Ecole nationale de la statistique
et de l'analyse de l'information

Projet de Fin d'étude

Étudiants 3A

Plans d'expérience optimaux pour le suivi dynamique de la croissance de couverts végétaux

Etudiants :
Cassandre Diaine
Hugo Perrichon
Etienne Rigaud
Romain Ruiz

Tuteurs :
Nicolas Parisey
Melen Leclerc

2020

Remerciements

Nous tenons à remercier nos deux tuteurs, Nicolas Parisey et Melen Leclerc, pour leur disponibilité et la qualité de leur encadrement malgré les contraintes imposées par la crise sanitaire actuelle. Nous avons beaucoup apprécié travailler avec eux sur ce projet qui nous a permis d'approfondir nos connaissances en optimisation et son application dans le domaine biologique.

Sommaire

1 Introduction **5**

2 Plans d'expérience : définitions des critères d'optimalité **6**
2.1 Plan D-optimal 6
2.2 Plan T-optimal 7
2.3 Plan DT-optimal 7
2.4 Application sur deux exemples simples 8
2.4.1 D-optimalité 8
2.4.2 T-optimalité 10
2.4.3 Plan DT-optimal 10

3 Modélisation de la croissance de couverts végétaux **12**
3.1 Présentation du cas d'étude 12
3.2 Description des trois modèles de croissance 12

4 Plans d'expériences séquentiels **15**
4.1 Principe général de la construction séquentielle de plans optimaux . . . 15
4.2 Evaluation de l'effet des plans 16

5 Mise en place de la D-optimalité **17**
5.1 Modèle 1 17
5.2 Modèle 2 19
5.3 Modèle 3 21
5.4 Synthèse sur la D-optimalité 23

6 GDT optimalité séquentielle **24**

7 Discussion **29**

8 Conclusion **31**

Bilbiographie **32**

Annexe **33**
A D-optimalité - Ellipses de confiance 33

Liste des figures

1 D-optimalité pour la défoliation de la canopée 9
2 D-optimalité pour le modèle de Verhulst 9
3 T-optimalité pour le modèle de Verhulst et le modèle de défoliation de la canopée 10
4 DT optimalité pour le modèle de défoliation de la canopée et le modèle de Verhulst 11
5 Proportion de sol recouvert par les feuilles des betteraves en fonction du temps en degrés-jours 14
6 Graphiques représentant le positionnement des points de mesure en fonction des différents plans étudiés précédemment pour le modèle 1 18
7 Graphiques représentant le positionnement des points de mesure en fonction des différents plans étudiés précédemment pour le modèle 2 20
8 Graphiques représentant le positionnement des points de mesure en fonction des différents plans étudiés précédemment pour le modèle 3 22
9 Graphique représentant 6 points GDT-optimum lorsque l'on applique directement la GDT sur les modèles 1 et 2 25
10 Graphique représentant 6 points GDT-optimum entre les modèles 1 et 2 pour la deuxième année après avoir appliqué une GDT séquentielle lors de la première année 26
11 Graphique représentant 6 points GDT-optimum en appliquant directement la GDT sur les modèles 1 et 3 27
12 Graphique représentant 6 points GDT-optimum entre les modèles 1 et 3 pour la deuxième année après avoir appliqué une GDT séquentielles lors de la première année 27
13 Ellipses de confiance à 95% pour les estimateurs du modèle 1 33
14 Ellipses de confiance à 95% pour les estimateurs du modèle 2 33
15 Ellipses de confiance à 95% pour les estimateurs du modèle 3 34

Abstract

To solve the optimization problems that arise when studying plant foliage growth, many solutions can be chosen. Specifically, Atkinson's paper explains the DT-optimal solution and its application to nonlinear growth models. The models studied are models with a fairly rapid growth phase followed by a more or less rapid decline phase. The DT-optimum is a combination of the D-optimum and the T-optimum. The D-optimum gives an estimate of the parameters and the T-optimum helps to discriminate two models. Thus, by combining these two aspects, the DT-optimum allows to optimize the time at which the observation should be performed. These three optimization methods are studied theoretically and then in a practical way with an application on foliage dynamics of sugar beet. To perform the optimization, local and sequential designs are used, with numerical resolution methods to obtain the results.

1 Introduction

La modélisation de la croissance de couvert végétaux est une question importante en biologie végétale, par exemple pour prédire les rendements [1] ou le développement des maladies [2, 3]. De nombreux modèles ont déjà été proposés pour décrire la dynamique non-linéaire de la croissance des plantes. Cette dernière peut être modélisée par des systèmes d'équations différentielles ordinaires dont les paramètres doivent être estimés à partir de données d'observation. L'ajustement et la comparaison de ces modèles n'est possible que lorsque les données, acquises par un expérimentateur ou un capteur, le permettent. Cependant, comme cette collecte de données expérimentales représente un coût humain et financier important, il est utile de pouvoir la raisonner et l'optimiser.

Une stratégie de collecte des données vise à établir des temps auxquels seront effectuées les mesures afin d'ajuster les paramètres et de sélectionner les modèles le mieux possible. Les plans d'expériences optimaux sont des outils permettant de répondre à cette problématique en cherchant à obtenir un maximum d'informations en un minimum d'expériences. Le problème des plans d'expérience consiste, le plus souvent, à définir le plan d'expérience de manière optimale pour minimiser la variance des estimateurs. La majorité des critères utilisés pour optimiser la collecte des données sont basés sur la matrice d'information de Fisher, qui mesure la quantité d'informations qu'une variable aléatoire observable porte sur des paramètres inconnus. Par exemple, la A-optimalité vise à minimiser la trace de l'inverse de la matrice d'information et la E-optimalité vise à maximiser la valeur propre minimale de la matrice d'information. Au cours de notre étude, nous nous focalisons uniquement sur les critères D, T et DT que nous introduirons par la suite [4, 5, 6] .

L'objectif de ce travail est d'appliquer des méthodes issues de la planification expérimentale optimale pour optimiser le suivi de la croissance des couverts dans le but d'ajuster et sélectionner des modèles de croissances. Dans un premier temps, nous définissons les différents critères d'optimalité utilisés (i.e. D-optimalité, T-optimalité, DT-optimalité). Dans un second temps, nous présentons trois modèles de croissance publiés qui vont nous servir de cas d'étude. Enfin, nous testons la construction de plans d'expériences séquentiels pour chacun des modèles à l'aide de données simulées.

Ce projet a été réalisé avec le langage R [7].

2 Plans d'expérience : définitions des critères d'optimalité

Prenons un modèle dynamique de régression non-linéaire :

$$y = \eta(t, \theta) + \epsilon \tag{1}$$

où t est le vecteur des variables indépendantes appartenant à un espace euclidien X, l'espace de design, θ est le vecteur des paramètres inconnus que nous voulons estimer, ϵ sont indépendants, normalement distribués avec une moyenne nulle et une variance constante.

L'objectif des plans optimaux est de trouver un ensemble de temps (observations) t_i avec $i = 1..m$ et le poids w_i associé à chaque observation. [8] Le plan peut être écrit comme :

$$\xi = \begin{pmatrix} t_1 & t_2 & ... & t_m \\ w_1 & w_2 & ... & w_m \end{pmatrix}$$

Par la suite, par souci de simplification, nous considérerons $\forall i, w_i = 1$.

2.1 Plan D-optimal

La D-optimalité est utilisée pour obtenir une estimation des paramètres avec une variance minimale. Le plan considéré comme D-optimal est le plan qui minimisera la variance des paramètres. Pour calculer le plan D-optimal, $\eta(t, \theta)$ est dérivé par la composante de θ. [6]

$$F(t, \theta) = \frac{\partial \eta(t, \theta)}{\partial \theta_j} \tag{2}$$

Ensuite, on peut obtenir la matrice d'information des estimations des moindres carrés pour θ :

$$M(\xi, \theta) = \sum_{i=1}^{m} w_i F(t_i, \theta)^T F(t_i, \theta) \tag{3}$$

Dans l'équation (3), nous voyons la matrice de Fisher des informations sur le paramètre θ. Elle peut être définie comme la matrice de covariance de la variance jacobienne de la fonction de log-vraisemblance. Si l'erreur est modélisée par une loi normale homoscédastique, utiliser l'estimateur du maximum de vraisemblance est équivalent à utiliser l'estimateur des moindres carrés. L'objectif est de minimiser la variance des paramètres afin de déterminer un plan statistiquement optimal. Si la matrice de Fisher est inversée, on obtient les limites inférieures de la matrice de covariance de θ, limites qui doivent être minimisées. Il s'agit donc bien de maximiser le logarithme du déterminant de la matrice de Fisher (le "D" dans D-optimal).

$$\xi_D^* = \max_{\xi} log|M(\xi, \theta)| \tag{4}$$

Afin de comparer les plans au plan D-optimal, les valeurs de D-efficacité peuvent être calculées. Le plan ayant la valeur la plus élevée est considéré comme le meilleur.

$$E^D(\xi) = \left(\frac{|M(\xi, \theta)|}{|M(\xi^*, \theta)|} \right)^{1/p} \tag{5}$$

où p est le nombre de paramètres à déterminer et ξ^* correspond au plan D-optimal.

2.2 Plan T-optimal

Contrairement aux plans D-optimaux qui sont utilisés pour obtenir une estimation des paramètres, les plans T-optimaux sont utilisés pour discriminer les modèles. Cette discrimination s'effectue entre deux modèles ou plus, dont l'un est considéré comme vrai. [6]

Considérons deux modèles différents :

$$\eta_1(t, \theta_1) \quad \eta_2(t, \theta_2)$$

Afin de déterminer le plan qui optimise la distinction des modèles, il faut fixer un modèle et étudier la maximisation de l'écart minimal entre les modèles η_1 et η_2. Par exemple, si η_1 est considéré comme fixe. [5] [4]

$$\Delta(\xi_T^*) = \max_{\xi} \Delta(\xi) \tag{6}$$

où

$$\Delta(\xi) = \sum_{i=1}^{k} w_i (\eta_1(t_i) - \eta_2(t_i, \hat{\theta_2}))^2 \tag{7}$$

et

$$\sum_{i=1}^{k} w_i (\eta_1(t_i) - \eta_2(t_i, \hat{\theta_2}))^2 = \inf_{\theta_2 \in \Theta_2} \sum_{i=1}^{k} w_i (\eta_1(t_i) - \eta_2(t_i, \hat{\theta_2}))^2 \tag{8}$$

Comme pour la D-optimalité, la T-efficacité peut être calculée comme suit.

$$E^T(\xi) = \frac{\Delta(\xi)}{\Delta(\xi^*)} \tag{9}$$

où ξ^* correspond au plan T-optimal.

