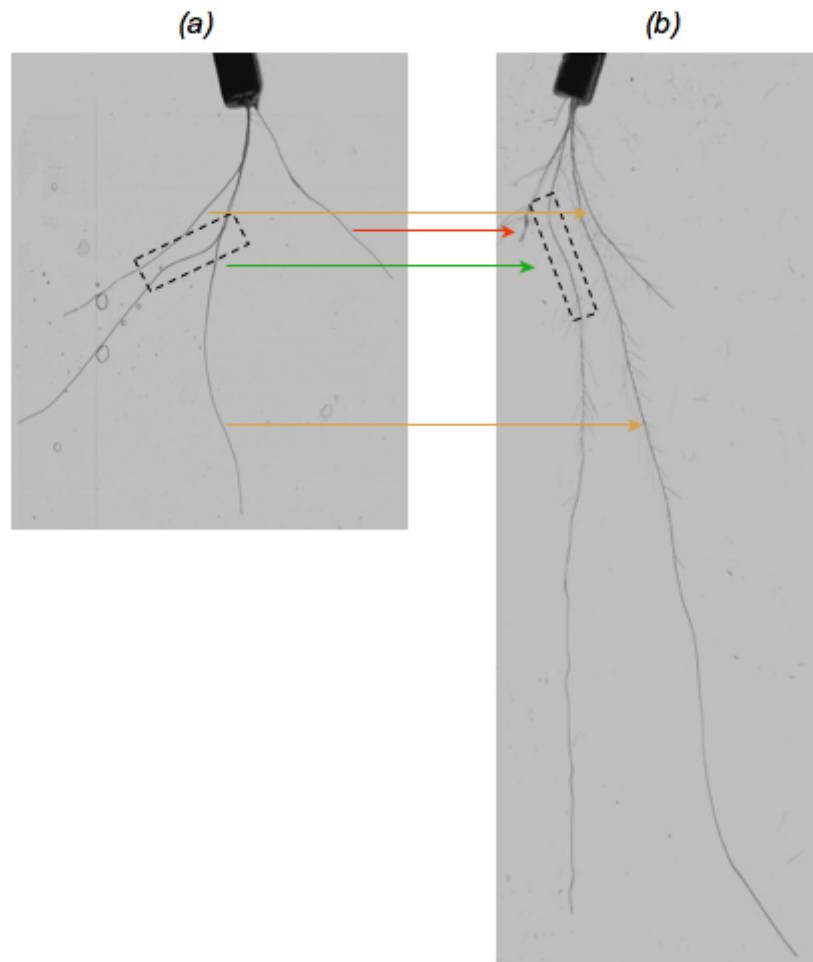
## IV. 1. Introduction et hypothèses.

### *Mise en situation*

La figure suivante (fig. IV.1 (a) et (b)) représente deux images obtenues par numérisation du même système racinaire à deux dates successives (fig. IV. 1.(a) : 20/10/2006 et fig. IV. 1.(b) : 24/10/2006) lors de la deuxième expérience préliminaire. Sur base de la configuration bi-dimensionnelle de certaines racines (courbures, longueurs, ..), nous pouvons faire correspondre une certaine racine sur deux dates. Cette reconnaissance n'est cependant possible que pour peu de racines.

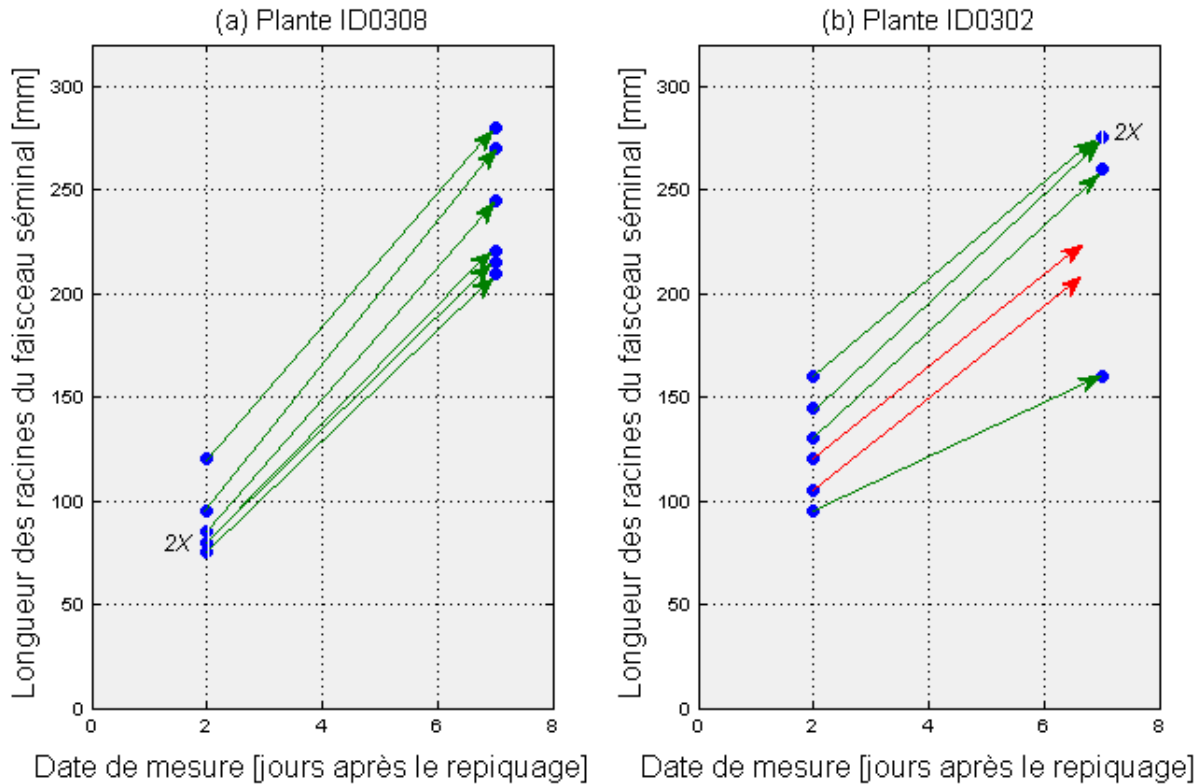
Il faut donc déterminer à quelle racine du second faisceau correspond chacune des racines du premier, en gardant à l'oeil plusieurs contraintes:

- (1) l'ordre du classement suivant la longueur n'est pas forcément conservé;
- (2) une ou plusieurs racines ont pu apparaître;
- (3) une ou plusieurs racines ont pu voir leur croissance stoppée, voire disparaître.



**FIGURE IV.1:** Images d'un système racinaire (Génotype SCRI 83) obtenues par numérisation à deux dates différentes: (a) le 20/10/2006 et (b) le 24/10/2006. **Légende :** →, correspondance certaine (en vert), probable (en orange) ou incertaine (en rouge) entre deux racines. Le quadrilatère pointillé indique le motif de reconnaissance de la racine sur deux dates. La correspondance rouge est incertaine en raison d'une cassure sur la racine correspondante à la deuxième date.

La figure qui suit est, elle, extraite de l'expérience principale. Chaque graphique illustre le cas d'un faisceau séminal ; y sont représentées les différentes mesures de longueurs racinaires du faisceau réalisées à deux dates différentes.



**FIGURE IV. 2.:** Evolution de la longueur du 21/02/2007 au 26/02/2007 des racines du faisceau séminal de deux plantes: (a) ID0308 et (b) ID0302. **Légende :** ● longueur d'un axe racinaire, →: correspondance supposée entre 2 mesures, 2 x : indique la présence de deux mesures de même longueur.

En outre, dans ce graphique, les flèches présupposent l'évolution de la longueur d'une racine séminale. Comme on le voit en Figure IV. 2. (a), chaque mesure effectuée le 21/02 est mise en correspondance avec une mesure effectuée le 26/2. On pourrait même penser que le cas se présente tel qu'il est illustré, c'est-à-dire en conservant l'ordre de longueur d'une date à l'autre. En Figure IV. 2. (b), la situation est différente. Dans ce cas-ci, 6 mesures à la première date doivent être mises en correspondance avec 4 mesures à la seconde date. La configuration de correspondance illustrée par les flèches repose sur une seule hypothèse : la vitesse de croissance (pente de la flèche) est semblable pour chacune des racines. Cependant, cette configuration ne prétend en rien être la réalité. De plus, deux racines, signalées par les flèches rouges ont soit disparu, soit n'ont pas été mesurées à la seconde date.



La question est dès lors de déterminer quelle configuration de correspondance est la plus probable pour un faisceau donné, connaissant les trois contraintes énoncées plus haut, contraintes toujours d'actualité pour cette seconde situation.

### *Hypothèses*

1. Hypothèse sur les vitesses de croissance : les vitesses de croissance des différentes racines sont normalement distribuées autour de la vitesse moyenne du faisceau de racines. Il n'y a pas de tendance en fonction de l'âge de la racine.
2. Hypothèse sur les racines apparues : les racines de la seconde date inférieure à toutes les racines de la première date doivent être considérées comme nouvelles, c'est-à-dire ayant apparu durant le laps de temps séparant les deux dates de mesure.
3. Hypothèse de non-régression : les plantes évoluent dans un milieu ayant toutes les conditions requises (température, lumière, nutrition minérale) pour observer une croissance normale. Des vitesses nulles ou régressives (inférieures à 0) ne peuvent résulter de la configuration optimale.
4. Hypothèse d'unicité de la configuration optimale : la configuration de correspondance jugée optimale par un certain indicateur sera toujours unique. Si plus d'une configuration devaient renvoyer une même valeur de cet indicateur, on choisira arbitrairement la première des deux afin de satisfaire cette hypothèse.

## IV.2. Algorithme paradigmatique de correspondance

Cette seconde section comporte trois parties : la première schématisera l'algorithme de correspondance au travers d'un cas paradigmatique (situation simplifiée et théorique de correspondance), la seconde présentera un second indicateur, utilisé dans l'implémentation du schéma paradigmatique en langage MATLAB® et la troisième reprendra les étapes de cette implémentation à travers un pseudo-code.

### *Schéma paradigmatique*

Prenons un cas théorique de deux faisceaux, mesurés à deux dates espacées de 4 jours. Les mesures de longueurs réalisées à ces deux dates se trouvent dans le Tableau IV. 1.

**TABLEAU IV. 1.** : Mesures théoriques ( $M_1$  à  $M_5$  en millimètres)d'un même faisceau à deux dates espacées de 4 jours.

	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$\overline{M}$
<b>Date <math>t</math> (A)</b>	100	120	140	150	240	150
<b>Date <math>t+4</math> (B)</b>	155	165	175	285	-	195

Selon l'hypothèse 1. de ce chapitre, les vitesses de croissance doivent être normalement distribuées autour de la vitesse moyenne du faisceau. Cette vitesse moyenne, en millimètres par jour, n'est autre que :

$$m = \frac{(\overline{M}_{t+4} - \overline{M}_t)}{4} = 11,25 \quad (4.1)$$

L'objectif, nommé  $\Omega$ , est donc de déterminer quelles correspondances entre mesures doivent être établies afin que les vitesses de croissance en découlant aient une dispersion minimale autour de la moyenne  $m$ .

Une configuration de correspondance est un ensemble de  $k$  couples de deux racines,  $k$  étant le nombre de mesures effectuées à la date où on en a effectué le moins ( $k = 4$  dans notre cas). Il faut donc déterminer la configuration de correspondance  $\Theta$ , telle que l'objectif de dispersion

$$\Omega = f(\Theta) \text{ soit minimal en } \Theta.$$

Le dénombrement des configurations de correspondances se base sur le calcul du nombre d'arrangements sans répétitions possibles de  $k$  objets choisis parmi  $n$  objets, ce  $k$  correspondant toujours au nombre de mesures à la date où il y en a le moins et  $n$  au nombre de mesures à la date où il y en a le plus. Une configuration valable est donc :

$$\forall x_1 \neq x_2 \neq \dots \neq x_k, \quad \forall y_1 \neq y_2 \neq \dots \neq y_k$$

$$\Theta = \begin{bmatrix} A_{x_1} B_{y_1} \\ A_{x_2} B_{y_2} \\ \dots \\ A_{x_k} B_{y_k} \end{bmatrix} \quad (4.2)$$

Dans notre situation,  $k = 4$  et  $n = 5$ . Donc,  $N$ , le nombre de configurations possibles, égale :

$$N = \frac{n!}{(n-k)!} = \frac{5!}{1!} = 120 \quad (4.3)$$

L'évaluation de l'objectif pour chacune des  $N$  configurations de correspondance possibles requiert l'élaboration d'un indicateur de dispersion. Nous présentons ici le premier indicateur développé, qui est une somme d'écart. Une section ultérieure de ce chapitre développera un autre indicateur.

Si  $A$  et  $B$  représentent respectivement les mesures  $M$  effectuées à chaque date  $t$  et  $t+4$ , alors chaque mesure  $A_i$  peut, théoriquement, être couplée à n'importe quelle mesure  $B_j$ , avec plus ou moins de succès. Ce succès peut se mesurer par  $e_{ij}$ , la valeur absolue de l'écart entre  $m$  et la vitesse de croissance calculée pour le couple de mesures proposé  $\{A_i B_j\}$  :

$$e_{ij} = \left| \frac{B_j - A_i}{\Delta t} - m \right| \quad (4.4)$$

On déduit donc que plus  $e_{ij}$  sera proche de 0, plus le couple  $\{A_i B_j\}$  aura une correspondance probable, sur base de l'hypothèse de distribution normale des vitesses.

Dans notre cas, nous pouvons calculer  $e_{ij}$  pour 20 couples sur 4 jours, car  $i = 1, \dots, 5$ ,  $j = 1, \dots, 4$  et  $\Delta t = 4$ . Ces valeurs définissent une matrice (5 x 4) présentée dans le tableau ci-dessous.

**TABLEAU IV. 2.** : Valeurs  $e_{ij}$  pour chaque couple de mesures de racines.

	$B_1$	$B_2$	$B_3$	$B_4$
$A_1$	2,5	5	7,5	35
$A_2$	2,5	0	2,5	30
$A_3$	7,5	5	2,5	25
$A_4$	10	7,5	5	22,5
$A_5$	32,5	30	27,5	0

Enfin, pour chaque configuration de correspondance  $\Theta$  élaborée sur le modèle de l'équation (4. 2) et dénombrée suivant l'équation (4. 3), on calcule la somme des valeurs  $e_{ij}$  de chaque couple composant la configuration. Cette somme est la valeur objectif  $\Omega$ . La meilleure configuration est donc celle qui a la valeur objectif la plus proche de 0, donc minimale.

### *Indicateurs implémentés*

#### (a) Somme des écarts absolus à la vitesse moyenne $m$ du faisceau

Cet indicateur n'est autre que l'objectif de dispersion  $\Omega$  tel que présenté ci-précédemment. C'est une mesure non pondérée de dispersion de la distribution normale des vitesses de croissances des racines du faisceau. Sa valeur ne peut donc se présenter comme un coefficient de satisfaction d'une configuration donnée. En effet, son domaine de variation étant non borné vers  $+\infty$ , sa valeur ne permet pas de comparaisons entre deux situations différentes, c'est-à-dire entre deux plantes différentes.

Pourquoi avoir développé cet indicateur en premier lieu ?

Simplement parce que nous pensons pouvoir déduire la configuration optimale par recherche itérative de minima dans le Tableau IV. 2., donc sans dénombrement exhaustif de toutes les

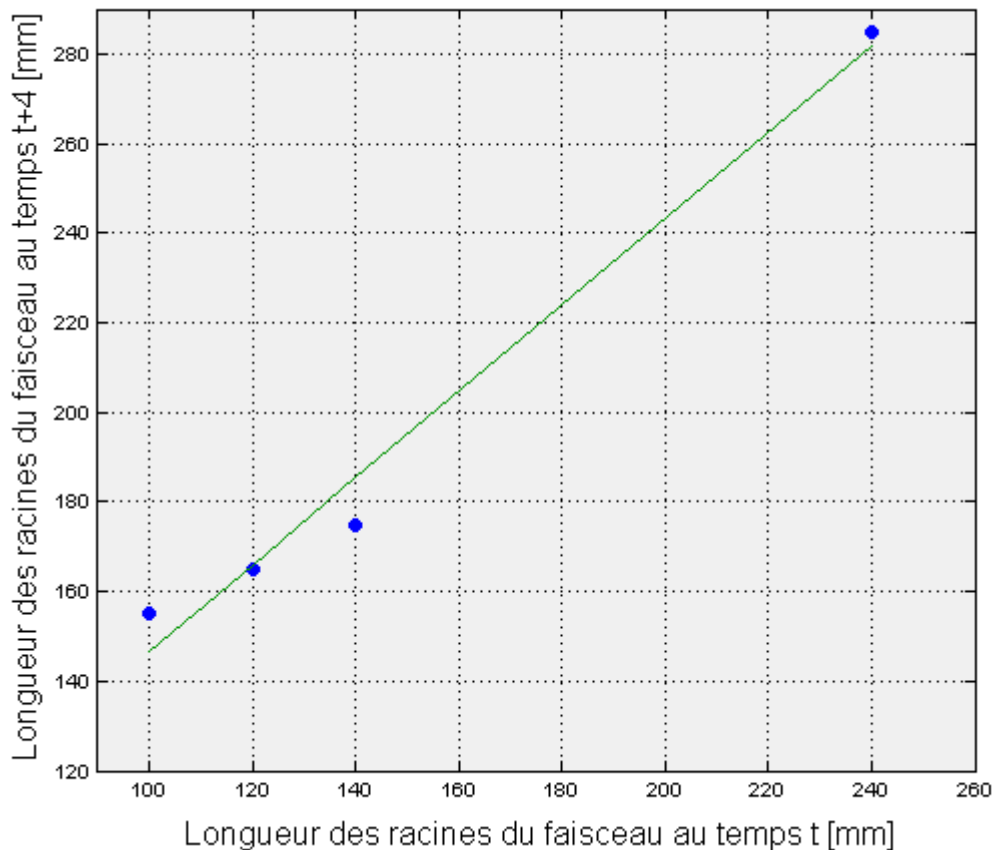
configurations. Cette méthode procède comme ceci :

- (1) Recherche du minimum du Tableau IV. 2.
- (2) Couplage des mesures  $A_i$  et  $B_j$  correspondant à la valeur minimale.
- (3) Elimination des colonnes correspondant aux mesures  $A_i$  et  $B_j$  du Tableau IV. 2.
- (4) Réitération de la démarche jusqu'à ce qu'il ne reste aucune cellule dans le Tableau IV.2.

Procéder de la sorte peut s'avérer tronqué dans certains cas car on exécute des recherches successives de minima locaux et pas un minimum global.

(b) Coefficient de détermination de la droite de correspondance

Cet indicateur calcule un objectif de dispersion égal au coefficient de détermination de la droite de correspondance. Celle-ci est la droite de régression des longueurs à la seconde date sur les longueurs à la première date (Figure IV. 3.). Si l'hypothèse de normalité des vitesses est respectée, on s'attend à avoir un coefficient angulaire égal à 1 et un intercepte égal à  $m$  (voir équation (4. 1)).



**FIGURE IV. 3.:** Droite de régression des longueurs des racines du faisceau à  $t+4$  sur les longueurs des racines du faisceau à  $t$  (en vert). Longueurs mesurées de racines (en bleu).

La configuration présentée dans la figure précédente est la configuration:

$$\Theta_1 = \begin{bmatrix} 100 & 155 \\ 120 & 165 \\ 140 & 175 \\ 240 & 285 \end{bmatrix} \quad (4.5)$$

Pour cette configuration, on calcule  $R^2$  à partir du coefficient de corrélation:

$$R^2_{\Theta_1} = \left( \frac{COV[A, B]}{\sqrt{VAR[A]VAR[B]}} \right)^2 = 0,9831 \quad (4.6)$$

On retiendra dès lors la configuration pour laquelle  $R^2$  est maximal.

