

**INTERET DES MODELES DES PLUS PROCHES VOISINS
POUR LE CONTROLE DE L'HETEROGENEITE SPATIALE :
APPLICATION A UN ESSAI DE PROVENANCES DE PIN
D'ALEP (*PINUS HALEPENSIS* MILL.) EN TUNISIE**

**T. SGHAIER ⁽¹⁾, J-J CLAUSTRIAUX ⁽²⁾
et Med A. BEJI ⁽³⁾**

الملخص

لدراسة أهمية السّحائيل الفضائية المعروفة بنماذج الأجوّار المتقاربة في مراقبة تباين خصوبة التربة داخل التجارب الحقلية، تمّ استعمال ثلاث طرق تحاليل إحصائية لطول الأشجار المقاس بعد 29 عاما من النّمّو وذلك داخل تجربة لمقارنة 30 مصدرا محلي وأجنبي للصنوبر الحلبي بالبلاد التونسية. تتمثل هذه الطرق في تحليل طول الأشجار بنون أي مراقبة لتباين خصوبة التربة (تحليل عشوائي) وتحليل مع مراقبة إجمالية لتباين خصوبة التربة (تحليل حسب نموذج المقاسم العشوائية الكاملة) وتحليل مع مراقبة موضعية لتباين خصوبة التربة (تحليل حسب نماذج الأجوّار المتقاربة). بالنسبة للطريقة الأخيرة التي تعتمد لتقدير خصوبة كل موضع شجرة على خصوبة مواضع الأشجار المجاورة لها مباشرة، تمّت دراسة عدة تركيبات للأجوّار المتقاربة مع الأخذ بعين الاعتبار في كل مرة بمعامل أو بمعا ملي إرتداد. بينت النتائج المتحصل عليها، مقارنة بالتحليل العشوائي، بأن تقليص متوسط مربع الرواسب يقدّر بنحو 30% بالنسبة للتحتيل مع المراقبة الإجمالية لتباين خصوبة التربة، ويتراوح بين 66% و 79% بالنسبة للتحليل مع المراقبة الموضعية لتباين خصوبة التربة وذلك حسب النموذج المستعمل. لقد لوحظت علاقات خطية إيجابية ذات مغزى عالي جدا بين رواسب مختلف مواقع الأشجار داخل مثال التجربة ومتوسط رواسب مواقع الأشجار المجاورة وذلك قبل تعديل الطول المقاس حسب تباين الخصوبة الموضعية للتربة. بعد تعديل الطول المقاس بالإعتماد على نماذج الأجوّار

- 1) Unité de Gestion et de Valorisation des Ressources forestières, Institut National de Recherches en Génie Rural, Eaux et Forêts, BP N° 10, 2080 Ariana, Tunisie.
- 2) Unité de Statistique et Informatique, Faculté Universitaire des Sciences agronomiques de Gembloux, passage des Déportés, 8, 5030 Gembloux, Belgique.
- 3) Institut National Agronomique de Tunisie. Avenue Charles Nicolle, 43, 1082 - Tunis Mahrajène, Tunisie.

المتقاربة، تقلصت هذه العلاقات الخطية بنسبة تتراوح بين 92% و 95% من جهة أخرى، تبين من بين الستة نماذج للأجوار المتقاربة التي تمت دراستها، أن النموذج الأنجع لتقليص مربع الرواسب هو النموذج الذي يعتمد لتعديل الطول المقاس لكل شجرة على طول الأشجار المجاورة الثمانية مع معامل إرتداد. وتبين أيضا أن الأخذ بعين الإعتبار للمقاسم العشوائية الكاملة بعد تعديل الطول المقاس بإستعمال نموذج الأجوار المتقاربة ليس له أي تأثير على تقليص متوسط مربع الرواسب.

ABSTRACT

To study the importance of the spatial analysis in the control of soil heterogeneity in experiment on field, three approaches of statistical analysis were applied to a total height of trees measured at 29 years of age in a comparative test of thirty origins of Aleppo pine (*Pinus halepensis* Mill.), autochthonous and introduced in Tunisia. Analysis without control of heterogeneity (totally random), with a global control of heterogeneity (complete blocks) and with a local control of heterogeneity (nearest neighbour models) were used. For this last approach using to estimate the plot fertility the one of the nearby plots, various combinations of neighbour were tested with one or two coefficients of regression. The obtained results showed that the reduction of the residual mean square by comparison to the analysis without control of heterogeneity is 30% to the complete blocks analysis and varied between 66 and 79% for spatial analysis. Significant positive correlations between plots residues and the overage of the residues of the nearby plots were observed before adjustment of data. After adjustment of data by the nearest neighbour models, these correlations were registered a reduction from 92 to 95% according to the used model. Among the six nearest neighbour models studied, the model bringing in eight neighbours with two coefficients of regression showed itself the most effective. It also seemed that after adjustment of data by nearest neighbour models, using the structure of blocks does not produce a reduction of the residual mean square.