2.3 Plan DT-optimal

Le plan DT-optimal donne une estimation des paramètres et une discrimination des modèles, c'est une combinaison de critères D et T. Le plan DT-optimal consiste à maximiser le critère D-optimal pour les deux modèles. Nous discutons du plan DT-optimum lorsque les modèles sont imbriqués, sinon nous discutons du plan GDT-optimum (DT généralisé). Les deux critères D et T sont différents dans la manière dont ils sont établis. Le critère T-optimal est une somme de carrés d'observations alors que le critère D-optimal est une maximisation d'un déterminant dépendant de variables. Par conséquent, pour comparer ces deux critères, une échelle commune est nécessaire, l'efficacité

et le produit pondéré des efficacités sont maximisés [6]. La fonction $\phi^{DT}(\xi)$ doit être maximisée :

$$\psi^{DT}(\xi) = \{E^T(\xi)\}^{1-\kappa}\{E^D(\xi)\}^{\kappa} \tag{10}$$

$$\psi^{DT}(\xi) = \left(\frac{\Delta(\xi)}{\Delta(\xi_T^*)}\right)^{1-\kappa}\left(\frac{|M(\xi,\theta)|}{|M(\xi_D^*,\theta)|}\right)^{\frac{\kappa}{p}} \quad 0 \leq \kappa \leq 1 \tag{11}$$

en appliquant la fonction logarithmique à l'équation (11),

$$log(\psi^{DT}(\xi)) = (1-\kappa)log(\Delta(\xi)) + \frac{\kappa}{p}log|M(\xi)| - (1-\kappa)log(\Delta(\xi_T^*)) + \frac{\kappa}{p}log|M(\xi_D^*)| \tag{12}$$

ξ_T^* et ξ_D^* sont constants. C'est pourquoi maximiser $log(\psi^{DT}(\xi))$ équivaut à maximiser $\phi^{DT}(\xi)$ de telle sorte que :

$$\phi^{DT}(\xi) = (1-\kappa)log(\Delta(\xi)) + \frac{\kappa}{p}log|M(\xi)| \tag{13}$$

Les plans présentés dans les parties 2.1 à 2.3 sont des plans localement optimaux, c'est-à-dire que le plan optimal est calculé selon des valeurs de paramètres supposées connues. Il existe des plans optimaux qui ne sont pas locaux, par exemple le plan en moyenne pour lequel les critères sont maximisés en espérance [9].

2.4 Application sur deux exemples simples

Pour illustrer l'utilisation des critères d'optimalité définis ci-dessus pour optimiser l'ajustement et la sélection de modèles de croissance, nous considérons deux modèles simples de croissance non-linéaire : un premier modèle qui a une croissance exponentielle suivie d'une phase de déclin (défoliation du couvert) (14), et un second qui correspond au modèle de Verhulst (c'est-à-dire la croissance logistique) (15). L'idée est maintenant d'appliquer la D-, T- et DT-optimalité à ces deux modèles joués. Nous considérons que l'expérience se déroule sur une période continue et fixe t, nous prendrons $t \in [0, 10]$.

$$f(t) = \eta_1(t,\theta_1) = u_0 exp((n - 0.5mt)t) \quad avec \quad \theta_1 = (n,m) \tag{14}$$

$$g(t) = \eta_2(t,\theta_2) = \frac{K}{1 + aexp(-rt)} \quad avec \quad \theta_2 = (K,r) \tag{15}$$

2.4.1 D-optimalité

Fixons $u_0 = 1$, θ comme $\theta = (n_0, m_0) = (0,5, 0,1)$ et considérons que seules deux mesures peuvent être effectuées. Grâce à l'équation (4), $M(\xi,\theta)$ peut être calculé et, en fonction de t_1 et t_2, le logarithme du déterminant de M peut être maximisé. Selon la D-optimalité, avec une méthode analytique, les deux mesures doivent être effectuées à $t_1 = 4,65$ et $t_2 = 8,65$ (Figure 1).

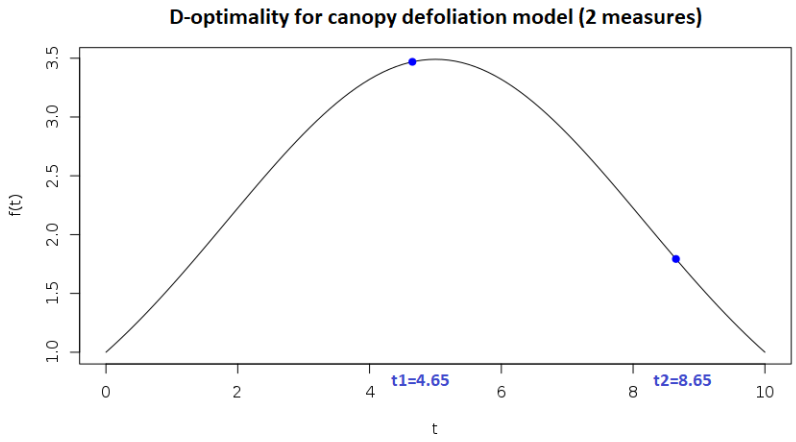

FIGURE 1 – D-optimalité pour la défoliation de la canopée

Ensuite, le plan optimal a été comparé à des plans non-optimaux obtenus par tirage aléatoire de t_1 et t_2 dans $\mathcal{U}(0, 10)$. Pour effectuer cette comparaison, la D-efficacité des plans aléatoires a été étudiée. Les estimateurs $\hat{n}$ et $\hat{m}$ ont été évalués 1 000 fois pour des valeurs aléatoires de t_1 et t_2 et une variance des observations $\sigma = 0.01$. Dans notre cas, seul 35% des plans non-optimaux ont un déterminant de la matrice de covariance observée inférieur à celui d'une simulation du plan optimal. Cela montre l'intérêt des plans optimaux qui sont en moyenne meilleurs que les autres modèles.

Le même raisonnement a été appliqué au modèle de Verhulst. Fixons $a = 1.9$, θ comme $\theta = (K, r) = (3, 1.2)$. Selon la D-optimalité, les deux mesures doivent être établies à $t_1 = 1, 55$ et $t_2 = max(t) = 10$ (Figure 2).

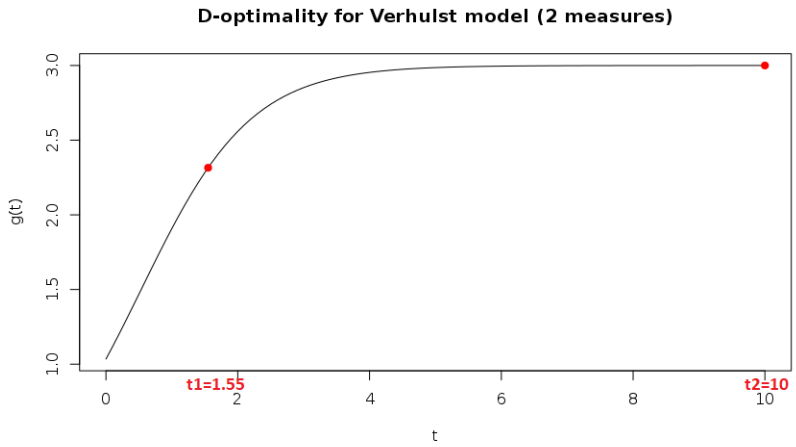

FIGURE 2 – D-optimalité pour le modèle de Verhulst

Comme attendu, il faut un point au cours de l'accroissement pour estimer r et un point à la fin pour estimer K.

2.4.2 T-optimalité

La T-optimalité peut être utilisée pour distinguer le modèle avec défoliation (14) et le modèle de Verhulst (15). Grâce à l'équation (6), le problème de maximisation à résoudre est :

$$\max_{t_1,t_2\in[0,10]} \Delta(\xi) = \max_{t_1,t_2\in[0,10]} \sum_{i=1}^{2} (f(t_i) - g(t_i))^2 \tag{16}$$

Par calcul numérique, on obtient que les solutions à ce problème sont $t_1 = t_2 = max(t) = 10$ (Figure 4)

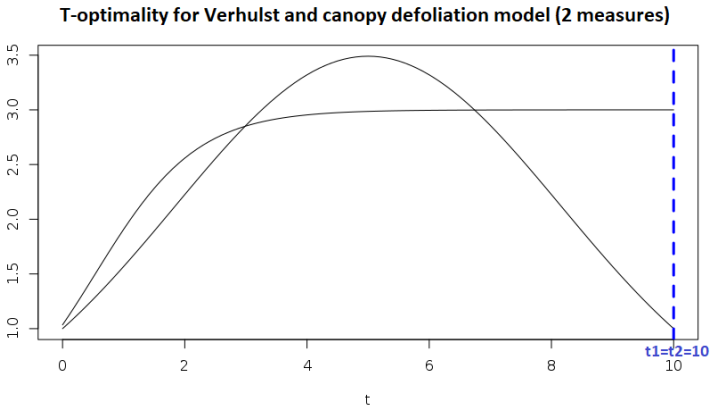

FIGURE 3 – T-optimalité pour le modèle de Verhulst et le modèle de défoliation de la canopée

2.4.3 Plan DT-optimal

Examinons maintenant le calcul de la DT-optimalité.

$$\max_{t_1,t_2\in[0,10]} \phi(\xi,\theta) = \max_{t_1,t_2\in[0,10]} (1-\kappa) log(\Delta(\xi)) + \frac{\kappa}{p} |M_{canopydefoliation}(\xi,\theta)|$$

Mettons $p = 2$ parce que deux mesures peuvent être établies et $\kappa = 0,5$ qui donne autant de poids à la D-optimalité (estimation des paramètres) qu'à la T-optimalité (distinction des modèles). La solution à ce problème est $t_1 = 6,23$ et $t_2 = max(t) = 10$, obtenue de manière analytique.