Cet indicateur peut être comparé d'une plante à l'autre car il varie dans un domaine borné :  $[0,1]$ . C'est donc celui qui sera retenu dans l'implémentation de l'algorithme.

#### *Pseudo-code de l'algorithme*

La fonction MATLAB® implémentée se nomme *corresvitR.m*. Elle admet trois inputs (*set1*, *set2* et *jours*) et renvoie quatre outputs (*tableau*, *maxim*, *vitesse* et *II*). Le code de cette fonction est disponible en version électronique sur le CD attaché. Cette fonction fait appel, entre autres, à une autre fonction *Kmatrix* (voir ultérieurement, dans cette section).

La difficulté de la méthode développée dans ce schéma réside dans l'implémentation du dénombrement des configurations de correspondance. Si l'on peut calculer aisément le nombre de configurations attendues au moyen de l'équation (4.3), il n'en demeure pas moins ardu de décrire chacune de ces configurations en terme d'objets la composant (i.e. de racines).

Par ailleurs, afin de rendre l'évaluation des correspondances plus rapide, ce nombre de configurations est réduit en éliminant les couples pour lesquels la mesure en seconde date (B) est inférieure à la mesure en première date (A). Cette réduction du problème est avantageuse dès que l'on travaille avec des nombres de racines supérieurs à 5. Par exemple, le nombre combinatoire correspondant à 8 racines à une date et 6 à une autre,  $A^8_2$ , égale 20160, c'est-à-dire autant de configurations à traiter pour une telle situation !

Les inputs de *corresvitR* sont *set1*, vecteur reprenant les longueurs racinaires, en millimètres, du faisceau à la première date, *set2*, vecteur reprenant les longueurs racinaires, en millimètres, du faisceau à la seconde date, et *jours*, la différence de temps entre le premier et le second jour.

#### Etape 1: *Traitement du vecteur set2.*

Le vecteur *set2* est expurgé de ses valeurs inférieures à toutes les valeurs de *set1*. Ces valeurs

correspondent plus que probablement à des racines apparues entre les deux dates de mesures (voir IV. 1. Hypothèse 2.). Etablir des correspondances entre celles-ci et celles de la première date est donc inutile, car les vitesses de croissance seraient négatives. En outre, cela permet de réduire le nombre de configurations possibles.

A la fin de cette étape, on assigne à chacun des éléments des vecteurs une identification qui est, simplement, sa position dans le vecteur, donc de 1 à  $k$  pour le vecteur le plus court et de 1 à  $n$  pour le vecteur le plus long. Il n'est pas possible, en effet, d'identifier chaque racine sur base de sa longueur car il est possible d'avoir plusieurs fois la même longueur.

Etape 2: Dénombrement des configurations possibles

Cette opération se réalise grâce à l'équation (4. 3).

Etape 3: Construction itérative des configurations.

De cette construction doit résulter un ensemble de matrices de dimension ( $k \times 2$ ), comportant en première colonne les identifications du vecteur le plus court, c'est-à-dire le vecteur-colonne (1, ...,  $k$ )' et en seconde colonne un arrangement de  $k$  identifications du vecteur le plus long parmi  $n$  identifications possibles. Par exemple, sachant que pour la situation développée plus haut  $k=4$  (à la seconde date) et  $n=5$  (à la première date), la configuration reprise en équation (4. 5) s'exprime donc :

$$\Theta_1 = \begin{pmatrix} 11 \\ 22 \\ 33 \\ 45 \end{pmatrix} \quad (4. 7)$$

La construction des configurations se résume donc à celle des vecteurs-colonnes ( $x_1, \dots, x_k$ )' dans lesquels chacun des éléments  $x_i$  sont différents les uns des autres et appartiennent à  $[1, \dots, n]$ .

Nous avons développé, pour y procéder, une fonction MATLAB® :  $K=Kmatrix(V, z)$ , pour laquelle  $V$  est une matrice ( $p \times q$ ), remplie avec des éléments  $x_i$ , et  $z$  est le vecteur (1, ...,  $n$ ). Le nombre  $q$  est toujours inférieur ou égal à  $n$ .

Soit  $d$  la différence ( $n-q$ ),  $Kmatrix$  produit une matrice  $K$  de dimension ( $dp \times q+1$ ) pour laquelle chaque ligne est un vecteur ( $1 \times q+1$ ) correspondant à une ligne de  $V$  complétée par chacun des  $d$  éléments de  $z$  qui ne se retrouvent pas dans cette ligne de  $V$ .

Le schéma de la page suivante traduit ces « borborygmes » en images.

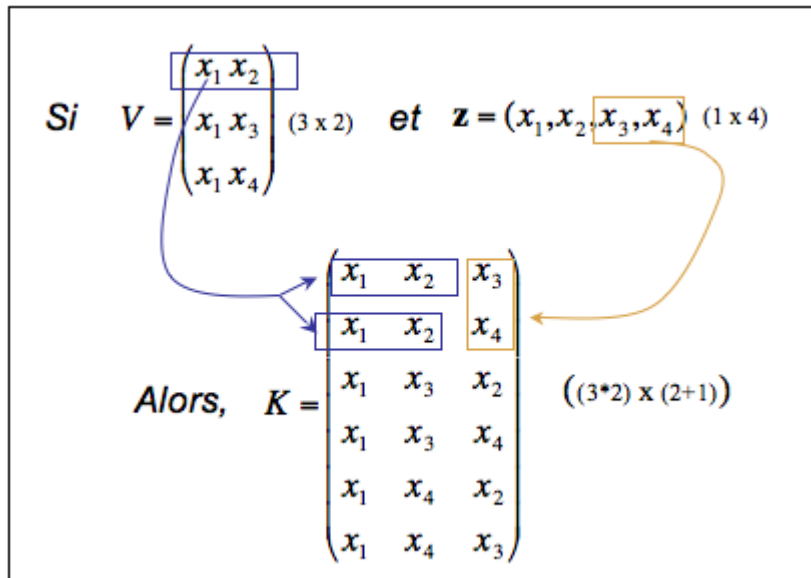


FIGURE IV. 4. : La fonction Kmatrix (V, z) : détail.

Le recensement des configurations utilise cette fonction en effectuant  $k$  itérations, suivant ces pas :

- (1)  $V = z, K = Kmatrix(V, z)$ ;
- (2)  $V = K, K = Kmatrix(V, z)$ ;
- (3) retour en (2) avec une nouvelle valeur de  $K$ .

Etape 4 : Evaluation de chaque configuration.

Chaque configuration est maintenant évaluée selon l'équation (4. 6). Il en ressort un vecteur objectif nommé *combival*.

Etape 5 : Recherche du maximum du vecteur *combival*.

La ou les configurations correspondant à ce vecteur sont celles qui minimisent la dispersion autour de la droite de régression. S'il devait y en avoir plusieurs, on en choisirait une arbitrairement.

Les outputs de la fonction *corresvit* sont:

- *tableau* : configuration optimale, donc structurée comme l'équation (4. 7);
- *maxim* : valeur maximale du vecteur *combival*, c'est-à-dire valeur de l'indicateur  $R^2$  pour la configuration reprise dans *tableau*;
- *vitesses* : vitesses de croissance en longueur des racines du faisceau, sur base de la configuration reprise dans *tableau* (vitesses exprimées en cm/jour);
- *ll* : longueurs racinaires de la date où l'on a  $n$  mesures selon l'arrangement optimal repris dans *tableau* (longueurs en mm).

La configuration optimale de la situation présentée se trouve ci-dessous sous forme graphique.

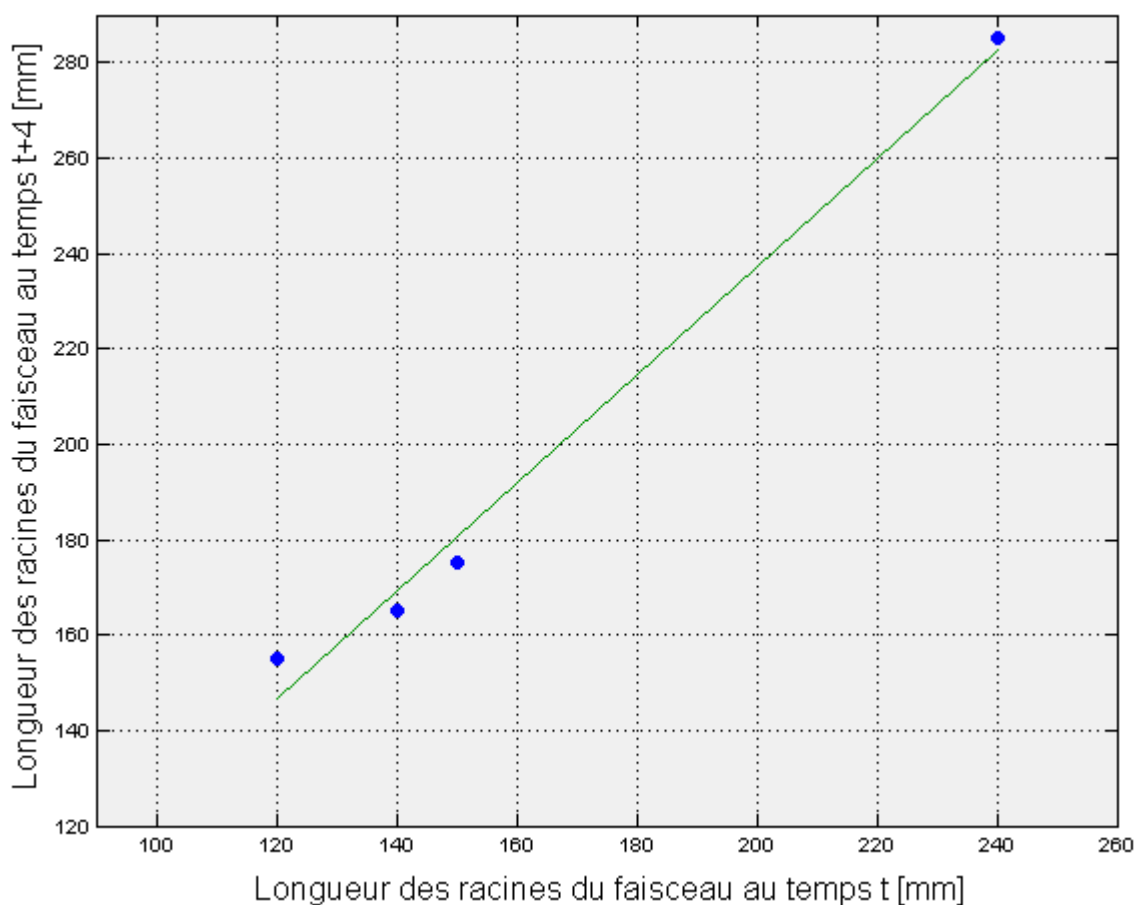


FIGURE IV. 5. : Droite de régression (en vert) et couples  $\{A_i B_j\}$  (en bleu) de la configuration optimale.

### IV.3. Cas particuliers et avatars de l'algorithme

#### *Cas particuliers.*

En fonction de la valeur des paramètres  $n$  et  $k$  (avec  $k$  toujours inférieur ou égal à  $n$ , voir section IV. 2. § *Schéma paradigmatique*), respectivement longueurs des vecteurs d'entrée  $set1$  et  $set2$ , il peut :

- ne pas y avoir de solution, si  $n \cdot k = 0$  ;
- y avoir une solution incalculable, si  $k=1$  et  $n \geq k$ ;
- y avoir une solution indécidable, si  $k=2$  et  $n \geq k$ .

La première restriction est évidente, la seconde et la troisième le sont moins.

Dans la seconde, le problème est incalculable. Prenons le cas où  $k=1$  et  $n=1$ . Nous avons donc, selon toute vraisemblance, une seule racine. Il est tout à fait possible d'en calculer la vitesse de



croissance et le problème de correspondance ne se pose pas (il y a une seule configuration possible). Or, l'algorithme présenté ci-dessus évalue une configuration sur son coefficient de détermination (équation 4. 6). S'il n'y a qu'un seul couple, ce coefficient ne peut donc pas être calculé. Donc, la solution existe; une correspondance pouvant s'établir *de facto*, mais l'algorithme ne peut la calculer.

Prenons ensuite le cas où  $n$  est supérieur à 1. Il y a donc plusieurs configurations, mais ces configurations sont chacune composée d'un seul couple, ce qui revient au problème de calcul posé par  $n=1$ . Dans cette situation-ci, le problème reste en outre entier, car l'algorithme est *a fortiori* incapable de donner une indication sur la configuration à choisir.

Enfin, dans la situation où  $k=2$  et  $n$  supérieur ou égal à  $k$ , l'algorithme teste une ou plusieurs configurations. Cependant, comme chacune des configurations comprend deux couples, le coefficient de détermination est toujours égal à 1. L'algorithme ne permet donc pas de prendre une décision sur les correspondances et la solution qu'il renvoie est arbitraire (telle qu'expliquée en Etape 5. du pseudo-code). Notons à cet effet que ce problème ne se rencontre pas avec l'indicateur  $e_{ij}$  car celui-ci se base sur une comparaison à partir de la vitesse moyenne du faisceau.

#### *Avatars*

Outre la version de l'algorithme de base comprenant l'indicateur  $e_{ij}$ , nous avons également développé un avatar excluant des couples improbables afin de faciliter le dénombrement et la construction des configurations. Cette version s'est avérée peut concluante et n'a pas été déboguée; elle n'est donc que citée.

#### IV.4. Conclusions.

L'algorithme développé et présenté dans ce chapitre n'a pas fait l'objet d'une mise en oeuvre sur des données observées émanant des expériences menées. La seule évaluation de son fonctionnement porte sur les différents cas possibles abordés pour sa construction. Ces tests de cas particuliers ont révélés des vitesses de croissance légèrement supérieures à celles calculées lors des expériences préliminaires. Ces vitesses de croissance sont de l'ordre de 35 – 40 mm/j pour les racines séminales et de l'ordre de 45 – 50 mm/j pour les racines adventives.

Néanmoins, l'implémentation de la méthode, et particulièrement celle de recherche des configurations possibles, a permis d'esquisser sa mise en oeuvre dans un problème élargi à plusieurs dates de mesure.

Ce dernier problème fera l'objet du chapitre suivant.

## *Chapitre V*

# *Correspondances sur plus de deux dates par calcul combinatoire*

Ce chapitre présente brièvement un élargissement de l'algorithme du Chapitre IV à plus de deux dates de mesure. La méthode de recensement et d'évaluation des configurations de correspondance demeure donc basée sur le calcul combinatoire.

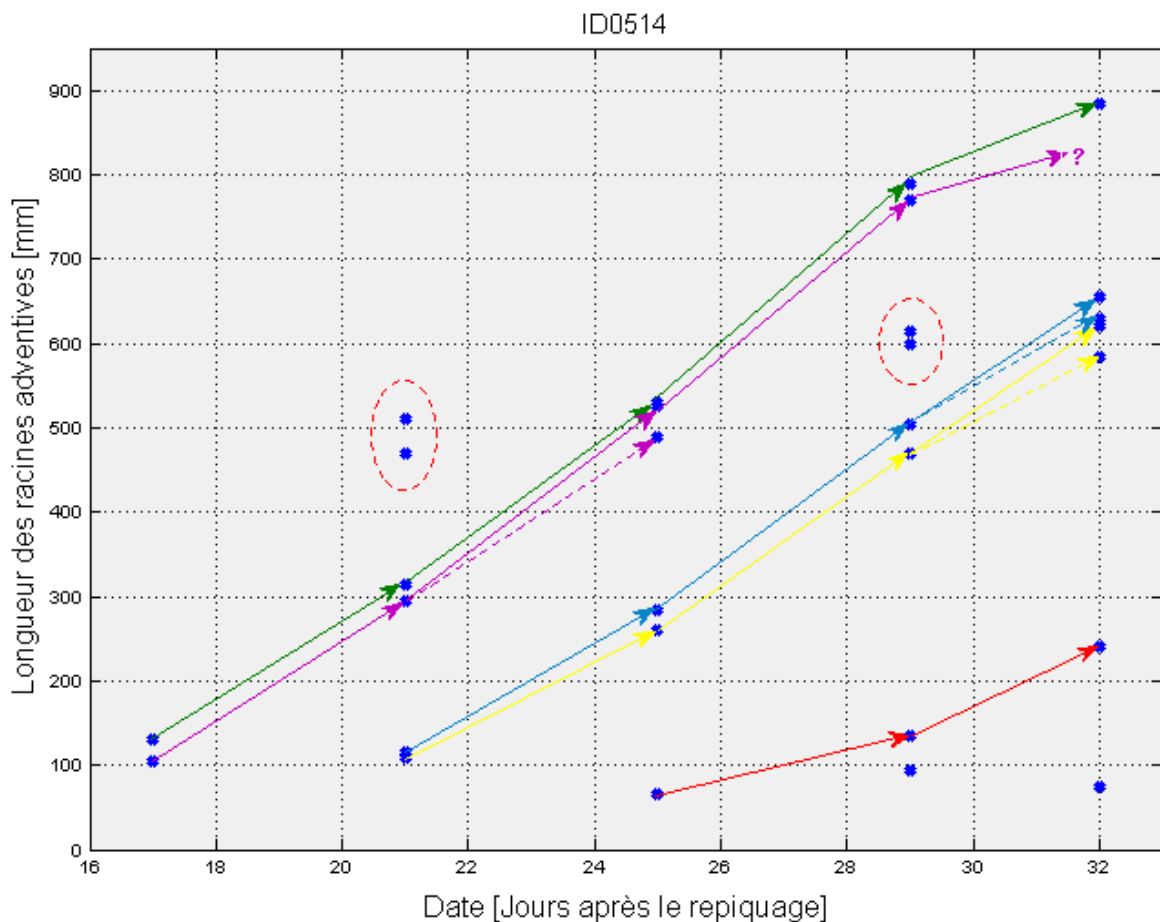
Cependant, afin de réduire le nombre potentiellement élevé de configurations générées, nous adjoignons à l'algorithme une restriction sur les configurations comprenant des couples de mesure dont les vitesses sont improbables. C'est la raison pour laquelle le dénombrement des configurations de correspondance se fera parallèlement sur les identifications et sur les longueurs de racines.

Dans ce chapitre, nous progresserons de la même manière que dans le précédent. D'abord, une mise en situation illustrée et les hypothèses de l'algorithme. Ensuite, l'algorithme paradigmatique de correspondance, à travers son schéma de base et son pseudo-code. Enfin, les cas particuliers et la conclusion.

## V. 1. Introduction et hypothèses.

### *Mise en situation*

Sur la figure V. 1. sont retranscrites les mesures d'un faisceau adventif à plusieurs dates successives (points bleus). Ces dates sont le 8 mars 2007 (17 ème jour après le repiquage), le 12 mars 2007, le 16 mars 2007 et le 20 mars 2007. Comme on peut le remarquer, le nombre d'adventives augmente progressivement au cours du temps. Il est par ailleurs possible d'indiquer l'évolution d'une racine de date en date, ce que nous faisons en les reliant au moyen de flèches.



**FIGURE V. 1.:** Evolution du faisceau adventif de la plante ID0514 du 8/03/07 (jour 17 après le repiquage) au 23/03/07 (jour 32 après le repiquage). Légende : → racines adventives retracées sur plusieurs dates (chaque couleur correspond à une racine), ● longueur d'une racine adventive.

Sur ce graphique, on peut observer que:

- Les flèches d'une couleur donnée retracent chacune l'évolution d'une racine, évolution supposée par l'expérimentateur, qui correspond donc à une combinaison de correspondance. Cette supposition peut amener à des indécisions (flèches pointillées).
- Certaines mesures de racines peuvent avoir été omises par l'expérimentateur. Ce cas est

indiqué par un point d'interrogation à la fin du traçage mauve.

– Certaines mesures peuvent être parasites. C'est le cas des mesures indiquées par les deux ellipses rouges, néanmoins plus flagrantes au jour 21 qu'au jour 29 car elles se trouvent, dans le premier cas, hors de l'ensemble de mesures effectuées ce jour-là.

– Les combinaisons tracées portent sur 3, 4 ou 5 dates.

On peut interpréter les mesures parasites du jour 21 comme étant celles de racines séminales considérées par l'expérimentateur comme appartenant au faisceau adventif. Celles du jour 29 sont plus difficiles à interpréter. Soit il s'agit également de séminales, soit il s'agit effectivement d'adventives, qui n'ont cependant pas été relevées aux autres dates, ce qui est moins probable.

De ces observations, on peut déduire des contraintes équivalentes à celles énoncées au Chapitre IV. 1. En effet,

(1) L'ordre du classement suivant la longueur n'est pas forcément conservé, en raison notamment des parasites;

(2) de nouvelles racines apparaissent (cfr. traçage rouge ci-dessus);

(3) d'autres racines disparaissent (cfr. traçage mauve ci-dessus);

La contrainte (2) ajoute une quatrième contrainte:

(4) l'évaluation de la configuration de correspondance d'une plante donnée s'effectue par évaluations successives des combinaisons de correspondance ayant le même nombre d'éléments, donc la même longueur. C'est-à-dire, combinaisons de longueur 5 pour les racines tracées mauve et verte, 4 pour les bleue et jaune et 3 pour la rouge.

### *Hypothèses*

1. Hypothèse de distribution des longueurs composant une combinaison : les longueurs mesurées, composant une combinaison donnée, sont normalement distribuées autour d'une droite de régression linéaire de ces longueurs sur le temps. La combinaison est d'autant plus probable que cette distribution est peu dispersée.

2. Hypothèse sur les combinaisons : Pour former une combinaison d'une certaine longueur, donc d'un certain nombre d'éléments, il faut un élément à chaque emplacement de cette combinaison. En d'autres termes, dans une combinaison donnée, il ne peut y avoir d'emplacement vide.

– *Hypothèse corrolaire* : le nombre maximal de racines potentiellement retraçables est le nombre de mesures à la date où il y en a le moins. On supposera par la suite que c'est à la première date de la combinaison qu'on en aura le moins.

3. Hypothèse de non-régression: les racines apparues à une certaine date ne peuvent en aucun cas être plus courtes à la date de mesure suivante.

4. Hypothèse sur les couples candidats: une certaine mesure sera couplée à une autre de la date suivante pour autant que la vitesse de croissance résultant du couple formé par ces deux mesures soit comprise dans un intervalle délimité ]5 ; 65[ (vitesses de croissances exprimées en mm/jour).

5. Hypothèse d'unicité de la configuration optimale: la configuration de correspondance choisie comme optimale sera unique. Si plus d'une devaient être optimales, on choisirait arbitrairement la première d'entre elles.

Il importe de bien comprendre que l'hypothèse 1., portant sur la normalité de la distribution des longueurs, est sensiblement différente de l'hypothèse de normalité posée dans le chapitre précédent. Lorsque, dans ce dernier, on se basait sur une vitesse identique pour toutes les combinaisons binaires possibles, en l'occurrence, sur la vitesse moyenne du faisceau, on se base ici sur une vitesse qui peut être différente pour chaque combinaison. La probabilité de l'évènement « *Correspondre effectivement à une racine* » se calcule non pas sur la dispersion autour d'une vitesse identique, mais sur la dispersion d'une vitesse différente pour chaque combinaison.

En outre, l'utilisation de la régression linéaire implique qu'une combinaison n'est évaluable que si elle contient trois mesures.

Les hypothèses 2. (sur les combinaisons) et 4. (sur les couples candidats) permettent de réduire le nombre de configurations de correspondance valables car elle n'admettent comme combinaisons valables que celles qui résultent de couples ayant une vitesse de croissance comprise dans un certain intervalle.

Une particularité majeure de l'algorithme proposé ici est de générer les configurations de correspondances concomitamment en termes d'identifications de racines (ce que nous faisons au Chapitre IV) et en termes de longueurs de ces racines, afin de pouvoir évaluer les vitesses de croissance des couples de correspondances proposés.

## V. 2. Algorithme paradigmatique de correspondance

### *Schéma paradigmatique*

Soit le faisceau mis en situation ci-précédemment (faisceau adventif de la plante ID0514). La matrice des données mesurées est retranscrite dans le Tableau V. 1., dans lequel 0 et NaN (Not-a-Number) sont les deux valeurs indicatrices d'absence de mesure. L'utilisation de deux valeurs

d'absence provient du pré-traitement de l'information (Chapitre III. 5. e.). Nous y avons en effet utilisés deuxvaleurs. Initialement, NaN indique que la mesure n'existe pas et 0, qu'elle vaut 0. Or, une mesure de valeur 0 n'est pas possible avec notre règle graduée en codes-barres.

TABLEAU V. 1.: *Matrice des données mesurées ( $M_1$  à  $M_8$ ) du faisceau adventif de la plante ID0514.*

	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$	$M_7$	$M_8$
<b>08-Mar-07 (A)</b>	105	130	0	0	0	0	0	0
<b>12-Mar-07 (B)</b>	110	115	295	315	470	510	NaN	NaN
<b>16-Mar-07 (C)</b>	65	260	285	490	525	530	0	0
<b>20-Mar-07 (D)</b>	95	135	470	505	600	615	770	790
<b>23-Mar-07 (E)</b>	75	240	585	620	630	655	885	0

On peut supposer que les mesures  $A_1$  et  $A_2$ , c-à-d les mesures de la première date, révèlent chacune une racine (en vert). C'est là l'application de l'hypothèse 2.

Sur base de l'hypothèse 3., sont considérées, dans ce tableau, comme étant des nouvelles racines apparues entre deux dates de mesure les racines inférieures à toutes celles retenues à la date précédente (en bleu). Ces mesures sont dès lors éliminées du tableau.

Sur base de l'hypothèse 4., on peut également deviner quelles mesures (en orange) n'entreront pas dans les combinaisons valables, car situées hors de l'intervalle ]5 ; 65[ (mm/j) sur les vitesses de croissance. Cette sélection simplifie le problème de mise en correspondance, néanmoins, elle n'intervient que lors de la génération des combinaisons, en posant des emplacements vides dans ces combinaisons, donc, en les invalidant. Par exemple, en date B, 470 mm est supérieur à 390 mm. 390 mm s'évalue en faisant le calcul 130 mm + 4 jours \* 65 mm/jour.

Par ailleurs, chacune des mesures de ce tableau reçoit un identificateur correspondant à la colonne dans laquelle elle se trouve. L'élaboration des configurations se fera simultanément sur ces identifications et sur les longueurs. Une configuration valide est donc double; chaque version étant construite comme ceci, si on a  $x$  dates de mesures  $A, B, \dots, X$ , auxquelles on a réalisé respectivement  $na, nb, \dots, nx$  mesures:

$$\forall xa_1 \neq \dots \neq xa_{na}, \forall xb_1 \neq \dots \neq xb_{nb}, \dots, \forall xx_1 \neq \dots \neq xx_{nx},$$

$$\Theta = \begin{bmatrix} A_{xa_1} & B_{xb_1} & \dots & X_{xx_1} \\ A_{xa_2} & B_{xb_2} & \dots & X_{xx_2} \\ \dots & \dots & \dots & \dots \\ A_{xa_{na}} & B_{xb_{nb}} & \dots & X_{xx_{nx}} \end{bmatrix} \quad (5.1)$$

$A, B, \dots, X$  représenteront soit des identifications soit des longueurs de racines en millimètres.

Le dénombrement de toutes les configurations possibles se fait en deux étapes.

Premièrement, le dénombrement de toutes les combinaisons de correspondances faisables, c'est-à-dire de racines faisables. Lors de cette étape, on générera toutes les combinaisons qui ne contiennent pas d'emplacement vide. La base de construction de ces combinaisons est l'ensemble des mesures réalisées à la dernière date de mesure. Nous allons donc opérer rétrospectivement, supposant que chaque mesure de la dernière date plus longue que plus d'une mesure de l'avant-dernière date indique une racine retraçable.

La fonction, développée en MATLAB®, *KLmatrix.m* dénombre les combinaisons à partir des données du Tableau VI. 1. et les organise en un ensemble de deux matrices  $\{R_1, \dots, R_z\}$ , pour les mesures, et  $\{I_1, \dots, I_z\}$ , pour les identifications; chaque matrice ayant comme élément de départ l'une des mesures (ou identification) de la dernière date, donc  $z \leq nx$ . Donc, chaque matrice représente toutes les « rétro-évolutions » possibles d'une racine mesurée en dernière date. Cette fonction, dont le schéma de fonctionnement est proche de celui de *Kmatrix* (développée au Chapitre IV.), applique par ailleurs le filtre sur les vitesses (cfr. hypothèse 4. sur les couples candidats). Par exemple, dans la situation qui nous occupe, la première des matrices de mesures renvoyée par *KLmatrix* se développe comme ceci:

$$R_1 = \begin{pmatrix} 885 & 770 & 525 & 295 & 105 \\ 885 & 770 & 525 & 295 & 130 \\ 885 & 770 & 525 & 315 & 105 \\ \dots & & & & \\ 885 & 770 & 530 & 315 & 130 \end{pmatrix} \quad (5.2)$$

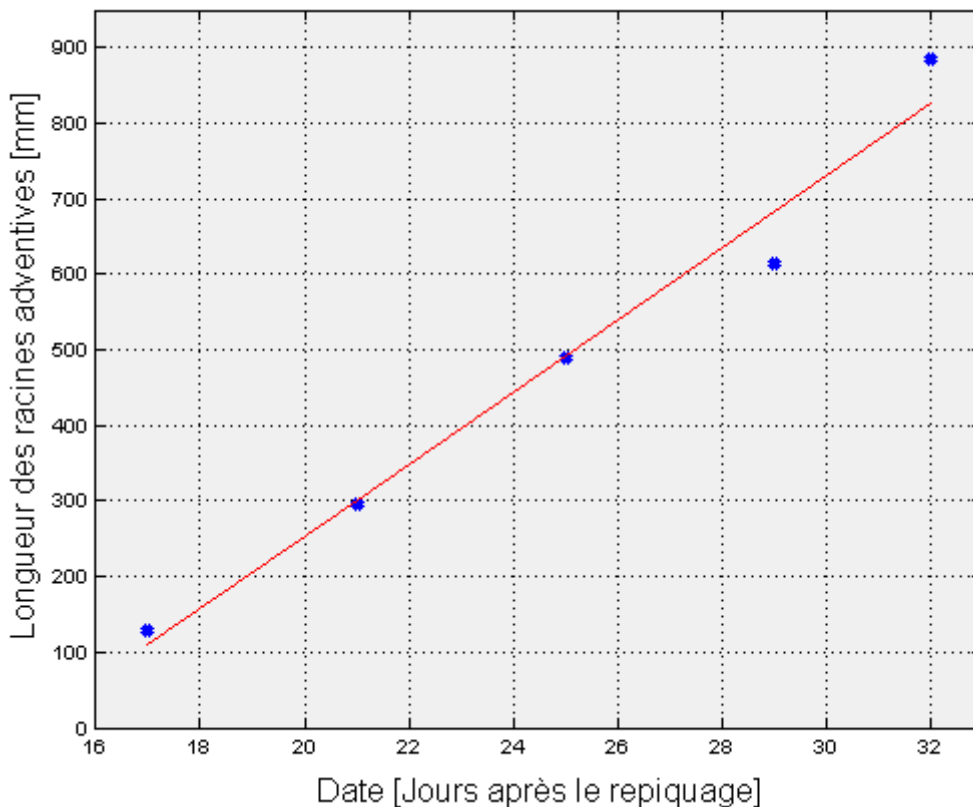
La rétro-évolution débutant par (885, 790, ...) n'est pas reprise car mènerait à des emplacements vides en troisième position. En effet,  $530+65*4 \geq 790$ ; la vitesse de croissance de ce couple invalide sa candidature. Rappelons aussi que cet ensemble de matrices est double : pour les mesures (comme ci-dessus) et pour les identifications.

Deuxièmement, cette paire de matrices  $\{R_1, \dots, R_z\}$  et  $\{I_1, \dots, I_z\}$  est soumise à l'élaboration des configurations. Dans cette seconde étape, plusieurs combinaisons sont rassemblées en une même configuration pour autant qu'elle ne contiennent pas de mesures identiques (exclusion basée sur les combinaisons d'identifications). Cette condition est vérifiée par la fonction *Rmatrix.m*, qui, à partir de 4 matrices  $R_i, R_j, I_i$  et  $I_j$  (doubles deux à deux), renvoie deux combinaisons pour lesquelles les mesures contenues dans l'une ne le sont pas dans l'autre. Par itérations successives, à partir de deux combinaisons, on construit une configuration valide la plus complète possible. Le nombre de configurations retenues peut éventuellement être fort minime par rapport au nombre de combinaisons possibles en raison de l'exclusion des mesures identiques.

Enfin, chacune de ces configurations est évaluée au moyen de l'indicateur développé dans la section suivante. On retient alors la meilleure de ces configurations.

*Indicateur implémenté*

Comme déjà mentionné dans les hypothèses de ce chapitre, l'évaluation d'une combinaison donnée est totalement déterminée par cette combinaison car il s'agit du coefficient de détermination de la droite de régression linéaire de la combinaison sur le temps. La figure suivante illustre ce calcul.



**FIGURE V. 2.** : Droite de régression linéaire d'une combinaison donnée [130, 295, 490, 615, 885] sur le temps. Légende : en bleu, éléments (mesures) de la combinaison; en rouge, droite de régression.

Le coefficient de détermination  $R^2$  se calcule par l'équation (4. 6).

Pour une configuration donnée, l'indicateur  $\Omega = f(\Theta)$  n'est autre que le produit des coefficients de détermination des combinaisons qui la composent. Donc, si  $\Theta$  est une configuration composée de trois combinaisons  $c_1$ ,  $c_2$  et  $c_3$ , c'est-à-dire trois racines retracées, contenant  $x$  mesures, alors:

$$\Theta = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \quad \text{et} \quad \Omega_{\Theta} = R^2_{c_1} \cdot R^2_{c_2} \cdot R^2_{c_3} \quad (5. 3) \text{ et } (5. 4)$$



### *Pseudo-code de l'algorithme*

La fonction MATLAB® implémentée se nomme *corresvitnR.m*; elle admet trois inputs (*sets*, *dates* et *date0*) et renvoie 4 outputs (*valmax*, *modelmax*, *configINDmax* et *configRACmax*). Le troisième input permet le renvoi d'une option graphique; c'est une amélioration que nous détaillerons dans la section V. 4.

Avant de passer en revue les différentes étapes du pseudo-code, il importe de remarquer que cette fonction, n'ayant pas fait l'objet d'une évaluation sur l'ensemble des données de mesures racinaires disponibles mais seulement sur quelques faisceaux adventifs, n'a pas été déboguée. Des erreurs imputables aux valeurs d'entrée pourraient en effet encore survenir.

Le code est toutefois disponible sur le CD attaché à ce mémoire.

La fonction *corresvitnR* nécessite d'autres sous-fonctions *indic.m*, *KLmatrix.m*, *Rmatrix.m* et *choixmodele.m*.

La difficulté de l'implémentation de cet algorithme est analogue à celle de l'algorithme du Chapitre IV. Il s'agit une fois encore du dénombrement des combinaisons. La résolution de ce problème fait l'objet de la fonction *KLmatrix* (comme expliqué précédemment) qui, en outre, applique le filtre des vitesses sur les couples candidats.

Les inputs de *corresvitnR* sont:

- *sets* : si  $x$  est le nombre de dates de mesures et  $nx$  le nombre de mesures à la date où il y en a le plus, *sets* est une matrice ( $x \times nx$ ) reprenant les données de mesures racinaires telles que présentées dans le Tableau V. 1. Aux emplacements vides, on retrouve soit la valeur NaN (Not-a-Number) soit 0. Cette dernière valeur n'est rendue possible que par le seul fait que notre dispositif expérimental (voir Chapitre III.) ne permet pas de mesurer des longueurs de 0 mm.
- *dates* : vecteur ( $x \times 1$ ) de  $x$  dates formatées comme ceci : *jj-(abréviation du mois)-AA*, dans lequel format l'abréviation du mois rassemble les trois premières lettres du mois en anglais (Jan, Feb, Mar, Apr, ..).

#### Etape 1: Choix des valeurs candidates.

A partir de l'input *sets* revisité (réarrangement si nécessaire des différentes lignes en ordre ascendant), son double contenant les identifications de racines est créé. Ces identifications correspondent aux positions de chaque mesure dans la ligne où elle se trouvent. Ces deux matrices (mesures et identifications) sont alors expurgées des valeurs correspondantes à des racines apparues

entre deux dates de mesure. Cette opération est réalisée par la sous-fonction *indic*. A terme, deux matrices *sets\_cand* (mesures) et *indsets\_cand* (identifications) contiennent les valeurs candidates, c'est-à-dire les mesures et identifications à partir desquelles on pourra établir des combinaisons.

Etape 2: Recensement des combinaisons possibles.

Ce recensement se base sur les matrices résultant de l'étape précédente et sur le filtre des vitesses. La fonction *KLmatrix* effectue ce recensement d'une manière, analogue à *Kmatrix*, déjà évoquée dans le schéma paradigmatique. Elle renvoie deux ensembles de matrices  $\{R_1, \dots, R_z\}$  et  $\{I_1, \dots, I_z\}$ , contenant respectivement mesures et identifications de racines, construites sur le modèle de l'équation (5. 2).

Etape 3: Construction des configurations possibles.

Cette étape a également déjà été détaillé dans le schéma paradigmatique. Elle s'appuie sur la fonction *Rmatrix*, qui renvoie une collection de paires de matrices dans laquelle chaque paire reprend les deux versions (mesure et identifications) d'une même configuration valide telle que décrite en équation (5. 1).

Etape 4: Evaluation des configurations valides.

La fonction *choixmodele* renvoie, pour une configuration valide donnée, le coefficient angulaire et l'intercept de la régression linéaire de chaque combinaison la composant sur le temps (une matrice  $(n \times 2)$ , pour laquelle  $n$  est le nombre de racines retracées, donc de combinaisons composant la configuration) ainsi que le coefficient de détermination de chaque régression (vecteur  $(n \times 1)$ ). En multipliant pour chaque configuration ces coefficients, on obtient l'indicateur décrit plus haut en équation (5. 4).

Au terme de cette étape, un vecteur contenant une valeur d'indicateur pour chaque configuration valide est obtenu.

Etape 5: Choix de la configuration optimale.

La valeur maximale du vecteur obtenu à la fin de l'étape 4 est celle correspondant à la configuration la plus probable, c'est-à-dire la configuration qui minimise la dispersion autour des combinaisons correspondant à des racines retracées sur plusieurs dates de mesures (et ayant une mesure à chacune de ces dates).

La fonction *corresvitnR* renvoie 4 outputs:

- *modelmax* : cette matrice ( $n \times 2$ ) rassemble, à chaque ligne, le coefficient angulaire et l'intercepte de la droite de régression d'une des  $n$  combinaisons de la configuration optimale.
- *valmax* : ce nombre est la valeur de l'indicateur calculée au moyen de l'équation (5. 4).
- *configRACmax* : cette matrice ( $n \times x$ ) reprend à chaque ligne, dans l'ordre décroissant, les mesures d'une des  $n$  combinaisons de la configuration optimale. L'ordre décroissant provient du traitement des combinaisons par *KLmatrix*.
- *configINDmax* : cette matrice ( $n \times x$ ) est la paire jumelle de la matrice *configRACmax* : elle contient à chaque ligne les identifications de racines (c-à-d positions dans une ligne de l'input *sets*) d'une des  $n$  combinaisons de la configuration optimale.