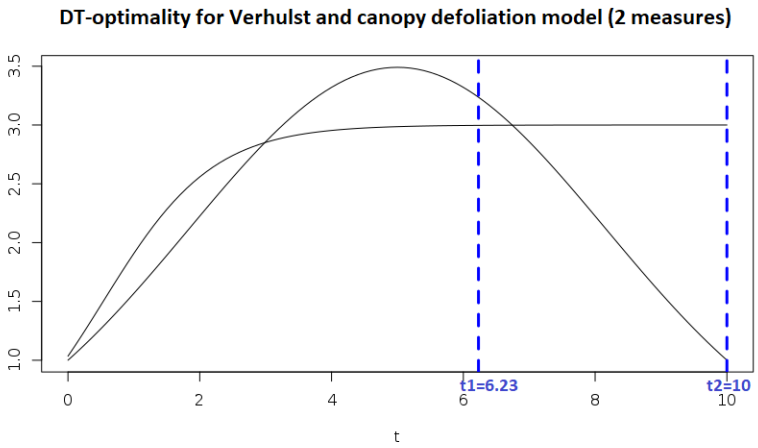

FIGURE 4 – DT optimalité pour le modèle de défoliation de la canopée et le modèle de Verhulst

La D-optimalité place pour les deux modèles un point lors de la phase de croissance et un à la fin de l'expérience. La T-optimalité place les points à la fin de l'expérience. On aurait donc ainsi pu s'attendre à ce qu'on ait, avec la DT-optimalité, un point sur la phase de croissance et un à la fin de l'expérience ce qui n'est pas le cas. Cela peut être du au fait qu'on est sur un critère qui est une moyenne ou au fait qu'on ait pris la DT-optimalité et non la GDT-optimalité.

Il aurait également pu être intéressant de faire varier la valeur de κ. En effet, κ permet de donner plus de poids au critère D ou au critère T, ce qui peut être intéressant dans des cas plus concrets.

3 Modélisation de la croissance de couverts végétaux

3.1 Présentation du cas d'étude

Pour de nombreuses plantes annuelles, au cours d'une saison elles vont croître jusqu'à une taille maximale avant que celles-ci ne diminuent (i.e. phénomène de sénescence). Les taux de croissance et de sénescence varient au cours de la saison en réponse à des changements dans le cycle de vie de la plante et dans l'environnement.

Nous considérons ici trois modèles de croissance proposés par Werker & Jaggard [10] dans un travail visant à analyser la dynamique de la fermeture du couvert (proportion de sol recouvert par le feuillage) de betteraves sucrières dans un champ cultivé. Les trois modèles prennent en compte la phase d'accumulation de la biomasse suivie d'une phase de défoliation (ou sénescence) et permettent donc de décrire toutes les phases de la dynamique du couvert observé dans les données.

3.2 Description des trois modèles de croissance

La composante mesurée de la plante (proportion de sol recouvert) est notée y . Le taux de croissance est séparé en un terme de croissance relative β et un terme de sénescence relative δ tel que $\mu = \beta - \delta$, où μ est une fonction du temps dont l'expression est différente pour chaque modèle.

Modèle I

Le premier modèle est défini par le système d'équations différentielles suivant :

$$\frac{dy}{dt} = \mu y$$
$$\frac{d\mu}{dt} = -k\left(\mu - \mu_{\min}\right)$$

où k est une constante déterminant la vitesse à laquelle μ atteint sa valeur finale μ_{min}. La résolution de ce système mène à une expression de $y(t)$ possédant quatre paramètres (k, μ_{min}, μ_0, y_0) avec μ_0 la croissance initiale et y_0 la valeur initiale de y à t_0 (temps de germination).

Modèle II

Le second modèle possède un paramètre en plus μ_{max} par rapport au précédent. Il est défini par le système d'équations différentielles suivant :

$$\frac{dy}{dt} = \mu y$$
$$\frac{d\mu}{dt} = -k\left(\mu - \mu_{\min}\right)\frac{\mu_{\max} - \mu}{\mu_{\max} - \mu_{\min}}$$

où μ_{max} est la limite de μ en $t \to -\infty$.

Modèle III

Le troisième modèle a cinq paramètres, il est défini par le système d'équations différentielles suivant :

$$\frac{dy}{dt} = \mu y\frac{(\kappa - y)}{\kappa}$$
$$\frac{d\mu}{dt} = -k\left(\mu - \mu_{min}\right)$$

où κ est la capacité limite. En théorie, $\kappa = 1$.

La figure 5 représente y en fonction du temps en ayant ajusté les paramètres avec des données issues de 12 ans de collecte sur une culture de betteraves. Les valeurs des paramètres sont fournies par Werker & Jaggard. Ici, y est la proportion de sol recouvert par les feuilles des betteraves et le temps est exprimé en degré-jours. L'estimation des paramètres est un problème d'optimisation globale. les auteurs ont ici choisi de partir de valeurs initiales qui ont été obtenues à l'aide d'une transformation linéaire des modèles. Les paramètres y_0, μ_0 et t_0 étant interdépendants, y_0 a été fixé pour pouvoir estimer les autres.

Degré-jours

Le degré-jour (DJ) est une unité de mesure utilisée pour décrire l'effet cumulé de la température sur la croissance de la plante au cours du temps. En notant T_{base} la température à partir de laquelle la plante se développe, le nombre de degré-jours N_{DJ} sur une journée est calculé de la manière suivante :

$$N_{DJ} = \frac{T_{max} + T_{min}}{2} - T_{base}$$

où T_{max} et T_{min} sont respectivement les températures maximale et minimale atteintes au cours de la journée. Le nombre de degré-jours étant minoré par 0, on fait l'hypothèse que T_{min} ne descend pas plus bas que T_{base}. Dans notre étude, la valeur de T_{base} est fixée à 3°C.

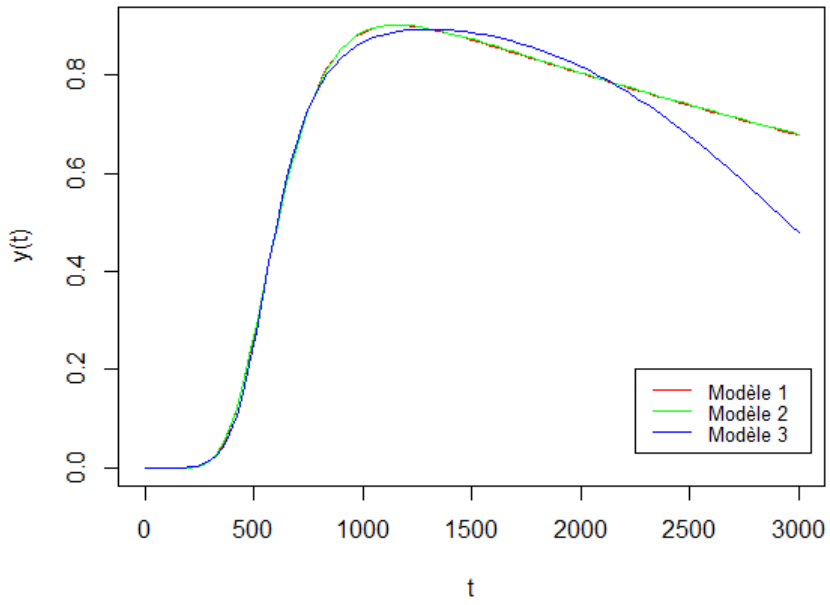

FIGURE 5 – Proportion de sol recouvert par les feuilles des betteraves en fonction du temps en degrés-jours

Les données ayant mené à l'ajustement des modèles ne sont pas disponibles, mais si on se réfère à la Figure 2 de l'article par Werker & Jaggard [10], on peut avoir une idée des données.

4 Plans d'expériences séquentiels

4.1 Principe général de la construction séquentielle de plans optimaux

La construction d'un plan séquentiel consiste à rajouter des points expérimentaux optimaux de manière itérative à un plan initial choisi avant le début de l'expérience. Plus le nombre d'observations augmente plus le modèle est précis, car les observations sont ajoutées à des temps d'observation optimaux du point de vue de la variance. Un plan séquentiel peut avoir un intérêt dans la pratique dans le cas où on ne sait pas quel budget allouer à l'expérience, le budget dépendant du nombre d'observations réalisées. Cette stratégie classique [11] permet d'avoir initialement un plan peu coûteux et d'enrichir itérativement ce plan .
La construction de plans séquentiels est particulièrement adaptée aux modèles non-linéaires car ils ont l'avantage de surmonter les problèmes liés à l'optimalité locale, ce qui n'est pas le cas des méthodes de calculs de plans optimaux présentées dans la partie 2 où le design optimal dépend du vecteur de paramètres. Mettre en place une stratégie de collecte séquentielle a donc à la fois des avantages théoriques et pratiques.

Construction d'un plan séquentiel simulé

La procédure générale est la suivante :

1. Choisir un vecteur de paramètres θ du modèle dit 'vrai' (à découvrir) et un vecteur de temps pour générer les premières observations
2. Estimer $\hat{\theta}$ avec la méthode des moindres carrés non-linéaires après avoir observé les données
3. Calculer le prochain temps d'observation optimal et mettre à jour le vecteur de temps et les observations.
4. Répéter les étapes 2 et 3 jusqu'à satisfaction

L'étape 3 est cruciale et va dépendre du critère d'optimalité que nous recherchons.

Soient deux modèles auxquels on veut appliquer un plan séquentiel DT-optimal. On note $\hat{\eta}_1(x, \xi_N)$ et $\hat{\eta}_2(x, \xi_N)$ les modèles estimés par moindre carrés non-linéaires après N temps observés. Atkinson a formalisé l'expression analytique de l'étape 3 pour la recherche d'un plan DT optimal . La prochaine observation est donnée par

$$x_{N+1} = \arg\max_{x \in \mathscr{X}} (1-\kappa) \left\{\hat{\eta}_1(x, \xi_N) - \hat{\eta}_2(x, \xi_N)\right\}^2 / \Delta(\xi_N) + (\kappa/p_1)\, d_1(x, \xi_{\mathrm{N}})$$

où p_1 est le nombre de paramètres du modèle 1 et $d_1(x, \xi_{\mathrm{N}})$ est le déterminant de la matrice d'information du modèle 1.