### V. 3. Cas particuliers et avatars.

#### *Cas particuliers*

Comme déjà mentionné précédemment (V. 1. *Hypothèses*), la configuration de correspondance doit porter sur un minimum de trois dates de mesures. En effet, sur le problème ne se pose pas pour une seule date et est indécidable pour deux dates car le coefficient de détermination de toutes les combinaisons possibles sera toujours 1.

#### *Avatars*

Plusieurs avatars de cet algorithme ont été envisagés, implémentés et testés sur quelques faisceaux adventifs.

(1) Ajout d'une option graphique. Cette option est rendue possible par les inputs *optgr*, *optgr2* et *date0* de la fonction de base *corresvitnR*. *optgr* est un nombre 0 (désactivation) ou 1 (activation de l'option graphique, *optgr2* peut également prendre comme valeurs 0 (utilisation d'une nouvelle figure) ou 1 (utilisation d'une figure préexistante) et *date0* est une date de référence choisie par l'utilisateur, formatée de la même façon que l'input *dates* (voir plus haut). L'option graphique consiste en l'impression à l'écran d'une figure similaire à la Figure V.2. pour l'ensemble des combinaisons de la configuration optimale. La Figure V. 3. de la page suivante est cette représentation graphique de la configuration optimale de la situation décrite dans le Tableau V. 1.

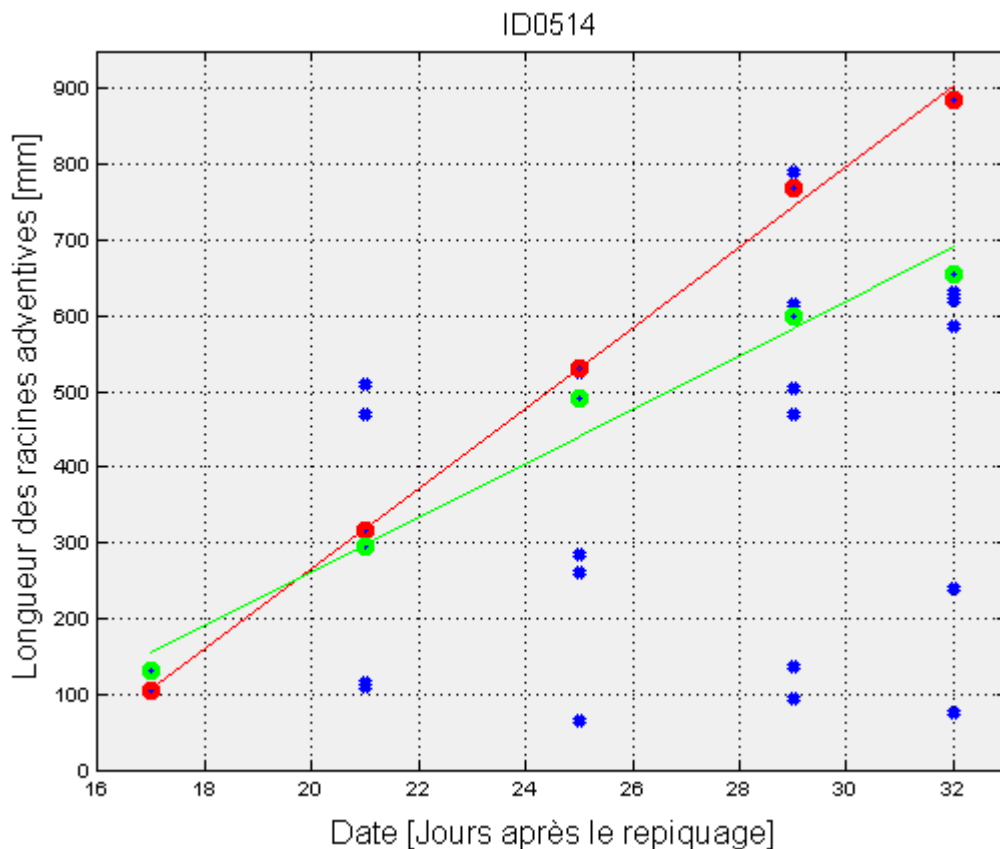


FIGURE V. 3.: Représentation graphique de la configuration optimale. Légende: droites rouge et verte : droites de régression linéaire des combinaisons de la configuration optimale; ● et ●: mesures composants les combinaisons rouge et verte; ●: longueur d'axes racinaires.

(2) Recherche de la vraisemblance sur d'autres modèles de croissance. Cet avatar consiste en une version différente de la sous-fonction choixmodele. Dans cette version, on évalue non seulement le coefficient de détermination de la régression linéaire d'une combinaison sur le temps mais également le coefficient de détermination de la régression logarithmique d'une combinaison donnée  $\mathbf{c}=[A_i, B_j, \dots, X_z]$  sur le temps  $\mathbf{t}$  :

$$\tilde{c} = \beta_1 \cdot t^{\beta_2} \quad (5.5)$$

Ce coefficient de détermination évalue l'adéquation de la combinaison à un modèle de croissance logarithmique (si  $\beta_2 < 1$ ) ou exponentiel (si  $\beta_2 > 1$ ) plutôt qu'à un modèle linéaire. Cependant, de tels modèles de croissance sont vraisemblablement inadaptés aux faisceaux que nous avons mesurés. C'est pour cette raison que nous avons délaissé cette évolution de l'algorithme.

## V. 4. Evaluation et conclusions

Comme il a déjà été dit plus haut, la méthode développée dans ce chapitre n'a été évaluée, ou plutôt testée, que sur quelques faisceaux adventifs. Ces tests ont permis de régler des problèmes de langage et de vérifier que les outputs correspondaient à nos attentes. Ils n'ont porté que sur des faisceaux mesurés sur 5 dates, ce qui correspond au maximum de dates possibles en ce qui nous concerne. C'est pour cette raison qu'on se doit de signaler que la fonction *corresvitnR* n'est pas totalement déboguée, car des erreurs pourraient encore survenir, notamment dans certaines conditions limites (3 dates par exemple).

Si l'on voulait poursuivre une évaluation avec cet algorithme, ce débogage serait nécessaire non seulement pour les raisons évoquées plus haut mais surtout parce qu'il ne permet pas de rendre compte de toutes les racines retraçables. En effet, comme cela a été évoqué dans la contrainte (4) de la situation développée en section V. 1., la fonction *corresvitnR* ne renvoie la configuration de correspondance que sur un nombre, nécessairement supérieur ou égal à 3, de dates donné préalablement. C'est donc une « semi-configuration » que l'on obtient, une partie de la configuration de correspondance optimale recherchée. Peu s'en faut pour parvenir à cette dernière; il suffit, en effet, de réitérer la fonction successivement sur un nombre décroissant de dates, en ayant pris soin, après chaque itération, d'ôter du tableau des mesures les mesures déjà répertoriées comme appartenant à une racine. Nous n'avons cependant pas développé la fonction jusqu'à ce point, délaissant le dénombrement et le recensement des combinaisons et configurations jugés trop lents au profit d'une méthode d'optimisation développée dans le chapitre suivant.

Enfin, dans le cadre de l'évaluation de cette fonction, deux critiques sont pertinentes.

Premièrement, on peut discuter de la place du filtre des vitesses. Le placer lors du dénombrement des combinaisons nous a paru judicieux car bien que cette façon de procéder amène à évaluer plus de combinaisons qu'il n'en faut, cela évite de recenser à deux reprises les couples candidats. Néanmoins, une autre stratégie serait bienvenue pour éviter cette évaluation surnuméraire.

Deuxièmement, la nécessité d'avoir une mesure à chaque date pour établir une combinaison valide pourrait s'avérer fort contraignante. De fait, il ne faut pas perdre de vue que certaines mesures peuvent être considérées comme parasites tandis qu'à certaines dates, une racine a pu ne pas être mesurée. Cette double contrainte (énoncées dans l'introduction de ce chapitre) peut induire l'algorithme *corresvitnR* en erreur en validant une combinaison contenant une mesure fautive ou en invalidant une combinaison parce qu'elle ne contient pas une mesure (attendue vraie).

## *Chapitre VI*

# *Correspondances sur plus de deux dates par optimisation des vitesses de croissance*

Ce chapitre utilise un algorithme de base différent de ceux développés et implémentés dans les chapitres précédents. En effet, la façon de procéder est inversée: il s'agira, dans ce cas-ci, de d'abord proposer des vitesses et ensuite, de déduire quelles combinaisons sont les plus probables sur base de ces vitesses. L'algorithme paradigmatique implémenté fait donc appel à une fonction MATLAB® de recherche de minimum de fonction objectif. Cela implique le développement d'une fonction objectif, ou indicateur, qui fera l'objet d'une discussion particulière.

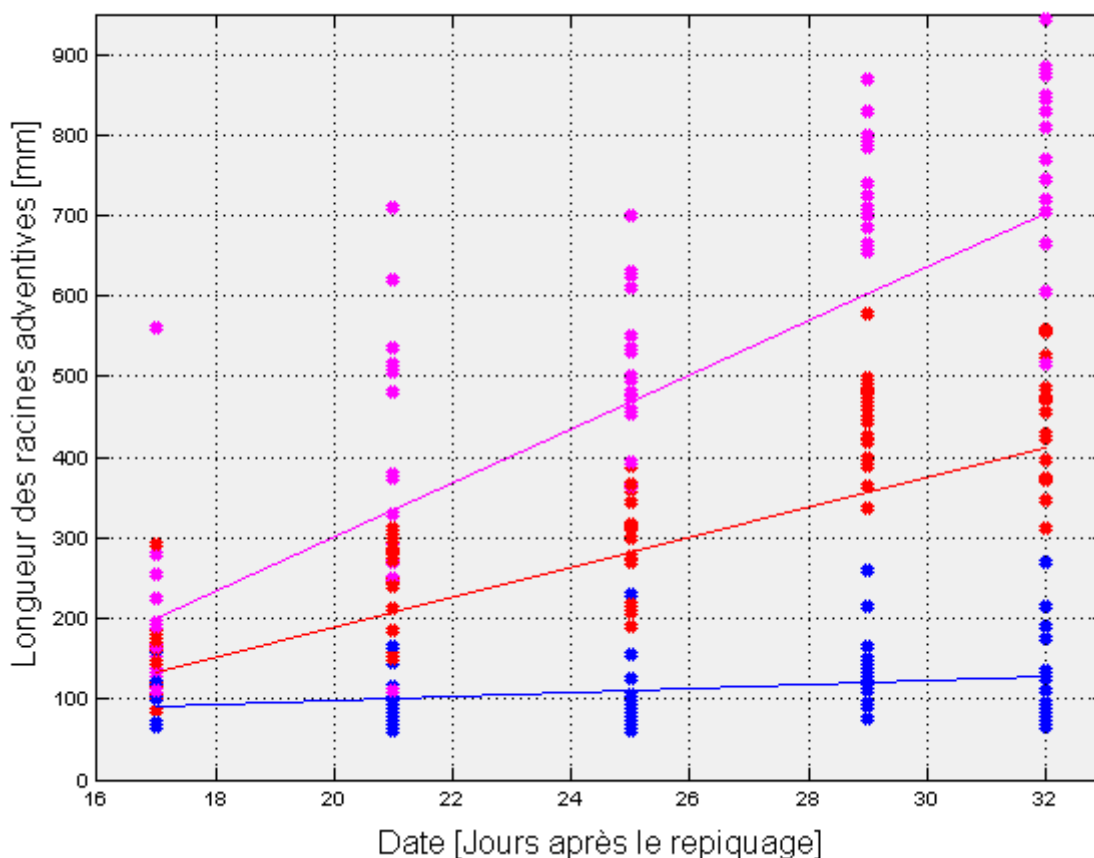
La version finale de l'algorithme est le résultat de plusieurs avatars, traitant chacun d'un cas particulier. L'ensemble des données sont évaluées sur base de cette version. Les vitesses de croissance et dates d'initiation de certaines racines prédites lors de cette évaluation sont discutées dans la dernière section de ce chapitre.

Ce chapitre est structuré de la même manière que les précédents. Après une mise en situation et les hypothèses posées dans l'introduction, vient le développement de l'algorithme paradigmatique, précédant les cas particuliers et, surtout, les avatars, dont est issue la version finale de cet algorithme, version utilisée pour prédire des vitesses de croissances soumises ensuite à une analyse de corrélation génétique (Recherche de QTL effectuée par Jean-François Dumasy). En dernier lieu, enfin, nous concluerons en discutant de ces implémentations et évaluations.

## VI.1. Introduction et Hypothèses

### *Mise en situation*

Sur la figure suivante (Figure VI. 1.), sont reprises des statistiques de 40 faisceaux adventifs provenant des plantes ID0495 à ID0534, mesurés à cinq dates (8/03/07, 12/03/07, 16/03/07, 20/03/07 et 23/03/07). Ces statistiques sont représentées par un point à chaque date. Il s'agit de la valeur maximale mesurée (en mauve sur le graphique), la valeur moyenne des mesures (en rouge sur le graphique) et la valeur minimale des mesures (en bleu sur le graphique). En outre, à chaque date, la valeur moyenne de la distribution de chaque statistique est calculée et une régression linéaire de cette valeur moyenne sur le temps est réalisée. Ces trois régressions linéaires sont les droites mauve (régression sur la moyenne des valeurs maximales à chaque date), rouge (régression linéaire sur la moyenne des valeurs moyennes à chaque date) et bleue (régression linéaire sur la moyenne des valeurs minimales à chaque date).



**FIGURE VI. 1. :** Evolution sur 5 dates des mesures effectuées sur 40 faisceaux adventifs (plantes ID0595 à ID0534). **Légende:** ● : mesure maximale d'un faisceau, ● : mesure moyenne d'un faisceau, ● : mesure minimale d'un faisceau.

D'apparence anodin, ce graphique peut néanmoins nous confirmer deux informations sur les

faisceaux adventives, déjà entrevues au chapitre précédent. Premièrement, on voit que la dispersion des mesures augmente avec le temps. Cela s'explique simplement par un plus grand nombre de racines de date en date; on retrouve des racines courtes à chaque date mais les plus longues le sont de plus en plus, d'où l'effet sur la dispersion. Deuxièmement, on remarque que les droites de régression indiquent une croissance marquée pour les moyennes de valeurs maximales et moyennes (coefficients angulaires respectivement de 38,16 mm/j et 18,57 mm/j) tandis que la régression linéaire sur la moyenne des valeurs minimales est moins, voire pas du tout, marquée. Le coefficient angulaire de cette dernière est égal à 2,36 mm/j, ce qui se situe hors de l'intervalle ]5, 65[ considéré comme filtre au chapitre précédent.

Elle permet en outre d'élargir le corrolaire de l'hypothèse 2. du chapitre précédent. Celle-ci supposait, pour nos développements ultérieurs, que l'on puisse considérer les mesures de la première date de mesure comme correspondant à autant de racines différentes. Sur base de la seconde observation, nous pouvons donc déterminer *a priori* un nombre correct de racines mesurées dans le faisceau.

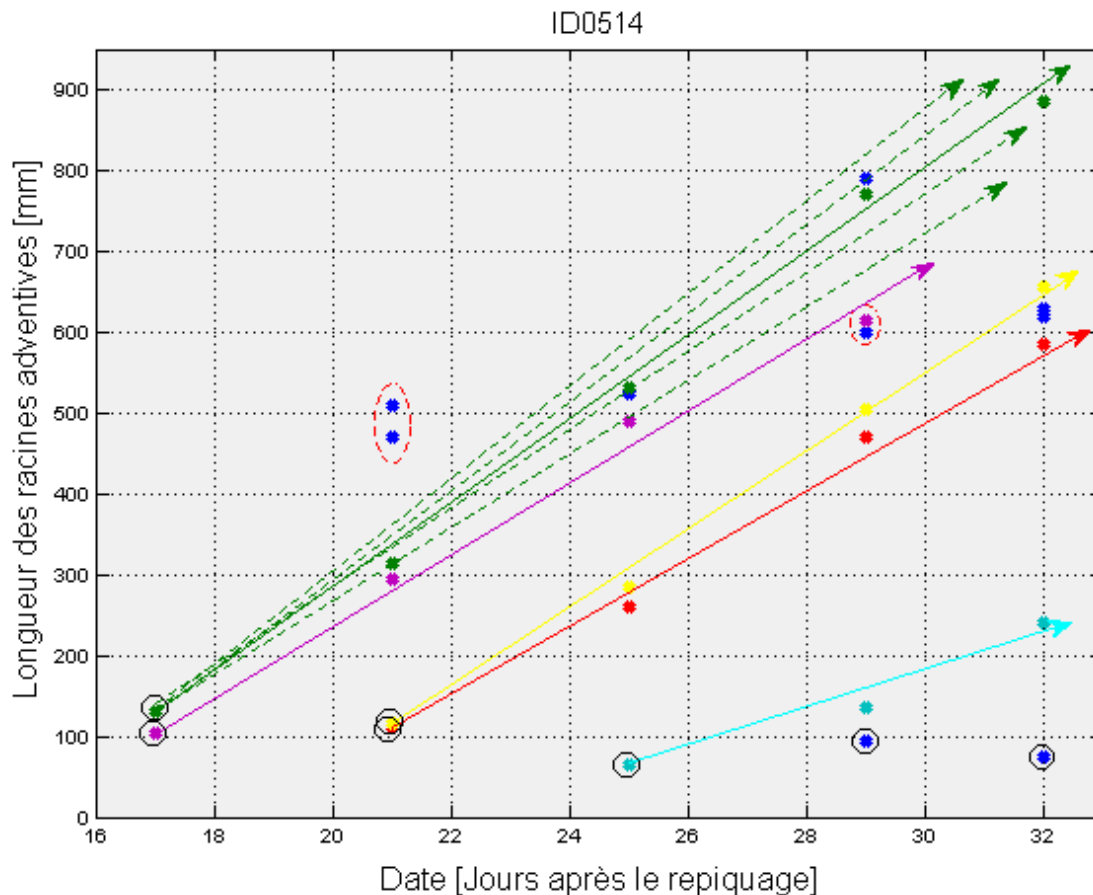
Ajoutons à cela que les mesures parasites, qui étaient un point d'achoppement relevé et non réglé dans le chapitre précédent, sont des erreurs expérimentales correspondant soit à une mesure nulle (erreur expérimentale : aucun objet n'est identifié mais une mesure est enregistrée) soit à une mesure d'un objet mal identifié. Ce second cas est plus probable que le premier car l'expérimentateur se trouve, lorsqu'il effectue ses mesures, en présence de deux faisceaux, l'adventif (à mesurer) et le séminal (à éviter). Des erreurs d'identification de racines ont donc pu survenir. Cependant, on s'attend à ce qu'elles soient plus fréquentes en ce qui concerne les mesures élevées car le faisceau séminal est en moyenne plus long que le faisceau adventif et, surtout, n'émet plus de nouvelles racines. Cela corrobore donc également le fait de considérer les mesures sous un certain seuil comme étant autant de nouvelles racines.

Cette nouvelle hypothèse, c'est-à-dire la considération de certaines mesures à chaque date comme autant de racines différentes, peut s'illustrer (Figure VI. 2., page suivante) sur le faisceau adventif mis en situation au chapitre précédent (plante ID0514). Considérer certaines mesures comme autant de racines différentes et savoir, sur base des évaluations réalisées aux chapitres précédents, à quelles vitesses de croissance on peut s'attendre permet en outre de supposer l'évolution de ces racines au fil du temps. Une configuration raisonnable sur base de cette supposition est illustrée par l'ensemble des droites verte, mauve, jaune, rouge et cyan sur la figure de la page suivante.

Chacune des droites représentées constituent donc une combinaison de correspondance dont les éléments sont les mesures de même couleur que la droite. On observe que ces combinaisons sont



toutes valides au sens qui a été donné à ce mot, c'est-à-dire ne contiennent pas d'emplacement vide.