4.2 Evaluation de l'effet des plans

Dans la pratique, il est possible d'évaluer l'effet d'un plan et de comparer plusieurs plans entre eux, outre les critères d'efficacité présentés dans la partie 2. Pour la D-optimalité, deux plans D-optimaux mènent à des variances différentes sur chaque paramètre. Il est donc possible de comparer les variances des paramètres des différents plans, par exemple en dimension 2 avec des représentations graphiques avec des ellipses de confiance à 95%.

Pour la GDT-optimalité, étant donné que le but d'un plan est de différencier deux modèles, l'AIC, un critère classique pour la sélection de modèle, sera ici utilisé pour comparer deux plans : celui introduisant la plus grande différence d'AIC entre les deux modèle est celui qui les différenciera le mieux. Même si l'AIC n'est pas le meilleur critère pour les modèles non-linéaires, notre attendu est que l'effet sur la composante en log-vraisemblance sera une source de contraste suffisante.

5 Mise en place de la D-optimalité

Dans un premier temps, la D-optimalité est appliquée aux trois modèles. Pour chaque plan, le nombre de mesures est fixé à six, et ces mesures doivent être réalisées entre 0 et 2 500 degrés-jours. Des problèmes d'identificabilité ayant été rencontrés, le nombre de variables à estimer a été réduit. On ne considère plus que k et μ_{min} comme variables à estimer pour ajuster le modèle aux données. Pour simuler les données, un bruit gaussien centré de variance 0,05 est généré. Différents plans ont été mis en place tels que les plans séquentiels, réguliers et D-optimaux. De nombreux plans ont été testés afin de voir si certains plans différaient significativement des autres. Pour chaque modèle, six plans différents ont été retenus :

— Un premier plan où les six points sont directement déterminés à l'aide de la D-optimalité.
— un plan où les trois premiers points sont déterminés à l'aide de la D-optimalité avec la contrainte qu'ils doivent être inférieurs à 1 100 degrés-jours, puis les trois points suivants sont construits séquentiellement.
— deux plans séquentiels.
— deux plans réguliers.

Pour de nombreux plans, on a été confronté à des problèmes d'identificabilité lors de la première estimation de $\theta = (k, \mu_{min})$, c'est pourquoi les plans sont assez proches. Les deux premiers plans supposent de connaître déjà la valeur de θ.

5.1 Modèle 1

Ces différents plans ont d'abord été appliqués aux modèles 1. Les résultats sont répertoriés dans le tableau ci-dessous (Table 1) ainsi que leur représentation graphique (Figure 6) :

Modèle 1	t1	t2	t3	t4	t5	t6	det(M)	D efficacité
D-optimalité pour les 6 points (a)	513.72	515.06	2496.81	2499.54	2499.98	2500	154521363451	1
D-optimalité pour les trois premiers points avant 1 100 puis séquentiel (b)	464.71	464.71	655.20	2500	2500	2500	111547938473	0,72
Plan séquentiel (c)	500	800	1100	2500	2500	2500	69277060951	0,45
Plan séquentiel (d)	400	820	1200	2500	2500	2500	20534683994	0,13
Plan régulier (e)	500	800	1100	1500	1850	2500	28357062303	0,18
Plan régulier (f)	400	820	1200	1660	2080	2500	17673978880	0,11

TABLE 1 – Tableau récapitulatif des différents plans pour la D-optimalité du modèle 1

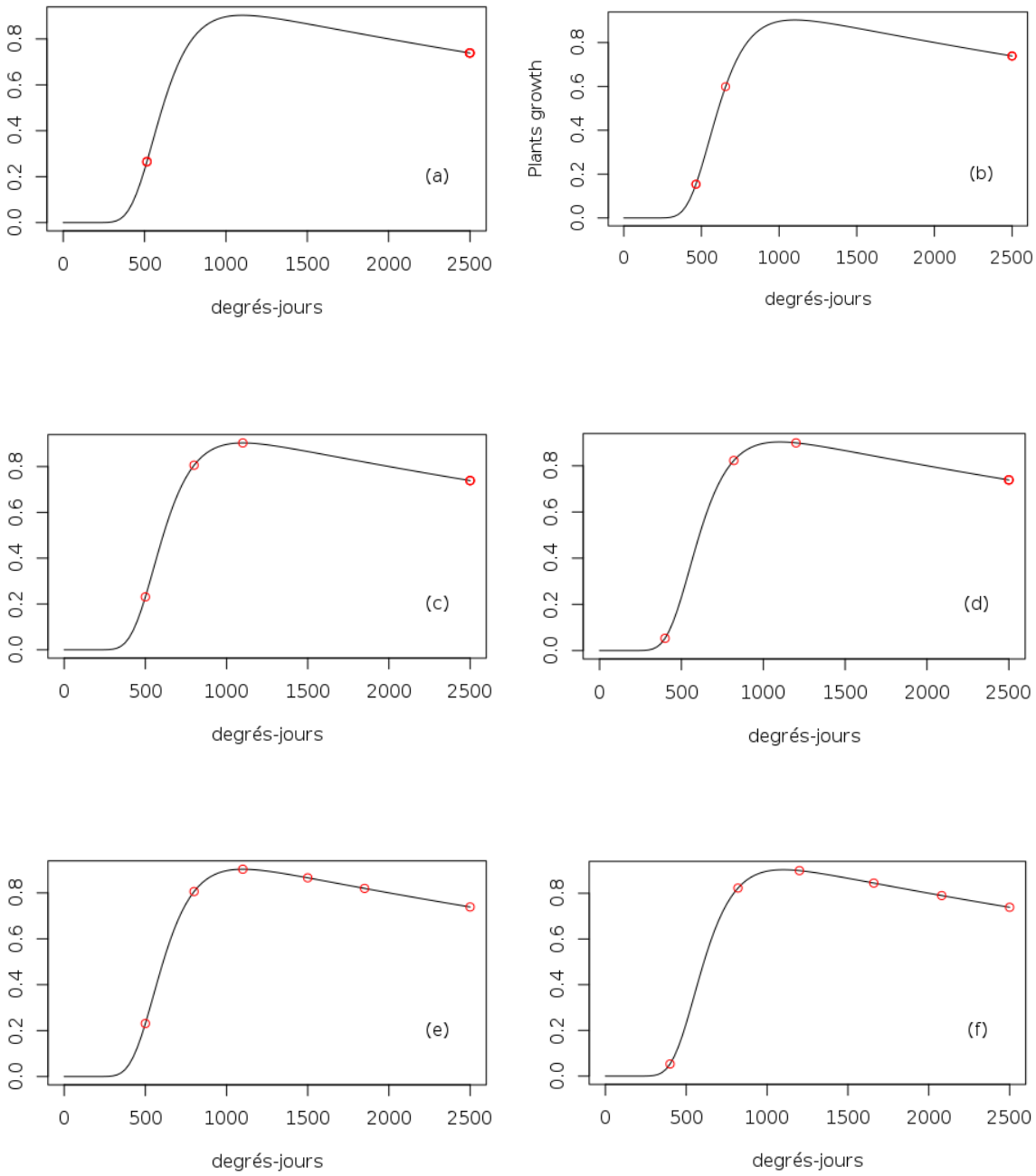

FIGURE 6 – Graphiques représentant le positionnement des points de mesure en fonction des différents plans étudiés précédemment pour le modèle 1
Les lettres sur les graphiques correspondent aux lettres des plans dans le tableau

Nous remarquons tout d'abord que pour les plans D-optimaux et les plans séquentiels, les trois dernières mesures se situent à la fin de l'expérience. On a essayé de mettre des bornes pour restreindre l'espace du plan pour voir si cela pouvait impacter les plans. Cependant, systématiquement, les points sont placés à la fin de l'expérience ce qui est assez logique étant donné la forme de la fonction. Concernant la D-efficacité des plans, le plan séquentiel (c) se démarque des autres plans séquentiels et réguliers, mais c'est

assez logiquement que le plan D-optimal est le meilleur.

Intéressons nous maintenant aux estimations de nos paramètres k et μ_{min} (Table 2). Au vu des résultats, la variance de μ est bien plus faible (divisée par 4) pour les plans séquentiels que pour les plans réguliers. Concernant k, les différences sont un peu moins significatives, mais les plans séquentiels restent meilleurs que les plans réguliers. Des ellipses de confiance à 95 % ont été réalisées afin de rendre plus visuel ce tableau (cf Annexes 13). Étudier la façon dont elles s'imbriquent peut-être intéressant.

Modèle 1	Estimateurs		Variance des estimateurs	
	μ_{min}	k	μ_{min} ($\times 10^{-10}$)	k ($\times 10^{-7}$)
D-optimalité pour les six points (a)	-0.00017717	0.00662124	6.3958	1.2674
D-optimalité pour les trois premiers points avant 1 100 puis séquentiel (b)	-0.00011939	0.00750106	3.2102	1.4069
Plan séquentiel (c)	-0.00012619	0.00743352	4.2891	1.6239
Plan séquentiel (d)	-0.0001258	0.0073967	4.2343	2.5493
Plan régulier (e)	-0.0001712	0.0066055	18.625	3.4578
Plan régulier (f)	-0.00018363	0.00639107	15.339	2.9195

TABLE 2 – Tableau récapitulatif de l'estimation des paramètres pour les différents plans pour le modèle 1

Il pourrait être intéressant de considérer une expérience sur deux années afin de ne pas avoir que des mesures à la fin du modèle. On considère un plan séquentiel en année 1. Le plan séquentiel (d) de la Table 1 est retenu car meilleur au sens de la D-efficacité. À la fin de ce premier plan, une estimation pour notre paramètre $\theta = (k, \mu_{min})$ est obtenue. C'est cette valeur de θ qui est utilisée pour faire de la D-optimalité localement l'année suivante. Donc au début d'année 2, les six mesures à réaliser en fonction de l'estimation obtenue de θ en année 1 sont générées. Cette modélisation peut être compliquée à la mise en place dans la réalité, car cette expérience est difficilement, parfaitement réplicable d'une année sur l'autre. Cependant, les mesures étant en degré-jours, cela reste possible. En mettant en place cette méthodologie, les mesures qui doivent être réalisées à : 498.95, 498.92, 2500.00 2500.00, 2500.00, 2500.00. Pour le modèle 1, il est bien visible que la majorité des points se trouvent à la fin en année 2. Ce plan est assez proche du plan (a) de la Table 1.

5.2 Modèle 2

On fait de même pour le modèle 2. Les résultats sont résumés ci dessous (cf Table 3 et Figure 7).

Modèle 2	t1	t2	t3	t4	t5	t6	det(M)	D efficacité
D-optimalité pour les six points (g)	485.41	486.20	2499.94	2499.94	2499.97	2499.98	114344238726	1
D-optimalité pour les trois premiers points avant 1 100 puis séquentiel (h)	465.19	732.19	732.20	2500	2500	2500	52653593446	0,46
Plan séquentiel (i)	500	800	1100	2500	2500	2500	44765233442	0,39
Plan séquentiel (j)	400	820	1200	2500	2500	2500	48012268653	0,42
Plan régulier (k)	500	800	1100	1500	1850	2500	18990149383	0,17
Plan régulier (l)	400	820	1200	1660	2080	2500	21433605141	0,19

TABLE 3 – Tableau récapitulatif des différents plans pour la D-optimalité du modèle 2

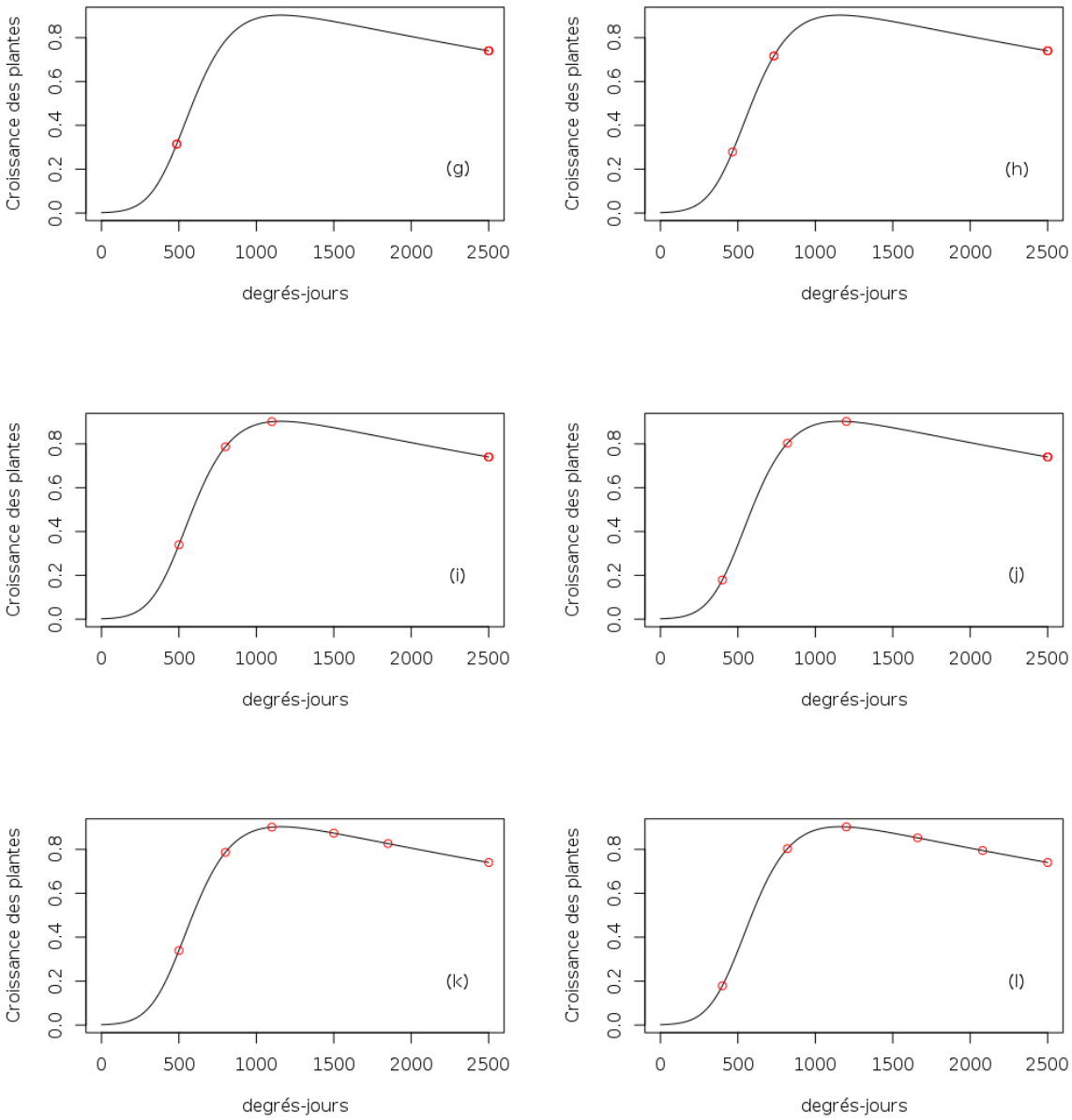

FIGURE 7 – Graphiques représentant le positionnement des points de mesure en fonction des différents plans étudiés précédemment pour le modèle 2

On remarque comme attendu que les résultats sont relativement similaires à ceux du modèle 1 étant donné qu'ils sont très proches. Cette fois-ci, au vu de la D-efficacité, les plans séquentiels se distinguent plus des plans réguliers. Comme pour le modèle 1 la majorité des points sont placés à la fin des modèles, soit à 2 500 degrés-jours.

Concernant les estimations de nos estimateurs k et μ_{min}, la variance de μ_{min} est bien plus faible pour les plans séquentiels et D-optimaux que pour les plans réguliers (cf Table 4). Pour k, la différence est moins marquée. Les ellipses de confiance se trouvent en Annexe 14. Les constats sont globalement les mêmes que pour le modèle 1.

Modèle 2	Estimateurs		Variance des estimateurs	
	μ_{min}	k	μ_{min} ($\times 10^{-10}$)	k ($\times 10^{-7}$)
D-optimalité pour les six points (g)	-0.00018909	0.00547470	7.8113	1.5214
D-optimalité pour les trois premiers points avant 1 100 puis séquentiel (h)	-0.00012874	0.00642389	4.6763	2.4727
Plan séquentiel (i)	-0.00013157	0.00647612	5.3354	1,405124
Plan séquentiel (j)	-0.00013642	0.00566758	5.4635	1,174052
Plan régulier (k)	-0.00017753	0.00552587	24.429	4.9414
Plan régulier (l)	-0.00019427	0.00531169	20.433	3.4347

TABLE 4 – Tableau récapitulatif de l'estimation des paramètres pour les différents plans pour le modèle 2

Comme pour le modèle 1, le plan qu'on obtiendrait en année 2 est étudié. Le plan séquentiel (j) de la Table 3 est appliqué pour l'année 1 car c'est le meilleur plan séquentiel au sens de la D-efficacité. Pour l'année 2, les mesures doivent être réalisées à 475.21, 475.21, 2500, 2500, 2500 et 2500 degré-jours. Cette fois-ci, les points sont toujours concentrés sur deux périodes, au début et à la fin de la courbe.

5.3 Modèle 3

Enfin, les calculs sont réitérés pour le dernier modèle, le modèle 3. Les résultats sont résumés ci-dessous (cf Table 5 et Figure 8).

Modèle 3	t1	t2	t3	t4	t5	t6	det(M)	D efficacité
D-optimalité pour les six points (m)	326.10	424.28	2500	2500	2500	2500	49973170338	1
D-optimalité pour les trois premiers points avant 1 100 puis séquentiel (n)	320.09	439.66	439.66	2500	2500	2500	44071473457	0.88
Plan séquentiel (o)	500	800	1100	2500	2500	2500	11178665540	0,22
Plan séquentiel (p)	400	820	1100	2500	2500	2500	20038173640	0,4
Plan régulier (q)	500	800	1100	1500	1850	2500	4437517304	0,09
Plan régulier (r)	400	820	1200	1660	2080	2500	8037632705	0,16

TABLE 5 – Tableau récapitulatif des différents plans pour la D-optimalité du modèle 3