**FIGURE VI. 2. :** Configuration possible du faisceau adventif de la plante ID0514. Les droites pleines représentent chaque combinaison, débutant en une mesure « ancre ». Ces mesures sont celles marquées d'un rond noir (○). Les droites pointillées représentent plusieurs combinaisons débutant à la même mesure (●) (voir texte).

Les mesures entourées d'un rond noir sont, sur base des suppositions précédentes, considérées comme autant de racines différentes. Nous appelons « ancrés » de racines ces mesures situées en-dessous d'un certain seuil.

Les droites pointillées montrent ce que pourrait être la droite déterminant la combinaison pour la mesure à 130 mm en jour 17 (ancre). En fonction du coefficient angulaire de cette droite, il y a donc une infinité de combinaisons possibles, quelques-unes étant représentées ici en pointillé. Si l'on étend ce raisonnement aux autres droites, nous aurons donc une infinité de configurations possibles. Choisir la meilleure configuration relève toujours du choix du meilleur ensemble de combinaisons. Pour déterminer lesquelles de ces dernières sont les meilleures, nous leur imposerons deux contraintes:

- (1) Les combinaisons doivent nécessairement contenir un point situé en dessous d'un certain seuil (une ancre de racine).

(2) La combinaison optimale minimise la dispersion des éléments la constituant autour d'une vitesse de croissance supposée. D'où la nécessité d'un critère (fonction objectif, indicateur) évaluant la dispersion.

Sous ces contraintes, être valide n'est plus une condition nécessaire pour être la combinaison optimale. Le problème des mesures manquantes est ainsi intégré à la recherche des correspondances.

Néanmoins, il faudra que les combinaisons émanant de différentes ancres ne contiennent pas d'éléments identiques. Cette contrainte de cohérence devra toujours être vérifiée.

Cette mise en situation nous permet d'esquisser ce que seront les principales étapes du fonctionnement de l'algorithme:

- la recherche des ancres;
- la recherche des vitesses de croissance minimisant la dispersion d'une combinaison autour d'une droite;
- l'utilisation d'un indicateur (ou fonction objectif) évaluant cette dispersion;
- la vérification de la cohérence des combinaisons entre elles et, si nécessaire, la résolution des conflits d'appartenance d'une même mesure à plusieurs combinaisons;
- l'estimation finale de la vitesse de croissance de la combinaison par régression linéaire sur le temps.

La façon de traiter le problème suit donc une procédure inverse de celles présentées dans les deux chapitres précédents en ce sens que nous recherchons d'abord une vitesse minimisant la dispersion des mesures d'une racine retracée, obtenons ensuite différentes combinaisons et déterminons enfin quels éléments appartiennent à quelles combinaisons.

### *Hypothèses*

1. Hypothèse sur les racines ancrées : Sont considérées comme ancrées toutes les mesures de la première date de mesure et les mesures des dates suivantes situées en-dessous d'un seuil correspondant à la longueur attendue de la plus petite mesure de la date précédente sur base de la moindre vitesse postulée.

2. Hypothèse sur la date d'apparition des racines : La date d'apparition d'une racine se déduit de la date à laquelle elle est mesurée la première fois et de sa longueur à cette date. L'ancre la plus longue de la première date est donc plus âgée que l'ancre la plus courte. Et les ancrées de la première date sont des racines plus âgées que celles de la seconde et ainsi de suite car l'on suppose qu'elles sont apparues entre les deux dates.

3. Hypothèse sur la vitesse de croissance des racines : La vitesse des racines est considérée comme inversement proportionnelle à leur âge. Cette hypothèse, nécessaire au processus de modélisation inverse, repose sur l'observation de plusieurs faisceaux adventifs et sur les tests et évaluations des précédents algorithmes. On postulera par exemple des vitesses de croissance de 50, 40, 30, ... mm/j pour les première, deuxième et troisième dates de mesure.

4. Hypothèse de non-régression : Dans cette situation, l'hypothèse exprimant le fait qu'une racine ne peut avoir régressé d'une date de mesure à la suivante porte sur la pente nécessairement croissante des droites de référence et de régression linéaire d'une certaine combinaison.

5. Hypothèse de cohérence sur l'indicateur : L'indicateur construit devra nécessairement prendre en compte une et une seule mesure (*a fortiori* la plus proche de la droite) par date lors de l'évaluation d'une combinaison donnée. En effet, prendre en compte plus d'une mesure peut entraîner une évaluation biaisée de la combinaison si plusieurs mesures se trouvent à proximité de la droite de référence.

6. Hypothèse de cohérence sur le nombre d'éléments d'une combinaison. Une combinaison ne doit pas nécessairement être valide mais doit contenir au moins trois éléments, sans quoi la droite de régression finale de la combinaison mènerait à une indécision (si deux éléments) ou serait incalculable (si un seul élément).

## VI.2. Algorithme paradigmatique de correspondance

### *Schéma paradigmatique*

Le tableau ci-dessous est identique au Tableau V. 1. Il reprend la matrice des mesures effectuées à cinq dates successives du faisceau adventif de la plante ID0514 et représentées dans la Figure VII. 2.

TABLEAU VI. 1. : Matrice S des données mesurées ( $M_1$  à  $M_8$ ) du faisceau adventif de la plante ID0514.

	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$	$M_7$	$M_8$
<b>08-Mar-07 (A)</b>	105	130	0	0	0	0	0	0
<b>12-Mar-07 (B)</b>	110	115	295	315	470	510	NaN	NaN
<b>16-Mar-07 (C)</b>	65	260	285	490	525	530	0	0
<b>20-Mar-07 (D)</b>	95	135	470	505	600	615	770	790
<b>23-Mar-07 (E)</b>	75	240	585	620	630	655	885	0

Sur base des hypothèses 1. et 3., on détermine le seuil des dates B et C, respectivement  $s_B$  et  $s_C$ , suivant la formule proposée en hypothèse 3.:

$$\begin{aligned} s_B &= \min(\Psi_A) + v_A \cdot \Delta t = 105 + (50 - 4,5) \cdot 4 = 287 \text{ mm} \\ s_C &= 110 + (40 - 4,5) \cdot 4 = 252 \text{ mm} \end{aligned} \quad (6.1.)$$

Dans l'équation (6. 1.),  $\Psi_A$  et  $v_A$  désignent respectivement les mesures ancrées et la vitesse de croissance en date A. La vitesse moyenne supposée est minorée de 4,5 mm/j, pour éviter de rencontrer des ancrées « parasites », c'est-à-dire des ancrées que l'on considère comme première mesure d'une racine alors qu'il s'agit de la seconde (voir VI. 3. § *Approximation globale* et VI. 3. § *Ancres parasites*).

Après les seuils, nous pouvons déterminer les ancrées, correspondant chacune au premier point de mesure de racines différentes. Ces ancrées sont indiquées en vert dans le tableau et sont au nombre de six. Les mesures en bleu dans ce tableau pourraient également prétendre à ce statut. Cependant, sous respect de l'hypothèse 6., elles ne nous sont pas nécessaires.

Il nous incombe maintenant d'évaluer différentes vitesses de croissance pour chacune des six racines. Comme supposé dans l'hypothèse 3., chaque racine reçoit une vitesse de départ  $v_A$  imposée en fonction de son âge, c'est-à-dire en fonction de la date à laquelle elle est considérée comme ancre (hypothèse 2.). En faisant varier le coefficient angulaire (c-à-d la vitesse de départ) de la droite passant par l'ancre d'une des six racines, on peut évaluer une combinaison (c-à-d une droite d'une certaine vitesse passant par l'ancre) sur sa proximité avec les mesures des dates suivantes.

Si le Tableau VI. 1. est la matrice  $S$ , l'évaluation de la combinaison  $\mathbf{c}$ , sur  $x$  dates, créée à partir d'une ancre et d'une vitesse repose sur une fonction objectif définie telle que :

$$\Omega_c = f(\Psi_{A_1}, v_A, S) \quad (6.2)$$

Si  $\mathbf{z}$  est le vecteur des valeurs attendues sur base de la droite définie par l'ancre  $\Psi_{A_1}$  et la vitesse  $v_A$ , c'est-à-dire les valeurs attendues de la racine retracée à chaque temps  $\mathbf{t}=(t_A, \dots, t_X)$ :

$$\mathbf{z} = v_A \cdot \mathbf{t} + (\Psi_{A_1} - v_A \cdot t_A) \quad (6.3)$$

Alors,

$$\Omega_c = P(c_1=z_1) \cdot P(c_2=z_2) \cdot \dots \cdot P(c_x=z_x) \quad (6.4)$$

On suppose qu'à la date de mesure  $t_i$ , les éléments  $w$  de la  $i$ ème ligne de  $S$  (reprennant toutes les mesures susceptibles de constituer le  $i$ ème élément de la combinaison  $\mathbf{c}$ ) situés dans le voisinage de  $z_i$  sont normalement distribués autour de cette valeur suivant la distribution  $W_i$ , de moyenne  $z_i$  et de variance  $\sigma_i^2$  :

$$W_i \sim N(z_i, \sigma_i^2) \quad (6.5)$$

Le vecteur  $\mathbf{c}$  évalué à l'équation (6. 4.) est constitué des éléments  $c_i$  égaux à  $w_i$  tels que la

probabilité  $P(w_i = z_i)$  soit maximale, ce qui signifie que  $c_i$  est l'élément  $w_i$  le plus proche de  $z_i$ , sa valeur attendue.

Cette évaluation de combinaisons s'opère sur les  $n$  combinaisons composant la configuration  $\Theta$  ; on peut donc définir sur base de l'équation (6. 2) une fonction multi-objectif de la configuration:

$$\Omega_{\Theta} = f((\Psi_1, \dots, \Psi_n), (v_A, \dots, v_{X-2}), S) \quad (6. 6)$$

dont une expression plus claire est:

$$\Omega_{\Theta} = f(\Omega_{c_1}, \dots, \Omega_{c_n}) = \Omega_{c_1} \cdot \dots \cdot \Omega_{c_n} \quad (6. 7)$$

Les modalités de l'implémentation de cette fonction feront l'objet d'une section suivante (VI. 2. § *Indicateur implémenté*).

Dans la situation qui nous occupe, la configuration jugée optimale par cet indicateur est la configuration  $\Theta$  égale à :

$$\Theta = \begin{pmatrix} \mathbf{105} & 315 & 525 & 770 & 885 \\ \mathbf{130} & 315 & 530 & 770 & 885 \\ 105 & \mathbf{110} & 285 & 470 & 620 \\ 105 & \mathbf{115} & 285 & 470 & 585 \\ 105 & 110 & \mathbf{65} & 135 & 240 \end{pmatrix} \quad (6.8)$$

Dans chaque ligne de cette matrice (6. 8), la mesure reprise en gras est l'ancre de la racine potentiellement retracée par la combinaison constituant cette ligne.

Cette configuration est peut-être optimale mais n'est pas valide dans le sens où, à une même date, on retrouve plusieurs mesures identifiées comme étant identiques. Ce problème est d'ailleurs loin d'être isolé.

Chaque ligne devant conserver l'ancre qu'elle possède, on peut déjà effectuer un premier tri en invalidant les combinaisons possédant une mesure identifiée comme ancre d'une autre combinaison. De la matrice (6. 8) résulte la matrice:

$$\Theta = \begin{pmatrix} \mathbf{105} & 315 & 525 & 770 & 885 \\ \mathbf{130} & 315 & 530 & 770 & 885 \\ NaN & \mathbf{110} & 285 & 470 & 620 \\ NaN & \mathbf{115} & 285 & 470 & 585 \\ NaN & NaN & \mathbf{65} & 135 & 240 \end{pmatrix} \quad (6. 9)$$

Un second tri s'avère néanmoins nécessaire. Ce tri se base sur une matrice identique à  $S$  et reprenant, à chaque position  $(i,j)$  dans cette dernière, l'occurrence de la mesure  $S_{ij}$  dans la ligne  $S_i$  par rapport à son occurrence dans la colonne  $\Theta_i$ . Cette matrice est nommée  $O_{\Theta,S}$  et, dans notre cas, égale à:

$$O_{\theta,s} = \begin{pmatrix} 1 & 1 & NaN & NaN & NaN & NaN & NaN & NaN \\ 1 & 1 & 0 & 2 & 0 & 0 & NaN & NaN \\ 1 & 0 & 2 & 0 & 1 & 1 & NaN & NaN \\ 0 & 1 & 3 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 & 2 & NaN \end{pmatrix} \quad (6.10)$$

Chaque élément de la matrice (6. 10) supérieur à 1 doit donc faire l'objet d'une attention particulière. La façon la plus simple de procéder est de distribuer ces éléments à la combinaison dont ils sont le plus proche. Suivant cette procédure, la matrice (6. 9) devient:

$$\Theta = \begin{pmatrix} \mathbf{105} & 315 & 525 & NaN & 885 \\ \mathbf{130} & NaN & 530 & 770 & NaN \\ NaN & \mathbf{110} & NaN & 470 & 620 \\ NaN & \mathbf{115} & 285 & NaN & 585 \\ NaN & NaN & \mathbf{65} & 135 & 240 \end{pmatrix} \quad (6.11)$$

D'autres façons de procéder seront développées dans les avatars de cet algorithme. Ces versions permettront notamment d'intégrer le problème posé par les ancres qui se révèlent être parasites. Les combinaisons sont toutes cohérentes sous respect de l'hypothèse 6.

La configuration finale est une configuration de correspondance valide, composée de combinaisons invalides, au sens de l'algorithme du Chapitre V. Ce terme « invalide » n'indique aucune péjoration sur ces combinaisons. Au contraire, elles sont plus pertinentes que celles déduites de l'algorithme du Chapitre V. car ces dernières n'incluaient pas la possibilité qu'une mesure soit parasite ou que la racine n'ait pas été mesurée.

Nous pouvons donc déterminer finalement les vitesses de croissance de 5 racines par régression linéaire des combinaisons de la matrice (6. 11) sur les dates dont sont issus les éléments les composant. Ces vitesses de croissances sont:

$$v = \begin{pmatrix} 51,99 \\ 52,86 \\ 46,08 \\ 42,74 \\ 24,59 \end{pmatrix} \quad (6.12)$$

La configuration de correspondance finale est illustrée sur la Figure VI. 3, à la page suivante.

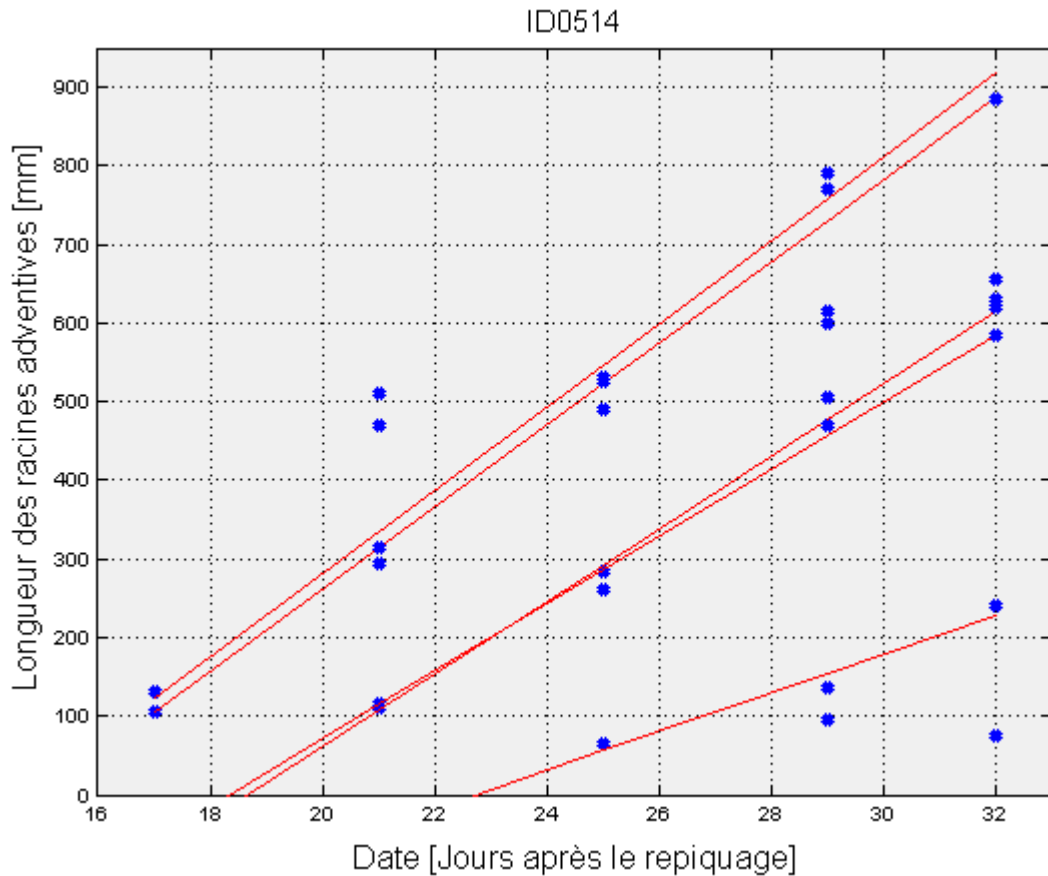


FIGURE VI. 3.: Droites de régression linéaire (en rouge) des combinaisons constituant la configuration de correspondance finale du faisceau adventif de la plante ID0514.

#### Indicateur implémenté

La fonction objectif implémentée en MATLAB® et utilisée dans l'implémentation de l'algorithme décrit ci-dessus se nomme *indicateur2.m*.

Le développement de cette fonction, conceptualisée en équation (6. 6), porte sur (1) l'implémentation de la distribution décrite en équation (6. 5) et (2) sur la façon dont on intégrera les objectifs multiples résultant chacun de l'équation (6. 4).

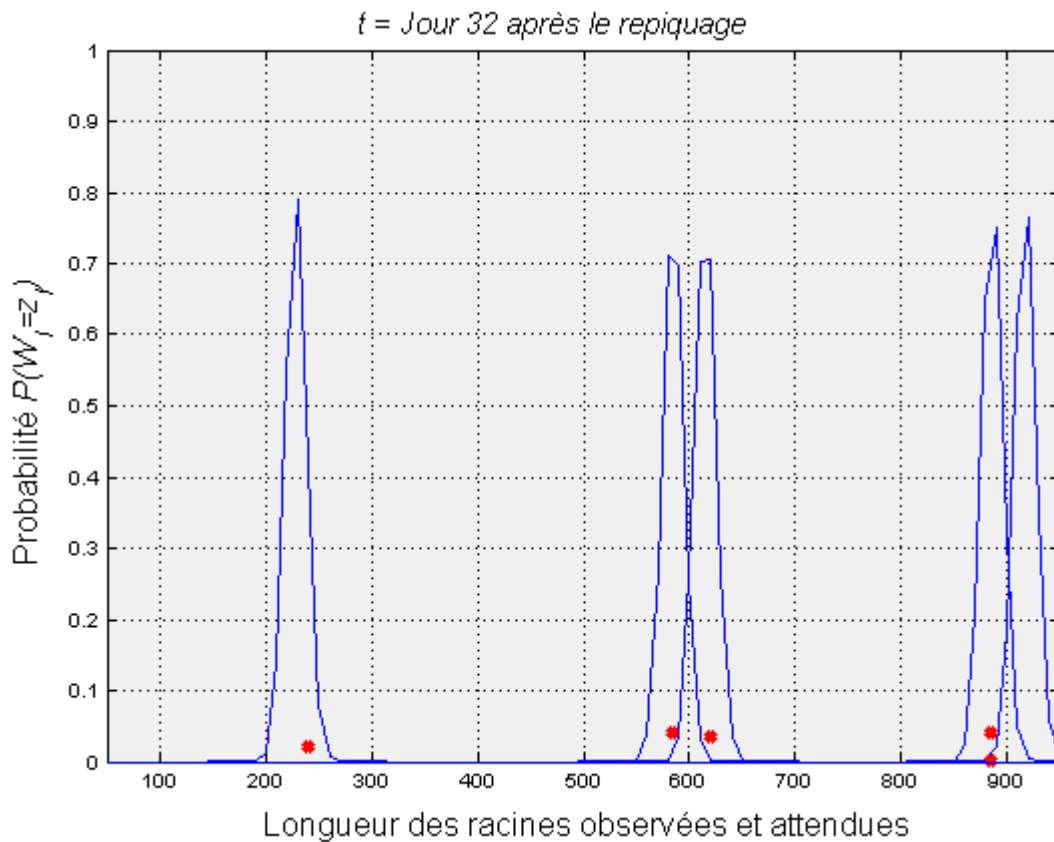
#### (1) Implémentation de la distribution (6. 5)

Pour rappel, nous avons supposé que les éléments  $w_i$  au voisinage d'une longueur attendue  $z_i$  d'une racine mesurée à un temps  $t_i$  suivaient une distribution:

$$W_i \sim N(z_i, \sigma_i^2) \quad (6. 5)$$

La question est de savoir quelle valeur attribuer à la variance de cette distribution. Nous choisissons un écart-type de 10 mm, donc une variance de 100 mm<sup>2</sup>. Le seul fondement de ce choix est la graduation de la règle utilisée qui est de 5 mm. Ce qui signifie qu'une mesure est au maximum précise à 2,5 mm. Si une imprécision est faite à une certaine date, alors le manque de précision attendu après 4 jours de croissance (l'intervalle commun à la plupart des dates de mesures) est donc

de 10 mm. Par ailleurs, afin d'avoir une probabilité maximale proche de 1, nous multiplions cette distribution par 20. La Figure VI. 4. illustre les distributions  $20.W_i$  pour chacune des droites retracée en jour 32.



**FIGURE VI. 4.:** Distribution (en bleu) de l'indicateur pour la combinaison 5 ( $t=32$  jours après le repiquage). En rouge, les mesures réalisées à cette date.

Dans l'implémentation de cette fonction objectif, les mesures ancrées ont forcément une probabilité maximale. Cependant, nous les excluons de notre évaluation. En effet, on sait a priori que les ancrées peuvent entraîner des redondances dans plusieurs combinaisons autres que celle à laquelle on veut, par définition, qu'elles appartiennent (voir matrices (6. 8) et (6. 9)). Afin de ne pas biaiser l'évaluation, nous préférons donc ne pas les inclure dans l'évaluation.

## (2) Agrégation des objectifs multiples en un seul indicateur

Chacune des fonctions objectif calculées par l'équation (6. 2) doit être agrégée en une formule structurée comme en équation (6. 6)

Pour une combinaison donnée, la fonction objectif est un produit de probabilités (valeurs comprises entre 0 et 1). Cet agencement a un inconvénient majeur en cas de comparaison entre faisceaux. En effet, si  $S_1$  est un faisceau mesuré sur 5 dates et  $S_2$ , un faisceau mesuré sur 4 dates, et si nous considérons deux combinaisons à tester provenant de ces deux faisceaux, la valeur objectif de la



droite émanant du premier risque fortement d'être plus faible que celle émanant du second, du fait qu'il y a 5 éléments dans le premier et 4 dans le second. Afin d'inclure cet effet de grandeur, nous préférons la somme des probabilités plutôt que le produit. Pour une combinaison  $\mathbf{c}$ , l'équation (6. 4) devient:

$$\Omega_c = P(c_1=z_1) + P(c_2=z_2) + \dots + P(c_x=z_x) \quad (6. 13)$$

La valeur de la fonction objectif d'une combinaison peut être dès lors supérieure à 1 et sera d'autant plus grande qu'il y a de points dans la combinaison (ce qui est un indicateur de sa validité).

L'agrégation des valeurs objectif de plusieurs combinaisons se fera quant à elle en multipliant chacune des valeurs objectif. Ces valeurs, propres à une configuration donnée, se dispersant sur plusieurs ordres de grandeurs, nous choisissons comme indicateur  $I$  (i.e. Résultat de la fonction objectif) l'opposé du logarithme naturel, ce qui équivaut, si  $\mathbf{c}$ ,  $\mathbf{d}$ , et  $\mathbf{e}$  sont les 3 combinaisons d'une configuration  $T$ , à:

$$I_T = -\ln(\Omega_c \cdot \Omega_d \cdot \Omega_e) \quad (6. 14)$$

La fonction *indicateur2* admet 4 inputs:

- **vits**: ce vecteur ( $n \times 1$ ) reprend les vitesses des  $n$  racines à retracer;
- **t**: ce vecteur ( $x \times 1$ ) est le vecteur temps des  $x$  dates de mesure;
- **t2**: ce vecteur ( $n \times 1$ ) reprend les dates auxquelles s'ancrent les  $n$  racines;
- **ancres**: ce vecteur ( $n \times 1$ ) reprend les ancres mesurées aux temps **t2** correspondant;
- **sets**: cette matrice ( $x \times m$ ) est la matrice  $S$  telle que présentée en Tableau VI. 1.

Elle renvoie bien entendu un seul output : *indic*, une valeur scalaire calculée suivant l'équation (6. 14) et dont la valeur minimale indique la configuration de vitesses optimales

#### *Pseudo-code de l'algorithme de base*

La fonction MATLAB® implémentée s'appelle *correst03.m*. Elle admet 4 paramètres d'entrées, détaillés, pour  $x$  dates de mesure, comme suit:

- **num**: valeur scalaire, numéro de la plante dans l'expérience (de 1 à  $N$ );
- **table**: tableau ( $N \times x$ ) de vecteurs de longueur variable reprenant l'ensemble des mesures effectuées sur le faisceau adventif à chacune des cinq dates de mesure;
- **t**: vecteur ( $x_l \times 1$ ) reprenant les  $x_l$  différentes dates ( $x_l \leq x$ ) auxquelles au moins une mesure a été effectuée sur le faisceau adventif;
- **reftable**: vecteur ( $x \times 1$ ) reprenant les dates correspondantes à chacune des colonne du tableau *table*, c'est-à-dire chacune des dates de mesure;

Pour  $n_l$  racines retracées, elle renvoie 3 outputs:  $p$  (matrice ( $n_l \times 2$ )),  $\mathbf{t0}$  (vecteur ( $n_l \times 1$ )) et  $fval$  (valeur scalaire). Ces trois outputs seront détaillés à la fin de cette section.

Etape 1: Chargement de la matrice  $S$  reprenant l'ensemble des mesures pour la plante identifiée par num.

Cette matrice est structurée comme le Tableau VI. 1.; elle contient donc des NaN ou 0 aux emplacements vides.

Etape 2: Détermination des seuils.

Cette étape se base sur les mesures de la première date et les vitesses de départ [50,40,30]. On considère, s'il n'y a pas de mesure à la date concernée que le seuil est de 300 mm.

Etape 3: Détermination des mesures ancrées.

Les mesures situées en dessous des seuils définis à l'étape 2 sont considérées comme ancrées, c'est-à-dire comme représentatives d'autant de mesures différentes. Elles sont recensées dans un vecteur ( $n \times 1$ ). Un vecteur de même taille reprend les temps auquel ces ancrées sont initiées (date de leur mesure) et un autre les vitesses de départ qu'il faut leur adjoindre en fonction de ces temps. Ces vecteurs sont nécessaires à la fonction *indicateur2* (voir section précédente).

Etape 4: Optimisation des vitesses par minimisation de la fonction objectif *indicateur2*.

La recherche de la valeur minimale de la fonction *indicateur2* est effectuée par une fonction MATLAB®, fournie par le constructeur (©The MathWorks), *fminsearch*. Cette fonction se base sur l'algorithme du simplex Nelder-Mead pour rechercher un minimum local d'une fonction multivariée. Dans notre cas, le nombre de variables est  $n$ , nombre d'ancres déterminées à l'étape précédente. En outre, *indicateur2* admet également des paramètres identiques pour toutes les valeurs de ces  $n$  variables; ce sont les paramètres **ancres**,  $\mathbf{t}$ ,  $\mathbf{t2}$  et *sets*.

*fminsearch* renvoie la valeur optimale des  $n$  variables, le minimum  $fval$  de la fonction objectif et une valeur binaire *exitflag* concernant l'approche de la solution : *exitflag* = 1 si la fonction converge vers une solution et *exitflag* = 0 si le nombre maximal d'itérations a été atteint. Nous choisissons donc, si ce dernier cas échoue, de relancer la recherche avec les dernières valeurs atteintes pour les variables. Si la recherche devait à nouveau mener à *exitflag*=0, alors nous abandonnerions la poursuite et garderions les valeurs actuelles.

#### Etape 5: Détermination des combinaisons.

Sur base des vitesses et des ancrés, on détermine premièrement les droites.

Deuxièmement, les combinaisons sont élaborées telles qu'en équation (6. 8).

Troisièmement, on procède à l'exclusion des valeurs redondantes à la même date dans plusieurs combinaisons. Cette exclusion concerne en premier lieu les ancrés (voir matrice (6. 9)) et en second lieu les autres éléments redondants.

Enfin, les combinaisons contenant moins de trois éléments sont totalement exclues.

Il nous reste donc au final  $n_l$  combinaisons, correspondantes à autant de racines ( $n_l \leq n$ ).

#### Etape 6: Evaluation des vitesses et des temps d'initiation.

Par régression linéaire des combinaisons sur le vecteur des dates desquelles sont issus les éléments les composant, on prédit la vitesse de la racine retracée par la combinaison. En outre, il est possible de prédire également, sur base de cette régression, la date à laquelle la racine a été initiée (c'est-à-dire la date à laquelle la longueur était nulle). Si  $\tilde{y} = \beta_1 \cdot t + \beta_2$  est cette droite de régression, alors le temps d'initiation  $t_0$  est égal à  $-\beta_2/\beta_1$ .

Les outputs renvoyés par *correst03* sont:

- $p$  : matrice ( $n_l \times 2$ ) comprenant, à chaque ligne, le coefficient angulaire (vitesse de croissance) et l'intercept de la droite prédite sur base de la combinaison correspondante;
- $t0$  : vecteur ( $n_l \times 1$ ) comprenant les temps d'initiation des  $n_l$  racines retracées;
- $fval$  : valeur optimale de la fonction objectif.

### VI.3. Cas particuliers et avatars.

Dans ce sous-chapitre, nous passerons en revue les différents cas particuliers, relevés lors des tests de la fonction *correst03*, qui ont mené à des versions continuellement plus enrichies de cette fonction. Il convient préalablement de préciser que la version de la fonction *correst03* disponible sur le CD attaché à ce mémoire est la version finale, donc comprenant tous ces avatars.

#### *Approximation globale de l'optimum.*

La recherche du minimum de la fonction objectif *indicateur2* effectuée par la fonction *fminsearch* préfère parfois un minimum local à un autre plus bas mais situé hors du voisinage des valeurs de départ. Cet effet est d'autant plus remarqué que le nombre de variables est grand. Pour y

faire face, nous avons choisi d'implémenter une pré-recherche de minimum de la fonction sur un voisinage des valeurs de départ plus large. Cette approximation procède par pas de 0,5 mm/jour sur un intervalle de  $\pm 4,5$  mm/jour autour des vitesses de départ.

La procédure d'approximation est itérative. A partir du vecteur de vitesses de départ de chaque racines, on détermine quelle est la vitesse optimale de l'intervalle sus-mentionné pour la première racine. On garde cette valeur et on procède de la même façon pour la seconde racine et ainsi de suite.

Au final, de cette approximation ressort un vecteur de vitesse de départ différent, indiquant une certaine tendance du minimum à rechercher (pour autant que la fonction *indicateur2* converge en un tel minimum) e donc prêt à l'optimisation par *fminsearch*.

#### *Cas des ancrés parasites.*

Ce cas particulier est, de loin, le plus important. Il s'insère dans l'étape de détermination des combinaisons (étape 5). Soit la Figure VI. 5, reprenant le faisceau adventif de la plante ID0514, déjà mis en situation précédemment.

Sur cette figure (voir page suivante) sont seulement reprises deux combinaisons finales (droites pleines rouge et verte) ancrées en deux mesures (entourées toutes deux d'un rond bleu). Chacune de ces deux combinaisons sont composés de deux éléments. Sous respect de l'hypothèse 6., elles sont donc considérées comme incohérentes et doivent être éliminées. Cette perte se fait aux dépens de deux autres combinaisons possibles cependant.

En effet, si l'on suppose que l'ancre de la combinaison rouge, en troisième date, soit une mesure en réalité parasite (c-à-d ne correspond en aucun cas à la mesure d'une racine de ce faisceau à ce moment de l'expérience), on retrouve une combinaison contenant les deux éléments de la droite verte et le second élément de la droite rouge. Cette combinaison, illustrée par une droite pointillée sur le schéma, est donc tout à fait cohérente. Les éléments susceptibles de la composer sont indiqués par une flèche verte.

Inversément, l'ancre de la droite verte pourrait également être parasite et, de ce fait, masquer la combinaison illustrée par une droite pointillée rouge, dont les éléments sont indiqués par des flèches rouges. Ce second cas est moins probable cependant car l'on s'attend à rencontrer des parasites dans les mesures les plus longues.

Ces deux combinaisons n'auraient pas pu être envisagées par l'algorithme car elles se basent sur l'invraisemblance de l'hypothèse 1. de l'algorithme. Néanmoins, il est intéressant d'effectuer une vérification ultérieure des combinaisons postulées sur cette base et le cas échéant, de les dissoudre et d'envisager quelles autres combinaisons peuvent être élaborées.

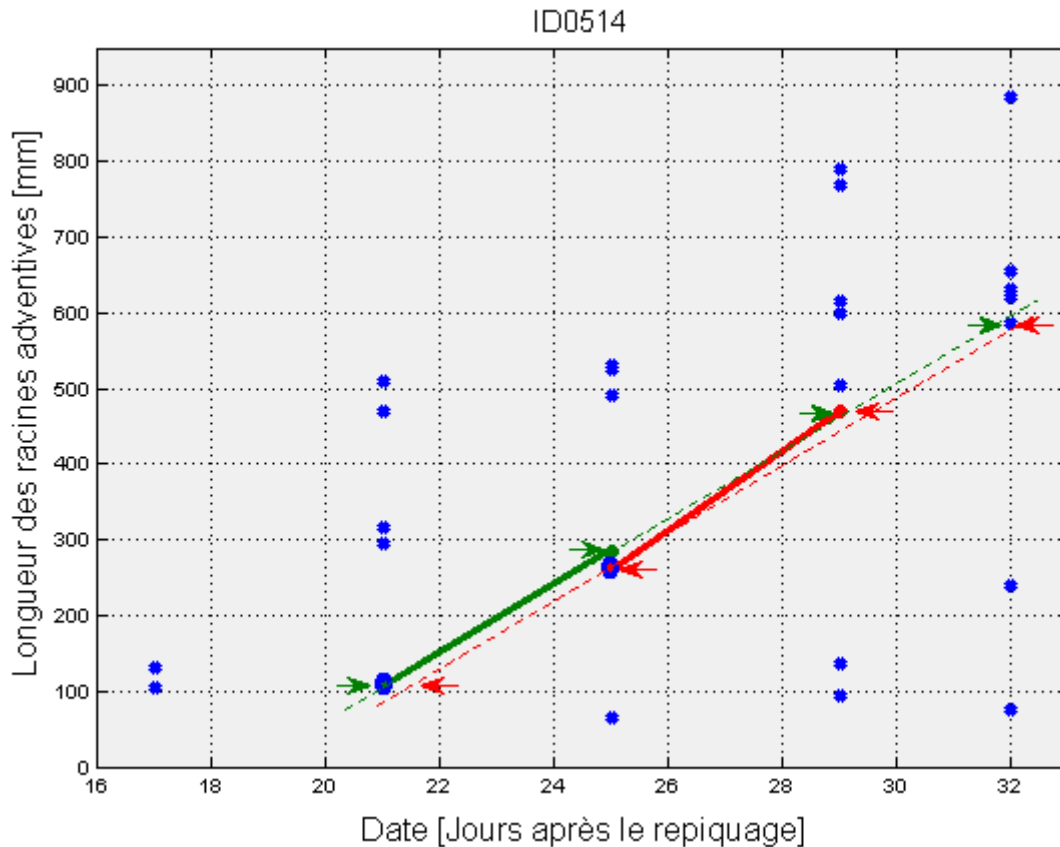


FIGURE VI. 5.: Illustration du cas particulier des ancrés parasites sur le faisceau adventif de la plante ID0514.

Le traitement de ce cas particulier nécessite une modification de l'algorithme de base, entraînant de fait la création d'une nouvelle version. Nous détaillons ici ce nouveau schéma et l'implémentation de ce dernier.

Une fonction dérivée d'*indicateur2* et nommée *indicateur3.m*, nous renvoie, en plus de la valeur optimale de la fonction objectif, les combinaisons résultant des droites créées à partir des ancrés et des vitesses optimales. Cet output est la matrice (6. 8). Nous allons la parcourir de ligne en ligne (donc, envisager chaque combinaison).

A chaque itération de ce parcours, nous construisons un duplicata de la matrice (6. 8), composé de 0 et de 1 pour les éléments respectivement différents et identiques à ceux de la ligne considérée à cette itération. Sur base de cette matrice dupliquée, nous recherchons quelle ligne présente le maximum de similarité avec celle que nous considérons. Si cette similarité porte sur trois éléments différents (comme nous avons 5 dates de mesure, deux lignes ayant deux éléments similaires peuvent être cohérentes) ou plus, alors on teste chacune des combinaisons sur son  $R^2$  et on choisit d'éliminer (ou pas, voir plus bas) les éléments redondants de la ligne ayant le coefficient de détermination le plus faible.

Ce traitement ne s'applique pas, bien entendu, aux ancrés de droites le cas échéant.

La procédure utilisée est donc différente de celle, détaillée dans le schéma de l'algorithme de base, utilisant une matrice d'occurrence (6. 10).

Elle nous permet de régler le problème posé par les ancrages parasites en évaluant au voisinage d'une mesure les éléments avec lesquels elle serait interchangeable car proches l'un de l'autre (par exemple, dans la Figure VI. 5., les éléments rouge et vert de la troisième date sont interchangeables).

Ce traitement est effectué sur un élément à éliminer par la fonction *interchange.m*, qu'il est inutile de décrire; il importe seulement de savoir que le seuil d'interchangeabilité est une distance de 100 mm entre deux éléments.

#### *Cas des fausses redondances*

Ce cas particulier provient de la modification de l'algorithme paradigmatique engendrée par le cas particulier précédent. En effet, de l'abandon du traitement des redondances par la matrice d'occurrence (6.10), il découle que les éléments redondants présents dans plusieurs combinaisons ont été examinés et exclus de certaines correspondances lors d'une nouvelle étape (décrite au cas précédent) qui ne tient pas compte du fait que certaines mesures peuvent avoir une occurrence supérieure à 1 dans la matrice *S* de départ. Ces mesures, considérées injustement comme redondantes, constituent ce que nous appelons les fausses redondances, qu'il convient de réhabiliter.

La réhabilitation des fausses redondances s'effectue par vérification de l'occurrence de chaque élément (fonction *occurrence.m*). Si un tel cas survient, on détermine l'occurrence de l'élément dans *S* (fonction *occurrence2.m*) et dans les combinaisons. A partir de cela, la fonction *interchange2* fait la part des choses entre vraies et fausses redondances et garde les meilleures combinaisons en les évaluant et en proposant des éléments interchangeables d'une façon analogue à *interchange.m*, dans le cas précédent.