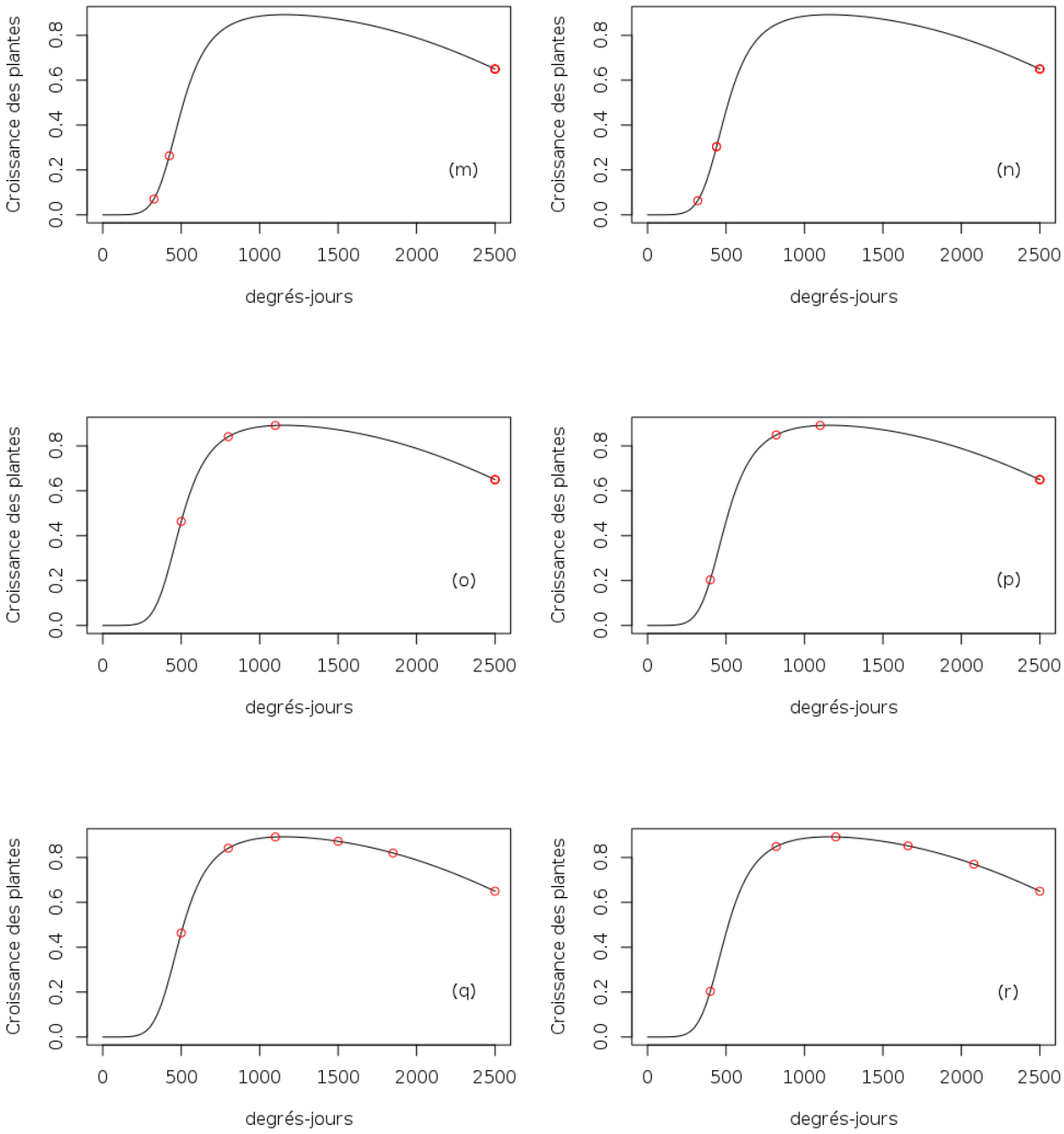

FIGURE 8 – Graphiques représentant le positionnement des points de mesure en fonction des différents plans étudiés précédemment pour le modèle 3
Les lettres sur les graphiques correspondent aux lettres des plans dans le tableau

Pour ce troisième modèle, la différence entre les plans séquentiels et réguliers est bien plus marquée au regard de la D-optimalité. Les points sont encore tous positionnés à la fin du modèle.

Ensuite, les estimateurs sont étudiés ci-dessous (cf Table 6).

Modèle 3	Estimateurs		Variance des estimateurs	
	μ_{min}	k	μ_{min} ($\times 10^{-10}$)	k ($\times 10^{-7}$)
D-optimalité pour les six points (m)	-0.0015851	0.0029881	148.72	0.18922
D-optimalité pour les trois premiers points avant 1 100 puis séquentiel (n)	-0.0012879	0.0032894	90.214	0.1617
Plan séquentiel (o)	-0.00136	0.00310	41.830	0.118436
Plan séquentiel (p)	-0.0014157	0.0030830	133.98	0.26233
Plan régulier (q)	-0.0013744	0.0031559	488.60	0.9167
Plan régulier (r)	-0.0014714	0.0030603	413.74	0.5670

TABLE 6 – Tableau récapitulatif de l'estimation des paramètres pour les différents plans pour le modèle 3

La variance de μ_{min} est beaucoup plus élevée pour ce modèle que pour les autres modèles. La variance est toujours plus faible pour les estimateurs issus des plans séquentiels que pour les plans D-optimaux.

Encore une fois, un plan en année 2 est étudié. Le plan séquentiel (p) est retenu. Les six mesures à réaliser en année 2 sont à : 422.08, 422.07, 2500.00, 2500.00, 2500.00 et 2500.00. Sur ce nouveau plan, les mesures devraient être réalisées à deux endroits, au tout début et à la fin.

5.4 Synthèse sur la D-optimalité

Pour le modèle 1, l'utilisation d'un plan séquentiel plutôt qu'un plan régulier est discutable mais le plan séquentiel donne tout de même des estimateurs avec une plus faible variance. Pour les modèles 2 et 3, l'intérêt de mettre en place un plan séquentiel est clairement visible comme le montre la D efficacité des plans et la faible variance des estimateurs. L'ensemble des plans pour les trois modèles conseillent de placer les 3 derniers points au même endroit soit à la fin de l'expérience ce qui est assez cohérent avec le fait que ce sont des modèles qui finissent avec une décroissance. Ce n'est néanmoins pas forcément réaliste dans des cas pratiques (notamment dans le cas des plans réguliers) alors que les modèles s'accordent sur l'utilité de réaliser plusieurs mesures en fin d'expérience.

6 GDT optimalité séquentielle

La GDT-optimalité est la DT-optimalité généralisée à l'ensemble des modèles. Le critère pour obtenir les points GDT-optimaux est écrit dans l'équation (17) et il doit être minimisé. Ce critère est constitué de la T-optimalité entre les deux modèles à laquelle on soustrait la D-optimalité concernant chacun des modèles. C'est ce qui la différencie de la DT-optimalité qui elle contenait seulement la D-optimalité pour un modèle et était plus adaptée à deux modèles emboîtés alors que la GDT est un critère pour deux modèles en général.

$$\Phi^{(\mathrm{GDT})}(\xi) = \sum_{j=1}^{c} a_j \log \Delta_j(\xi) - \sum_{k=1}^{m} (b_k/s_k) \log \left| A_k^{\mathrm{T}} M_k^{-1}(\xi) A_k \right| \tag{17}$$

a_j et b_k sont des coefficients permettant de donner plus ou moins d'importance à une partie de ce critère. Ici a_j et b_k sont posés tels que les deux parties du critères aient la même importance. La valeur m correspond au nombre de modèles dans ce critère tandis que la valeur c vaut ici 1 puisque ici nous considérons seulement la discrimination entre deux modèles.

La GDT-optimalité est utilisée ici car les modèles ne sont pas emboîtés. Dans un premier temps, une GDT locale a été réalisée pour les modèles 1 et 2 et pour les modèles 1 et 3. Ensuite, une GDT séquentielle a été faite pour ces mêmes couples de modèles. Les tables 7 et 8 présentent les résultats. Les temps d'observation obtenus avec la GDT séquentielle se situent à la fin de la courbe de croissance. C'est pourquoi nous avons décidé d'appliquer la GDT locale à une deuxième année de collecte, en prenant pour vecteur de paramètres les estimations obtenues lors de la GDT séquentielle de la première année.

Modèles 1 et 2	t1	t2	t3	t4	t5	t6
GDT optimalité pour les 6 points	393.83	524.20	1155.00	2118.15	2226.97	2383.28
GDT séquentielle	500	800	1100	2215.43	2500	2500
GDT deuxieme année	337.12	532.55	1429.48	1998.26	2079.90	2354.23

TABLE 7 – Tableau récapitulatif des différents plans pour la GDT-optimalité entre les modèles 1 et 2

Différents plans ayant été établis, il est maintenant possible d'évaluer la qualité de ces plans du point de vue de la variance des estimateurs et de l'AIC des modèles, comme présenté dans la partie 4.2. Nous calculons l'AIC des différents modèles, pour chaque plan. Deux plans correspondent à deux jeux de données différents, les valeurs ne sont donc pas comparable d'une ligne à l'autre. Nous faisons cependant l'hypothèse qu'un plan A discrimine plus les modèles qu'un plan B si la différence d'AIC entre les deux modèles pour le plan A est supérieure à la différence pour le plan B.

Les données simulées sont issues du modèle 1 auquel un bruit gaussien a été ajouté ($\epsilon \sim Normal(0, 0.01)$), il est logique d'avoir des AIC plus faibles pour le modèle 1 que pour le modèle 2.

[H]

Plan d'expérience \ Modèle	Modèle 1	Modèle 2
GDT pour les 6 points	-65.88	-54.49
GDT séquentielle	-73.28	-72.05
Plan régulier	-69.08	-65.59
GDT deuxième année	-85.15	-60.30

TABLE 8 – AIC des modèles 1 et 2 pour 3 plans d'expériences différents

L'AIC du modèle 1 est systématiquement plus faible que celui du modèle 2, ce qui est un résultat prévisible. Le plan GDT local pour la deuxième année de croissance est celui avec la plus grande différence d'AIC entre les deux modèles.