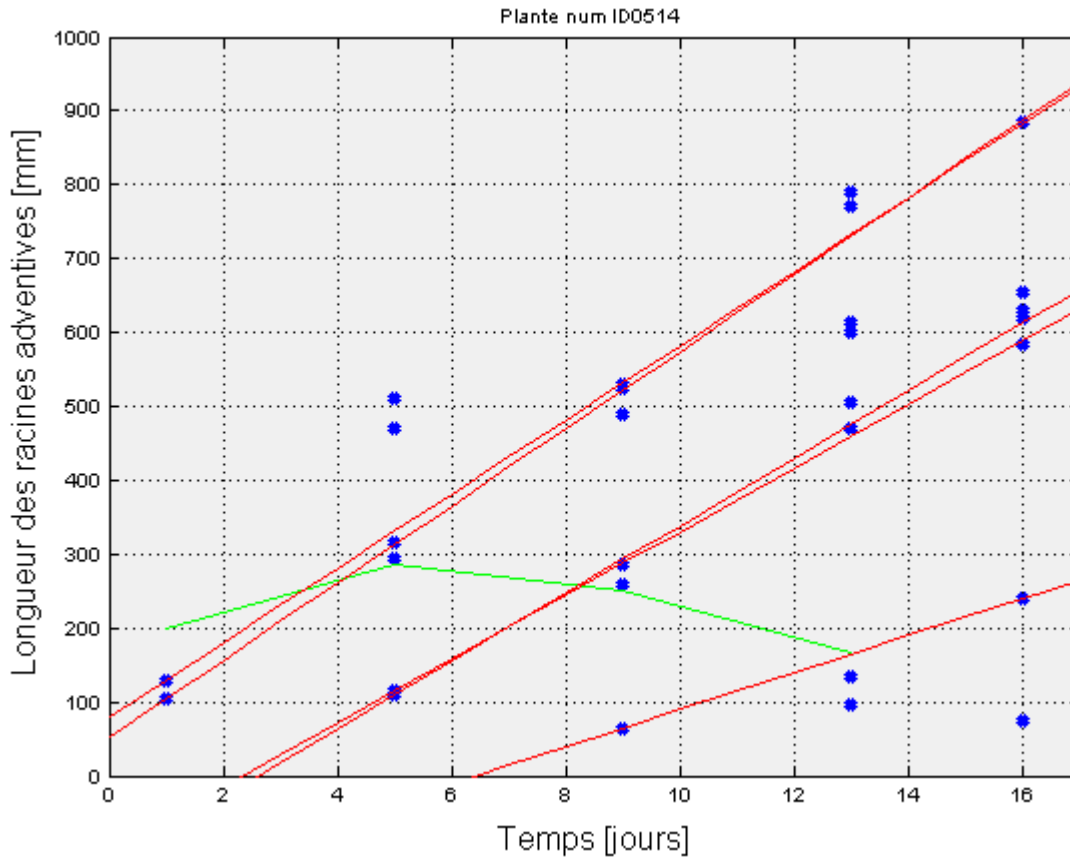
Au final, la prédiction des vitesses de croissance porte sur les combinaisons contenant au moins trois éléments.

Cette dernière version est donc la plus complète et est implémentée dans la fonction *correst03*.

#### *Option graphique.*

Nous avons également développé une version *correst0\_graphes3.m* de la dernière version de l'algorithme de base. Cette fonction comporte un input supplémentaire, *ymax*, valeur scalaire de longueur de l'axe des ordonnées (axe des longueurs racinaires). En outre, elle fait appel à la fonction *grapheplante.m* qui crée le graphe des vitesses d'une plante donnée.

La version graphique renvoie deux outputs graphiques. Le premier illustre les droites optimales ancrées et le second, les droites de régression des combinaisons correspondantes, c'est-à-dire la meilleure configuration. Dans les Figures VI. 6 et VI. 7 illustrant ces deux outputs, l'axe du temps est référencié à la date précédant la mesure la plus ancienne. Dans notre cas, le jour 1 sur les graphiques correspond donc au jour 17.



**FIGURE VI. 6:** Droites optimales ancrées (en rouge), seuil d'ancrage (en vert) et mesures de longueur (en bleu) du faisceau adventif de la plante ID0514.

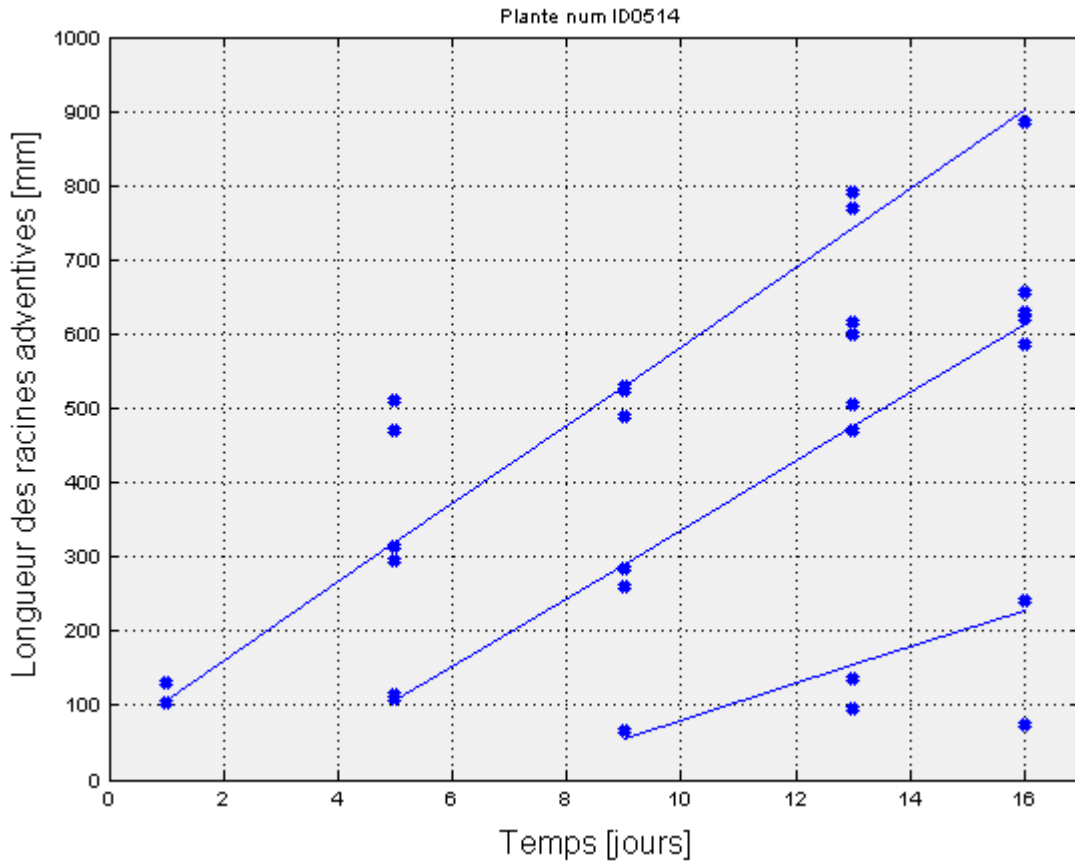


FIGURE VI. 7.: Droites de régression des combinaisons optimales (en bleu) et mesures des longueurs du faisceau adventif de la plante ID0514.

Dans la figure ci-dessus, la configuration optimale de correspondance est:

$$\Theta = \begin{pmatrix} 105 & 315 & 525 & 770 & 885 \\ NaN & 110 & 285 & 470 & 620 \\ NaN & NaN & 65 & 135 & 240 \end{pmatrix} \quad (6.15)$$

Les outputs renvoyés sont ceux de la version complète de l'algorithme, *correst03*:

$$p = \begin{pmatrix} 54,14 & 52,35 \\ 46,27 & -126,18 \\ 24,59 & -164,86 \end{pmatrix} \quad (6.16),$$

$$t_0 = \begin{pmatrix} -0,98 \\ 2,73 \\ 6,7 \end{pmatrix} \quad (6.17) \quad \text{et} \quad fval = -2,54 \quad (6.18)$$

Les temps d'initiation doivent bien entendu être ramenés au temps de référence (16 jours après le repiquage, dans notre cas).



## VI. 4. Conclusions et perspectives d'amélioration de l'algorithme

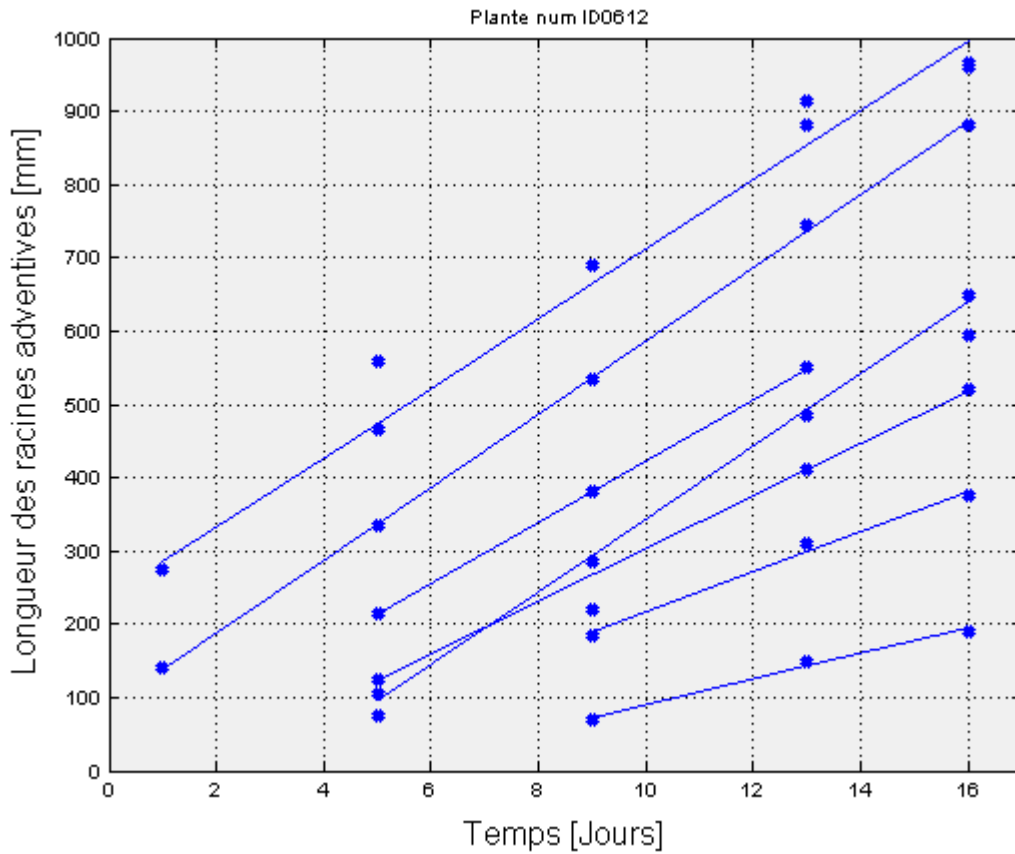
### *Conclusions*

Les vitesses de croissance et les temps d'initiation ont été prédits sur l'ensemble des données concernant les faisceaux adventifs (pour peu qu'ils respectent l'hypothèse 6. de cohérence) en utilisant la dernière version implémentée de l'algorithme. Comme on le voit sur la figure VI. 7., l'ensemble des racines supposées (c-à-d l'ensemble des ancrés) ne conduit pas forcément à l'évaluation d'une vitesse de croissance. Cela se déduit également de la façon dont l'algorithme traite les données. En effet, l'hypothèse de cohérence rejette une partie des combinaisons. D'autre part, des mesures manquantes sous les seuils d'ancrage amènent invariablement à considérer moins de racines qu'il n'est possible d'en déceler si l'on retrace des racines à partir d'autres mesures cohérentes.

Ces différentes raisons font que l'algorithme final répond au problème plus en terme de prédiction des vitesses pertinentes qu'en terme d'exhaustivité des racines retracées.

Cette conclusion peut s'illustrer par les cas du faisceau adventif des plantes ID0612 et ID0180.

La configuration de correspondance optimale du faisceau adventif de la plante ID0612 est particulièrement exhaustive. Comme on le voit sur la Figure VI. 8. (page suivante), une racine émerge de presque toute les ancrés recensées (seulement l'ancre la moins longue en deuxième date n'aboutit pas à une racine) et le retraçage de ces racines paraît fort convaincant. L'ancre non reprise reste malgré tout un point discutable; on remarque qu'il est possible de l'intégrer à la deuxième droite ancrée en jour 9.



**FIGURE VI. 8:** Droites de régression des combinaisons optimales et mesure des longueurs du faisceau adventif de la plante ID0612.

La Figure VI. 9 illustre la configuration de correspondance optimale du faisceau adventif de la plante ID0180. Sur ce graphique, le manque d'exhaustivité est flagrant: on remarque, en effet, que les trois racines provenant des ancrs 180 (jour 5), 140 et 210 (jour 9) ne sont pas retracées alors qu'il semble que des combinaisons cohérentes peuvent être formées.

Ce problème est à imputer principalement à l'optimisation des vitesses de croissance.

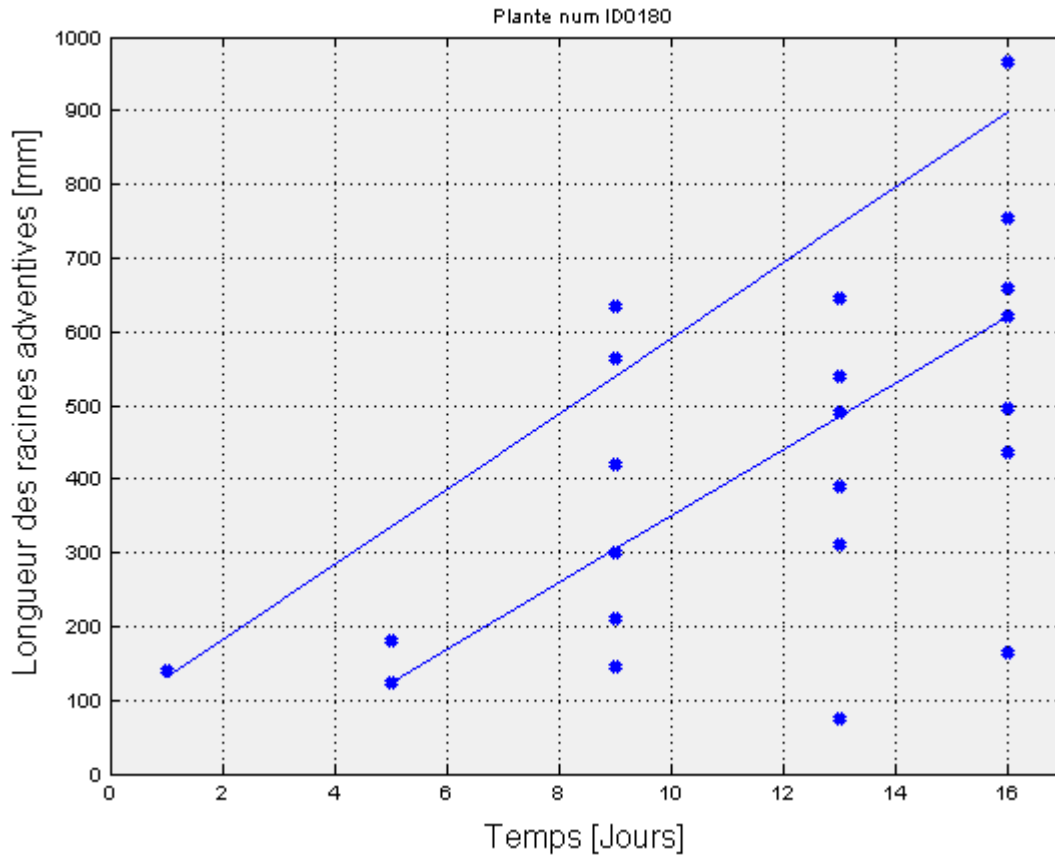
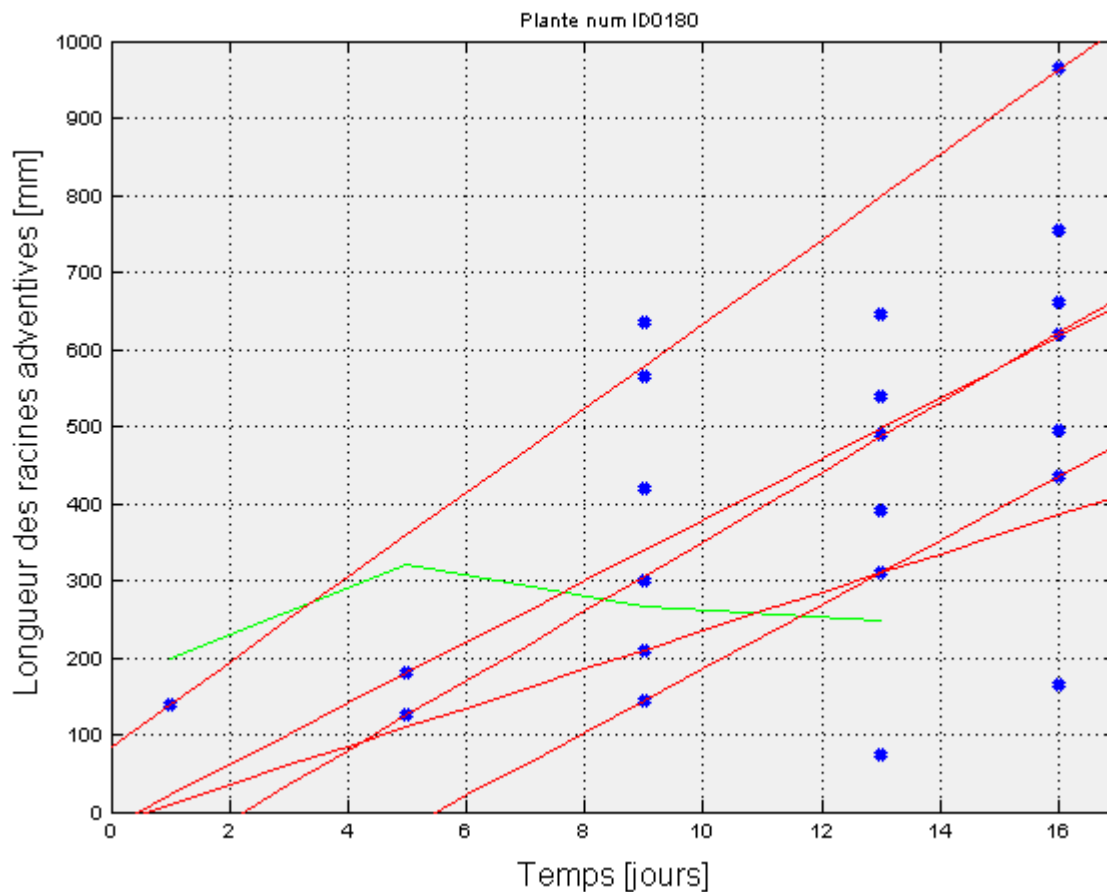


FIGURE VI. 9: Droites de régression et mesures des longueurs du faisceau adventif de la plante ID0180.

En effet, comme on le voit sur la Figure VI. 10 (page suivante), la droite ancrée en 180 (jour 5) est captée par trois des points de la droite ancrée en 130 au même jour. La construction des combinaisons se charge donc de l'éliminer.

Sur la même figure, on voit que les droites ancrées en jour 9 se partagent deux points identiques. Celle ancrée en 210 considère en outre l'ancre 130 du jour 5 comme étant parasite. Ces raisons font que la droite ancrée en 140, qui est plus probable, est éliminée du fait qu'elle ne remplit pas la condition de cohérence (au moins trois éléments).



**FIGURE VI. 10.:** Droites optimales ancrées (en rouge), seuil d'ancrage (en vert) et mesures de longueur (en bleu) du faisceau adventif de la plante ID0180.

Il importe également de remarquer que l'algorithme utilisé et évalué sur un grand ensemble de faisceaux adventifs n'a pas fait l'objet d'une validation. Une telle validation requiert, en effet, une répétition de l'expérience avec des racines identifiées afin de déterminer la pertinence des combinaisons retracées.

L'évaluation finale porte sur les 581 faisceaux adventifs (émanant des 680 plantes) qui ont été mesuré à trois dates ou plus. Notons que l'exhaustivité peut s'avérer réduite encore dans le cas où il n'y a eu que trois dates de mesure, dont une seule contenant des ancrés.

Les vitesses de croissance d'un certain nombre de racines de ces faisceaux sont donc prédites. Elles sont disponibles sur le CD attaché à ce mémoire dans le tableur *vitADV2.xls*. En outre, nous estimons, pour chaque faisceau, la moyenne des vitesses de croissance mesurées (reprises en tableur *vitADVm.xls*). Ces dernières sont des paramètres pertinents (car révélateurs de la vitesse de croissance moyenne du faisceau adventif) pour l'analyse génétique.

Elles ont ainsi fait l'objet d'une telle analyse génétique, réalisée par Jean-François Dumasy, qui a mis en évidence une région du génome vraisemblablement impliquée dans le déterminisme de

ce paramètre (« *Analyse génétique de la croissance racinaire de l'orge* », Dumasy, 2007).

### *Perspectives d'amélioration*

Le dernier algorithme présenté peut encore être amélioré. Le point majeur à améliorer est sans conteste l'exhaustivité des racines retracées. Nous présentons ici deux voies d'amélioration possible de cet algorithme.

#### (1) Détermination du rythme d'initiation des racines

Comme on l'a dit précédemment, les temps d'initiation  $t_0$  des racines retracées ont été prédits conjointement avec la vitesse de croissance de ces racines. Par ailleurs, la température a été mesurée tout au long de l'expérience. Son évolution, en moyennes journalières, est reprise à au Chapitre III.5.a. Nous pourrions, à partir de cela, retracer l'accumulation des températures tout au long de l'expérience et exprimer l'apparition des racines en fonction de la somme des degrés-jours mesurées à leur temps d'initiation.

Si les temps d'initiation sont renvoyés par l'algorithme implémenté, la racine à laquelle ils correspondent n'est pas spécifiée. L'algorithme ne renvoie pas le numéro de la racine, selon un classement d'âge basé sur les hypothèses 1. et 2.. Une modification de l'algorithme permettant ce renvoi est nécessaire.

#### (2) Affinage de la détermination des ancras parasites et manquantes

La première voie d'amélioration présentée ci-dessus rend possible cet affinage. En effet, on s'attend à voir une relation linéaire entre le nombre de racines apparues et la somme des degrés-jours à un instant donné de la croissance d'une plante.

Supposer des points probablement manquants ou parasites sur cette droite, en cohérence avec les ancras relevées, nous permettrait de déterminer si certaines de ces dernières sont manquantes ou parasites.

## Chapitre VII

# Conclusions générales

La culture d'un grand nombre de plants d'orge dans un système aéroponique original offre la possibilité d'effectuer, grâce à un système de lecture de codes-barres, un grand nombre de mesures et d'observations sur la croissance de ces plants dans cet environnement. Ces mesures et observations s'opèrent sans destruction ni invasion du matériel végétal et nécessitent très peu de manipulations. Les plants observés croissent donc d'une manière que l'on espère optimale dans un environnement contrôlé.

Le premier objectif de ce mémoire était de proposer plusieurs méthodes de retraçage de racines mesurées à plusieurs dates, dans le cadre d'une expérience de suivi (mesure et observation) de la croissance de 151 lignées d'orge, issues du croisement de deux variétés différentes.

Les méthodes de résolution proposées sont au nombre de trois. Les deux premières méthodes se basent sur un dénombrement des configurations de correspondance entre racines basé sur un arrangement combinatoire. Cette façon de procéder est appréciable pour deux dates de mesures, mais s'avère rapidement évasive pour mettre en correspondance des racines mesurées sur trois dates ou plus. En effet, nous observons que des oublis (données manquantes) et des parasites (données fausses) « polluent » nos données. Les méthodes basées sur le dénombrement imposent cependant qu'une racine ne peut être retracée qu'en ayant des mesures à chaque date. Cette contrainte peut mener à des erreurs du fait de la présence d'oublis et de parasites.

La troisième méthode fait appel à l'optimisation d'un indicateur de proximité entre les mesures réalisées pour le faisceau adventif d'une plante et l'ensemble des vitesses de croissance des racines de ce faisceau. L'hypothèse principale de cette méthode est l'évaluation *a priori* du nombre de racines

potentiellement retraçables. Cela nous amène à poser certaines hypothèses sur l'âge et la croissance de ces racines. Dans une première phase, cette méthode détermine les vitesses les plus probables de l'ensemble des racines potentiellement retraçables et, dans une seconde phase, elle élabore les racines retraçées (ou combinaisons de correspondance) en respectant une contrainte de cohérence (au moins trois mesures par racine) et en veillant à ce qu'une même mesure n'appartienne pas à deux racines.

Cette troisième méthode est celle que nous avons retenue. En supposant d'abord des vitesses et ensuite des combinaisons de correspondance, elle procède de manière inverse par rapport aux précédentes et permet, de ce fait, d'intégrer la double « pollution » sus-mentionnée. Néanmoins, du fait de la contrainte de cohérence, elle ne relève que les racines jugées pertinentes et n'est donc pas exhaustive dans le retraçage des racines.

Nous prédisons les vitesses de croissance des racines retraçées de cette manière ainsi que leur temps d'initiation. Ces vitesses prédites ont deux possibilités d'utilisation. La première, présentée dans la synthèse bibliographique, est l'utilisation de ces vitesses comme paramètre d'entrée dans un modèle de simulation racinaire, *Root Typ*. L'intérêt de ce modèle réside dans le fait qu'il peut, à partir d'un nombre restreint d'inputs, reproduire une image réaliste de l'ensemble du système racinaire laquelle peut être utilisée pour décrire l'exploration du sol. La seconde possibilité est une analyse du déterminisme génétique du caractère 'vitesse de croissance'. Cette analyse a mené à identifier un locus sur le chromosome 4H au marqueur *p17m62f* de l'orge (Dumasy, 2007), vraisemblablement impliqué dans le contrôle de la vitesse de croissance.

Quelques limitations du modèle retenu méritent néanmoins de retenir notre attention.

Premièrement, la méthode retenue a seulement été évaluée sur des faisceaux adventifs. Elle n'a pas été testée sur le faisceau séminal des mêmes plantes, principalement parce que les données dont nous disposons sur celles-ci étaient insuffisantes (3 dates de mesures). Appliquer cette évaluation aux faisceaux séminaux n'en reste du moins pas infaisable. Les hypothèses de croissance posées doivent cependant être revues en terme de vitesse de croissance supposée. En outre, dans notre cas, les racines séminales sont mesurées avant l'apparition de racines adventives. Les données parasites sont donc nettement moins fréquentes et seulement imputables à une prise de mesure erronée (et non pas à une interprétation erronée du type de racine par l'expérimentateur).

Deuxièmement, ce modèle évalue la pertinence d'une racine retracée sur base d'un modèle de croissance linéaire. Il peut s'avérer nécessaire d'envisager d'autres modèles de croissance qu'une croissance strictement linéaire. En effet, dans un environnement différent, on peut s'attendre à une évolution différente des racines (courbe logarithmique, du fait d'un manque en nutriment). Par ailleurs, la croissance linéaire n'est supposée telle que sur la durée de notre expérience, c'est-à-dire

jusqu'au stade trois talles.

Enfin, nous relevons deux perspectives d'application de la méthode retenue. La première concerne les temps d'initiation des racines retracées, également prédits. La connaissance de ce paramètre permettrait d'étudier le rythme d'apparition des racines adventives. En outre, il peut également contribuer à une amélioration de l'algorithme en devinant les données manquantes sur base d'un rythme d'apparition attendu. La seconde perspective est plus large. Elle concerne la problématique des mesures de racines menées dans un environnement différent du nôtre. Les mesures de racines en champ, par exemple, sont certes plus malaisées qu'en aéroponie, mais aussi bien souvent destructrices. Une méthode analogue à celle que nous retenons, appliquée au cas d'observations en champ, pourrait s'avérer être un soutien judicieux.



# Glossaire

**Ancre** : Mesure de longueur d'un axe racinaire considérée comme la première mesure de cet axe. Donc, cette mesure est « révélatrice » d'un axe racinaire.

**Combinaison de correspondance** : pour un individu-plante mesuré sur  $x$  ( $x \geq 2$ ) dates différentes, résultat de la mise en correspondance de  $x$  mesures ou  $x$  numéros identificateurs de racine, appartenant chacun à une des  $x$  dates. Il s'agit donc de la définition d'un *couple de correspondance* élargie à plus de deux dates. En notation vectorielle, c'est un vecteur  $\mathbf{c}$  ( $1 \times x$ ) tel que si  $A_1, \dots, A_m$  sont les mesures de la première date,  $B_1, \dots, B_n$  celles de la seconde et ainsi de suite jusque  $X_1, \dots, X_p$ , les mesures de la dernière date, alors

$$\forall i=1, \dots, m, \forall j=1, \dots, n, \dots, \forall z=1, \dots, p \\ \mathbf{c} = [A_i B_j \dots X_z]$$

Une combinaison est valide si elle ne contient aucun emplacement vide.

**Configuration de correspondance** : pour un individu-plante donné, ensemble des couples ou combinaisons de correspondance sur deux ou plusieurs dates de mesure. En notation matricielle, c'est une matrice ( $n \times m$ ) pour laquelle  $n$  est le nombre de racines pour lesquelles une correspondance est faisable sur les différentes dates et  $m$ , le nombre de dates de mesure différentes. Dans cette matrice, chaque ligne est un vecteur  $\mathbf{c}$ , *combinaison de correspondance* sur  $m$  dates de mesure.

Une configuration est valide si les couples ou combinaisons la composant sont constitués, à chaque date, d'éléments identifiés comme différents pour tous les couples et combinaisons.

**Couple de correspondance** : pour un individu-plante mesuré sur deux dates différentes, résultat de la mise en correspondance de deux mesures ou deux numéros identificateurs de racine, appartenant chacun à une des deux dates. En notation vectorielle, c'est un vecteur  $\mathbf{c}$  ( $1 \times 2$ ), tel que si  $A_1, \dots, A_m$  sont les mesures de la première date et  $B_1, \dots, B_n$  sont les mesures de la seconde date, alors

$$\forall i=1, \dots, m, \forall j=1, \dots, n \\ \mathbf{c} = [A_i B_j]$$

**Faisceau** : ensemble de racines d'un système racinaire fasciculé tel que celui de l'orge. On parlera de faisceau séminal ou faisceau adventif, distinguant ainsi les racines sur base de leur origine.

# Références bibliographiques

*Liste d'ouvrages et d'articles consultés et cités*

**Anderson, P.M., Oelke, E.A., Simmons, S.R. (2002).** Growth and Development Guide for Spring Barley. University of Minnesota. Twin Cities, USA.

**Atkinson, D. (2000).** Root Characteristics : Why and What to Measure. In: Smit, A.L., Bengough, A.G., Engels, C., van Noordwijk, M., Pellerin, S., van de Geijn, S.C. (éds). Root Methods: A Handbook. Springer, Heidelberg, Ch. 1, 1-32.

**Bogaert, P. (2003).** Probabilités pour Scientifiques et Ingénieurs. Cours (BIR1203). Université catholique de Louvain, Louvain-la-Neuve, Belgique.

**De Dorlodot, S., Forster, B., Pagès, L., Price, A., Tuberosa, R., Draye, X. (2007).** Root system architecture: opportunities and constraints for genetic improvement of crops. Trends in Plant Science (sous presse)

**De Dorlodot, S., Bertin, P., Baret, P., Draye, X. (2005).** Scaling up quantitative phenotyping of root system architecture using a combination of aeroponics and image analysis. Aspects of Applied Biology **73**: 41-54.

**De Dorlodot, S. (2007).** Root System Architecture: A genetic analysis in rice. Thèse de doctorat. Université catholique de Louvain, Louvain-la-Neuve, Belgique.

**Draye, X. (2005).** Productions Végétales (Partim I). Cours (BIRA2108A). Université catholique de Louvain. Louvain-la-Neuve, Belgique.

**Drouet, J.L., Pagès, L. (2007).** GRAAL: Growth, Architecture, Allocation. In: Vos, J., Marcelis, L.F.M., de Visser, P.H.B., Struik, P.C., Evers, J.B. (éds). Functionnal-Structural Plant Modelling in Crop Production. Springer. Heidelberg. Ch. 14, 165-174.

**Dumasy, J.F. (2007).** Analyse génétique de la croissance racinaire de l'orge (*Hordeum vulgare ssp spontaneum*). Mémoire de fin d'études. Université catholique de Louvain. Louvain-la-Neuve, Belgique.

- Ellis, R.P., Forster, B.P., Gordon, D.C., Handley, L.L, Keith, R.P., Lawrence, P., Meyer, R., Powell, W., Robinson, D., Scrimgeour, C.M., Young, G., Thomas, W.T.B. (2002).** Phenotype/genotype associations for yield and salt tolerance in a barley mapping population segregating for two dwarfing genes. *Journal of Experimental Botany*, **371**: 1163-1176.
- Kahn, F. (1977).** Analyse structurale des systèmes racinaires des plantes ligneuses de la forêt tropicale dense humide. *Candollea* **32**: 321-358.
- Lagarias, J.C., Reeds, J.A., Wright, M.H., Wright, P.E. (1998).** Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J. Optim.* **9**: 112-147.
- Lambot, S., Javaux, M., Hupet, F., Vanclooster, M. (2004).** Inverse Modeling Techniques to Characterize Transport Processes in the Soil-Crop Continuum. *In: Alvarez-Benedi, I., Munoz-Carpena, R. (éds). Soil-Water-Solute Processes in environmental systems . CRC Press, London. Ch. 20, 693-714.*
- Lejeune, A., Auquièrre, J.P., (2001).** Biologie Végétale. Cours (BIOL1150). Université catholique de Louvain. Louvain-la-Neuve, Belgique.
- Lynch, J. (1995).** Root Architecture and Plant Productivity. *Plant Physiology* **109**: 7-13.
- May, L.H., Chapman, F.H., Aspinall, D. (1965).** Quantitative studies of root development. I. The influence of nutrient concentration. *Australian Journal of Biological Sciences* **18**: 25-35.
- Pagès, L., Asseng, A., Pellerin, S., Diggle, A. (2000).** Modelling Root System Growth and Architecture. *In: Smit, A.L., Bengough, A.G., Engels, C., van Noordwijk, M., Pellerin, S., van de Geijn, S.C. (éds). Root Methods: A Handbook. Springer. Heidelberg, Deutschland. Ch. 4, 113-146.*
- Pagès, L., Kervella, J. (1990).** Growth and development of root systems: Geometrical and structural aspects. *Acta Biotheoretica* **38**: 289-302.
- Pagès, L., Lecompte, F., Ozier-Lafontaine, H. (2001) .** The relationships between static and dynamic variables in the description of root growth. Consequences for field interpretation of rooting variability. *Plant and soil* **236**: 19-31.
- Pagès, L., Vercambre., G., Drouet, J.L., Lecompte, F., Collet, C., Le Bot, J. (2004).** Root Typ: a generic model to depict and analyse the root system architecture. *Plant and Soil* **258**: 103-119.
- Pellerin, S., Tabourel, F. (1995).** Length of the apical unbranched zone of maize axile roots: its relationship to root elongation rate. *Environmental and Experimental Botany* **35**: 193-200.
- Pison, S. (2005).** Modélisation inverse pour l'optimisation des émissions surfaciques à l'échelle régionale. Thèse de doctorat. Université Paris XII – Val-de-Mame. Paris, France.
- Tarantola, A. (1987).** Inverse Problem Theory: Methods for data fitting and models parameters estimation. Elsevier (éd), Amsterdam, 614 p., pp 5-14.
- Tsegaye, T., Mullins, C.E., Diggle, A.J. (1995).** An experimental procedure for obtaining input

parameters for the « ROOTMAP » root simulation program for peas (*Pisum sativum* L.). Plant and Soil 172: 1-16.

**Vézina, A.F., Pahlow, M. (2003).** Reconstruction of ecosystem flows using inverse methods: how well do they work? Journal of Marine Systems **40-41**: 55-77.

*Sites Internet consultés*

**Institute for Protection and Security of the Citizen (IPSC), AGRIFISH Unit.**  
[agrifish.jrc.it](http://agrifish.jrc.it)

**Institut National de la Recherche Agronomique.**  
[www.inra.fr](http://www.inra.fr)

**Prat et Rubinstein (2005).** Structure et Croissance de la racine. Cours multimédia. Université Pierre et Marie Curie, Paris, France.  
[www.snv.jussieu.fr/bmedia/racine/index.htm](http://www.snv.jussieu.fr/bmedia/racine/index.htm)

**The Barley Modelling Page.**  
[www.mansfeld.ipk-gatersleben.de/bucksorlin/](http://www.mansfeld.ipk-gatersleben.de/bucksorlin/)

*Dictionnaire utilisé pour la rédaction du mémoire*

**Le Petit Larousse (1996)**, grand format. Editions Larousse, Paris, France. 1872 p.

*Logiciels utilisés pour l'analyse des données et pour la rédaction du mémoire*

Microsoft® Excel 2004 for Mac®. Version 11.0. ©2004, Microsoft Corporation.

Matlab® & Simulink®. Version 7.3.0.1266, Student Version. ©2006, The MathWorks.

OpenOffice.org. Version 2.2.0. ©2006, Sun Microsystems Inc.

# Annexes électroniques

Un CD d'annexes électroniques est joint à ce mémoire. Ce volume est dénommé « Mémoire Pierre FAUX (2007) ». Son contenu et son arborescence sont détaillées ici.

**Scripts MATLAB** relatifs au pré-traitement de l'information brute et aux algorithmes développés (dans l'ordre des chapitres):

```
// Mémoire Pierre FAUX (2007)/ ROOTMAT/...
    .../ rootmat.m
    .../ transcode.m
    .../ metainfo.m
    .../ feuille.m
    .../ pato.m
    .../ racine.m

// Mémoire Pierre FAUX (2007)/ CORRESVITR/...
    .../ corresvitR.m
    .../ Kmatrix.m

// Mémoire Pierre FAUX (2007)/ CORRESVITNR/...
    .../ corresvitnR.m
    .../ indic.m
    .../ Klmatrix.m
    .../ Rmatrix.m
    .../ choixmodele.m

// Mémoire Pierre FAUX (2007)/ CORREST03/ ...
    .../ indicateur2.m
    .../ indicateur3.m
    .../ interchange.m
    .../ occurence.m
    .../ occurence2.m
    .../ interchange2.m
    .../ correct0_graphes3.m
    .../ loinormale.m
```

**Tableaux contenant l'évaluation des vitesses** des racines du faisceau adventifs (ID0001 à ID0680)

```
// Mémoire Pierre FAUX (2007)/ EVALUATION/ ...
    .../ vitADVmf.xls
    .../ vitADV2.xls
```

**Mémoire de Jean-François DUMASY** (UCL-AGRO, 2007): « *Analyse génétique de la croissance racinaire de l'orge (Hordeum vulgare ssp spontaneum)* »

// Mémoire Pierre FAUX (2007)/ Mémoire J-F DUMASY (2007)/ ...  
.../ MémoireJF.pdf

Bonus-track : **Montage Vidéo de l'expérience**

// Mémoire Pierre FAUX (2007)/ BONUS/...

.../serres2.mov

# *Annexes*

**Annexe III. 1.:** Composition de la solution nutritive HEWITT (full et half).

**Annexe III. 2.:** Sessions de mesure détaillées.