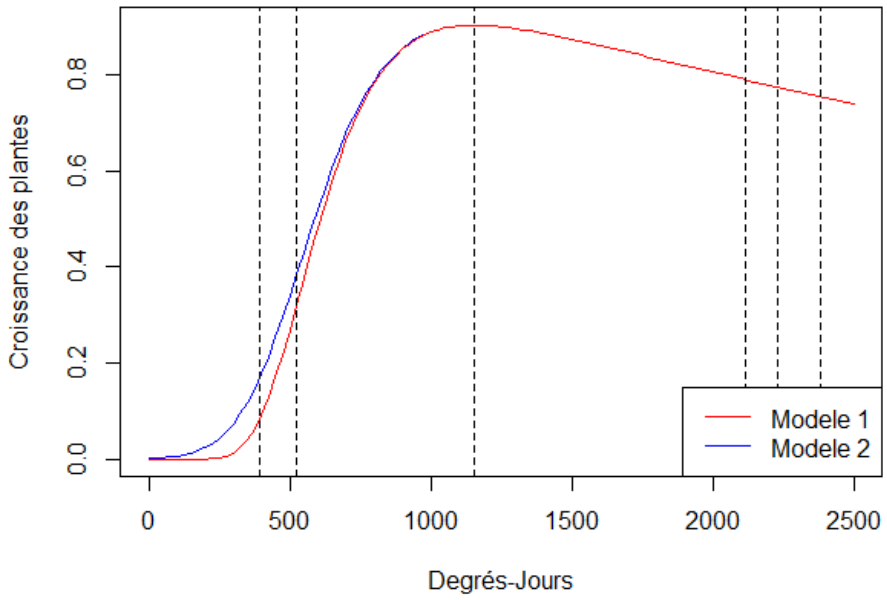

FIGURE 9 – Graphique représentant 6 points GDT-optimum lorsque l'on applique directement la GDT sur les modèles 1 et 2

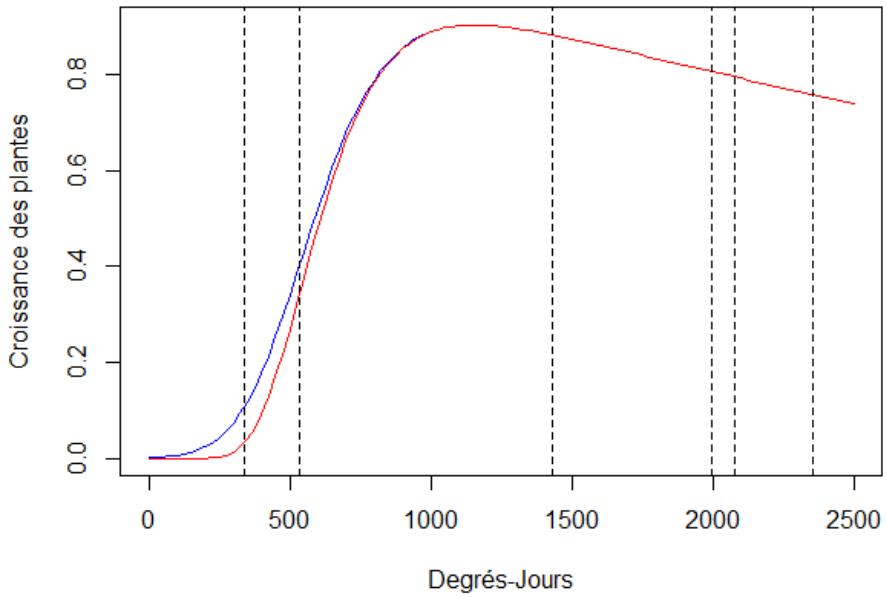

FIGURE 10 – Graphique représentant 6 points GDT-optimum entre les modèles 1 et 2 pour la deuxième année après avoir appliqué une GDT séquentielle lors de la première année

On fait ensuite le même travail pour le modèle 1 avec le modèle 3.

Modèles 1 et 3	t1	t2	t3	t4	t5	t6
GDT optimalité pour les 6 points	405.22	562.25	1155.00	1871.82	2254.48	2486.52
GDT séquentielle	500	800	1100	2500	2500	2500
Deuxieme année	505.2	505.2	2500.0	2500.0	2500.0	2500.0

TABLE 9 – Tableau récapitulatif des différents plans pour la GDT-optimalité entre les modèles 1 et 3

Modèle / Plan d'expérience	Modèle 1	Modèle 3
GDT pour les 6 points	-81.84	-66.50
GDT séquentielle	-76.68	-44.95
Plan régulier	-84.81	-46.55
GDT deuxième année	-84.20	-84.20

TABLE 10 – AIC des modèles 1 et 3 pour 3 plans d'expériences différents

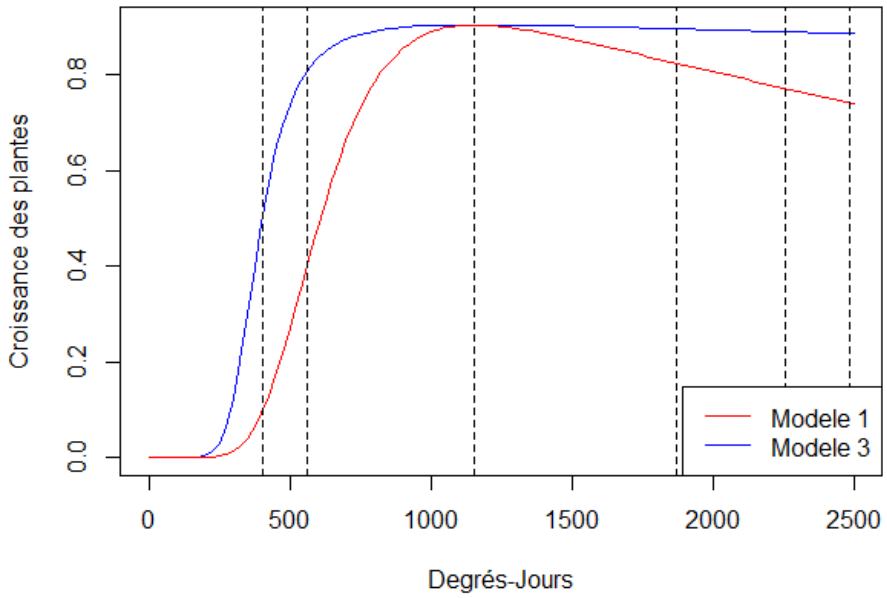

FIGURE 11 – Graphique représentant 6 points GDT-optimum en appliquant directement la GDT sur les modèles 1 et 3

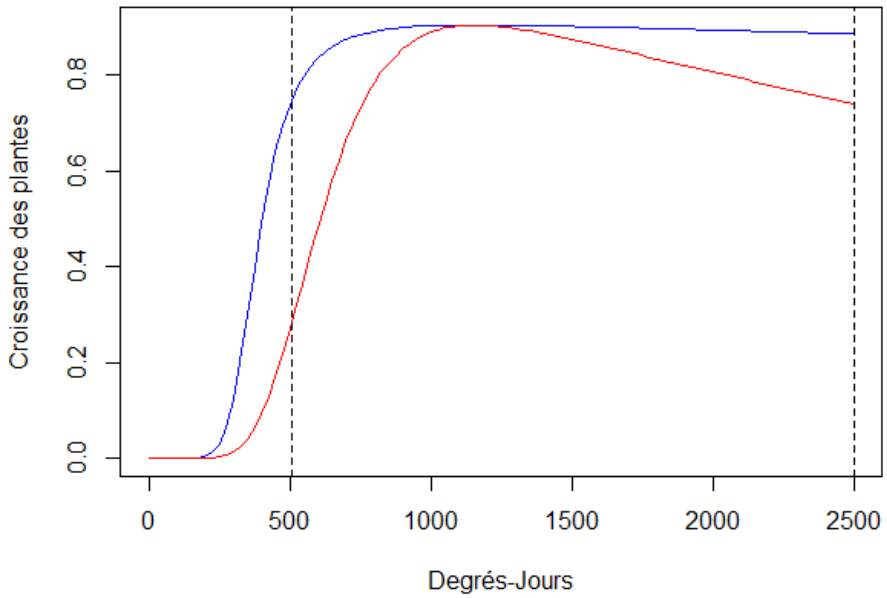

FIGURE 12 – Graphique représentant 6 points GDT-optimum entre les modèles 1 et 3 pour la deuxième année après avoir appliqué une GDT séquentielles lors de la première année

Compte-tenu de la différence entre l'allure des deux courbes des modèles 1 et 3, les résultats de la table 10 pour la deuxième année de collecte ne sont pas des résultats que nous attendions et sont peut-être dus à un problème d'optimisation numérique.

Avec le plan séquentiel, nous estimons les θ_i pour les deux modèles à partir d'un même signal, ici issu du modèle 1. Donc au fur et à mesure des itérations les prédictions du modèle 3 vont se rapprocher du signal issu du modèle 1. Ce qui augmente mécaniquement la partie log-vraisemblance de l'AIC. On s'attend donc, pour un nombre de points infini et parfaitement placés, à n'avoir un contraste entre les modèles, via l'AIC, que dans leur nombre de paramètres et la ou les partie(s) du signal produit par le modèle 1 que le modèle 3 ne pourra pas ou mal reproduire (géré au mieux par le critère GDT).

7 Discussion

Lors de la mise en place d'un plan optimal local, l'hypothèse est faite que la valeur des paramètres est connue, ce qui n'est pas toujours le cas. Le travail sur les plans séquentiels effectué à partir de la partie 4 vise à enlever la dépendance au fait de connaître les paramètres car la construction séquentielle de plans permet en partie de relâcher l'hypothèse. Il existe d'autres méthodes pour contourner cette hypothèse. Une méthode alternative est le Maximin [12] qui consiste à trouver le plan qui maximise le déterminant de la matrice d'information pour le vecteur de paramètres qui minimise cette même matrice, en d'autres termes on cherche le plan qui gère au mieux le pire des cas :

$$\max_{\xi} \min_{\theta \in \Theta} \log \mid M(\xi, \theta) \mid$$

Une autre méthode est celle du plan optimal en moyenne, consistant à borner les paramètres du modèle puis trouver le plan qui maximise l'espérance du déterminant de la matrice d'information, où le vecteur de paramètres est considéré comme une variable aléatoire, en d'autres termes on cherche le plan qui gère au mieux 'en moyenne' :

$$\max_{\xi} E_{\theta}(\log \mid M(\xi, \theta) \mid)$$

Un second point de discussion est le problème d'identifiabilité qui a été rencontré, nous menant à faire le choix d'estimer μ_{min} et k et de fixer les autres paramètres. On pourrait envisager de changer les paramètres qui sont fixés ou de modifier la structure du modèle pour aller plus loin dans l'étude. Il y a aussi un problème d'identifiabilité pour les plans séquentiels lorsque les premiers points des plans ne sont pas situés sur une bonne partie de la dynamique et ne permettent pas d'avoir une première estimation. Nous avons donc été confronté à des problèmes classiques pour des modèles applicatifs c'est-à-dire des problèmes d'identifiabilité structurelle (dans la formalisation mathématiques des processus) et d'identifiabilité en pratique (ici lié à l'échantillonnage).

Nous avons pris en compte les degré-jours comme unité de mesure dans tous nous plans. Si cette mesure permet d'être assez précis, cela peut se révéler assez compliqué à mettre en place dans la réalité. En effet, on ne peut pas réellement savoir quel jour la mesure va être faite. Un temps calendaire permet de planifier les mesures avec les personnes qui les réalisent à l'avance. Les plans en degré-jours sont plus faciles à mettre en place lorsque les mesures sont réalisées par des capteurs sans besoin d'une présence humaine. Si une personne doit se rendre dans le champ cela peut s'avérer plus laborieux. Il pourrait être intéressant de réfléchir à une conversion des degré-jour en un calendrier en fonction des températures moyennes sur la période de l'expérience.

Enfin, ce travail est réutilisable pour des expérimentations similaires. Une collecte de données faite à partir d'un réseau de capteurs intelligents permettrait de calculer de manière autonome les temps de mesure. Si c'est une personne qui réalise les mesures, il faut prendre en compte les contraintes liées aux disponibilité de cette personne, en

particulier se pencher sur la question des degrés jours, veiller à ne pas avoir des temps de mesure le week-end, prendre en compte les trajets dans le calcul du coût, etc.

8 Conclusion

Ce rapport s'appuie dans un premier temps sur les travaux d'Atkinson pour étudier la théorie des plans D, T et DT-optimaux et leur application sur les cas simples des modèles de sénescence et de Verhulst. Les plans D-optimaux sont utilisés pour obtenir des estimations de paramètres tandis que les plans T-optimaux sont utilisés pour discriminer les modèles. Le plan DT-optimal, ou GDT-optimal si les modèles ne sont pas emboîtés, est une combinaison des critères T et D.

Le rapport se penche ensuite sur un cas d'étude d'analyse de la dynamique de la croissance de betteraves sucrières, avec pour contraintes d'effectuer un maximum de six observations entre 0 et 2500 degrés-jours. En s'appuyant sur le papier de Werker et Jaggard, 3 modèles ont été considérés pour construire les différents plans d'expérience. Les plans obtenus par construction séquentielle D-optimale conseillent de placer les observations à la fin de cet intervalle de temps. Les plans localement D-optimaux ont une meilleure précision que les plans séquentiels, à condition de connaître préalablement une bonne estimation des paramètres, ce qui n'est pas forcément le cas. Le modèle 3 par exemple semble bénéficier grandement des plans séquentiels (en terme de variance des paramètres). Une solution est donc de réaliser un plan séquentiel sur la première année afin d'avoir une première estimation des paramètres, puis d'utiliser cette estimation pour réaliser un plan localement D-optimal.

Concernant la GDT-optimalité, nous avons mis en place des plans DT-optimaux et séquentiels afin de discriminer au mieux nos modèles et d'obtenir une estimation des paramètres avec une faible variance. Avec les plans séquentiels, plus on avance dans l'expérience, plus l'estimation des paramètres fait que les deux modèles sont de plus en plus proches. Encore une fois, il est intéressant de faire un plan séquentiel en année 1 pour obtenir une estimation des paramètre et faire un plan GDT-optimal local en année 2.

Finalement, de vraies expérimentations sont prévues sur des modèles biologiques. Ces expérimentations pourraient directement bénéficier de ce travail sur le sujet des plans d'expériences appliqué au suivi dynamique de la croissance de couverts végétaux.

Bibliographie

Références

[1] T. Setiyono, K. Cassman, J. Specht, A. Dobermann, A. Weiss, H. Yang, S. Conley, A. Robinson, P. Pedersen, and J. De Bruin, "Simulation of soybean growth and yield in near-optimal growth conditions," Field Crops Research, vol. 119, no. 1, pp. 161–174, 2010.

[2] P. Casadebaig, G. Quesnel, M. Langlais, and R. Faivre, "A generic model to simulate air-borne diseases as a function of crop architecture," PLoS One, vol. 7, no. 11, p. e49406, 2012.

[3] N. J. Cunniffe and C. A. Gilligan, "Invasion, persistence and control in epidemic models for plant pathogens : the effect of host demography," Journal of the Royal Society Interface, vol. 7, no. 44, pp. 439–451, 2010.

[4] E. Rady, M. Abd El-Monsef, and M. Seyam, "Relationships among several optimality criteria," Interstat, vol. 15, no. 6, pp. 1–11, 2009.

[5] H. Dette, V. B. Melas, P. Shpilev, et al., "T-optimal designs for discrimination between two polynomial models," The Annals of Statistics, vol. 40, no. 1, pp. 188–205, 2012.

[6] A. C. Atkinson, "Dt-optimum designs for model discrimination and parameter estimation," Journal of Statistical planning and Inference, vol. 138, no. 1, pp. 56–64, 2008.

[7] R Core Team, R : A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2017.

[8] M. Amo-Salas, E. Delgado-Márquez, L. Filová, and J. López-Fidalgo, "Optimal designs for model discrimination and fitting for the flow of particles," Statistical Papers, vol. 57, no. 4, pp. 875–891, 2016.

[9] A. P. L. Pronzato, Design of Experiments in Nonlinear Models. Springer-Verlag New York, 2013.

[10] A. Werker and K. Jaggard, "Modelling asymmetrical growth curves that rise and then fall : Applications to foliage dynamics of sugar beet (beta vulgarisl.)," Annals of Botany, vol. 79, no. 6, pp. 657–665, 1997.

[11] A. Wald, "Sequential tests of statistical hypotheses," The Annals of Mathematical Statistics, vol. 16, no. 2, pp. 117–186, 1945.

[12] R. T. A.C. Atkinson, A.N. Donev, "Optimum experimental designs with sas," Oxford University Press, pp. 258–259, 2007.

Annexe

A D-optimalité - Ellipses de confiance

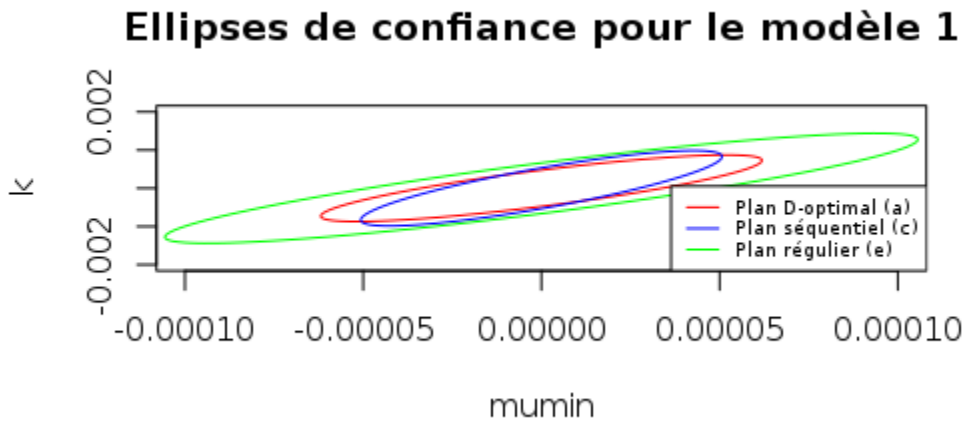

FIGURE 13 – Ellipses de confiance à 95% pour les estimateurs du modèle 1
La légende fait référence au plan a, c et e de la Table 1

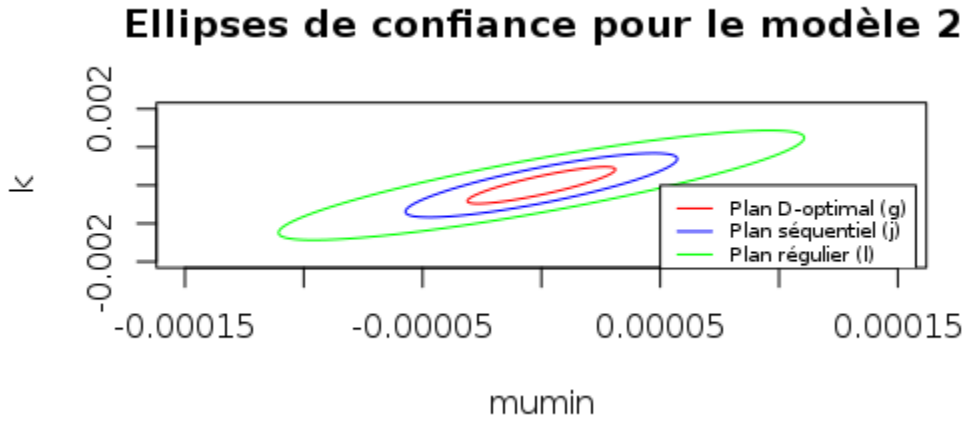

FIGURE 14 – Ellipses de confiance à 95% pour les estimateurs du modèle 2
La légende fait référence au plan g, i et l de la Table 3

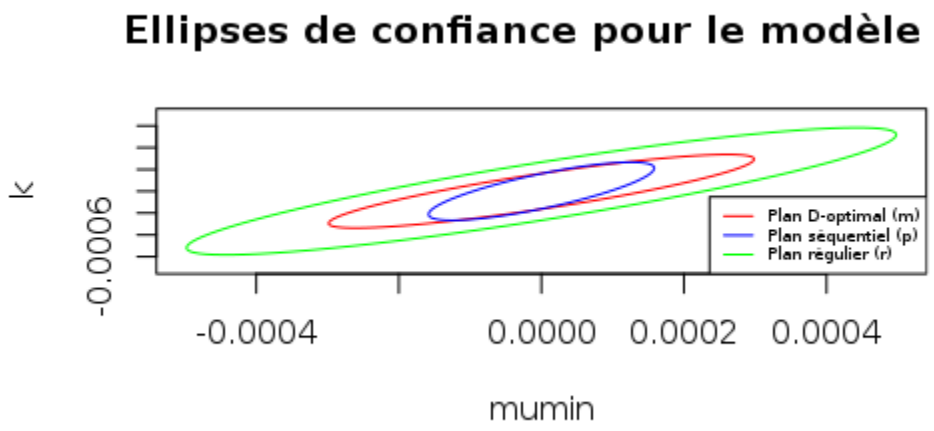

FIGURE 15 – Ellipses de confiance à 95% pour les estimateurs du modèle 3
La légende fait référence au plan n, p et r de la Table 5