RESUME

Afin d'étudier l'importance des analyses spatiales dans le contrôle de l'hétérogénéité du terrain en expérimentation sur champ, trois approches d'analyses statistiques ont été appliquées à la hauteur totale des arbres mesurée à 29 ans d'âge dans un essai comparatif d'une trentaine de provenances de pin d'Alep (*Pinus halepensis* Mill.) autochtones et introduites en Tunisie. Il s'agit des analyses sans contrôle d'hétérogénéité (Complètement Aléatoire - CA), avec un contrôle global de l'hétérogénéité (Blocs Aléatoires Complets - BAC) et avec un contrôle local de l'hétérogénéité du terrain (Modèles des Plus Proches Voisins - MPPV). Pour cette dernière approche qui

consiste à utiliser pour l'estimation de la performance propre de chaque parcelle celles des parcelles voisines, différentes combinaisons des voisins les plus proches ont été testées avec chaque fois un ou deux coefficients de régression. Les résultats obtenus ont montré que la réduction de la variance résiduelle par rapport à l'analyse sans contrôle d'hétérogénéité est de l'ordre de 30% pour l'analyse en BAC et varie de 66 à 79% pour les MPPV.

Des corrélations positives très hautement significatives entre les résiduelles des différentes parcelles et la moyenne des résiduelles des parcelles voisines ont été observées avant ajustement des données à l'effet terrain. Après ajustement des données à l'aide des MPPV, ces corrélations ont enregistré une réduction de l'ordre de 92 à 95% selon le modèle utilisé. Parmi les six modèles des plus proches voisins étudiés, le modèle faisant intervenir les huit voisins avec deux coefficients de régression s'est montré le plus efficace. Il est apparu également qu'après ajustement des données à l'aide des MPPV, la prise en compte de la structure en blocs n'entraîne pas une réduction de la variance résiduelle.

INTRODUCTION

Les essais de provenances in situ ont pour objectif de dégager les meilleures d'entre elles dans des conditions de milieu bien définies. Dans ce but, il est nécessaire que le dispositif expérimental adopté et les techniques d'analyses statistiques qui en découlent permettent d'obtenir des résultats non biaisés et de variabilité minimum. Cette variabilité appelée erreur expérimentale est évaluée par la variance résiduelle. Souvent, en raison de l'hétérogénéité des conditions expérimentales, plus un essai est de grande taille, plus l'erreur expérimentale augmente et moins l'expérimentateur a de chance de mettre en évidence des différences faibles entre les provenances qu'il compare.

En particulier, les dispositifs en blocs aléatoires complets ne permettent en général qu'un contrôle limité de l'hétérogénéité du terrain. Le phénomène est accentué pour des essais de longue durée tels que ceux utilisés dans le domaine forestier, caractérisés le plus souvent par une hétérogénéité du terrain et une mortalité importante. Dans ce cas, il est certain que la performance des individus est influencée par leur position géographique relative. Pour ce genre de dispositifs, la méthode d'analyse statistique est d'autant plus efficace qu'elle rend précisément compte des effets de positionnement. Dès lors, d'autres techniques d'analyse statistique des résultats telles que les modèles des plus proches voisins (MPPV) pourraient être

intéressantes à considérer. C'est pourquoi les résultats issus d'un essai comparatif de provenances de pin d'Alep (*Pinus halepensis* Mill.) sont utilisés pour mesurer l'intérêt des dernières techniques citées par comparaison aux méthodes classiques, à savoir les analyses pour un dispositif sans contrôle de l'hétérogénéité (Complètement Aléatoire - CA) et celles avec un contrôle global d'hétérogénéité (Blocs Aléatoires Complets - BAC). Concernant les analyses avec contrôle local d'hétérogénéité (MPPV), différentes combinaisons des voisins les plus proches sont également testées pour estimer la performance du terrain en tout point du dispositif.

A. - INTERET DE LA METHODE DES PLUS PROCHES VOISINS

Afin de contrôler l'hétérogénéité du terrain qui entraîne des imprécisions dans l'interprétation des résultats, de nombreuses méthodes ont été développées. En 1935, Fisher a tout d'abord émis les premiers fondements d'une planification des expériences permettant de tenir compte de l'hétérogénéité du terrain (GOUARI, 1985); il a introduit les dispositifs expérimentaux classiques, à savoir les dispositifs en blocs complets, en blocs incomplets, en lattices, en carrés latins, etc. Les principes de répétition, de randomisation et du contrôle local constituent la base de la construction et de l'utilisation de ces dispositifs.

Pour augmenter la précision des résultats des expériences, PAPADAKIS a proposé en 1937 sa célèbre méthode d'analyse. Elle consiste à calculer pour chacune des parcelles un rendement ajusté ou corrigé par régression, en fonction notamment des rendements des parcelles voisines (DAGNELIE, 1987). Cette méthode, connue par la suite sous le nom de méthode des plus proches voisins a été reprise par différents auteurs. Afin d'affiner les estimations respectives des facteurs et des effets de terrain, BARTLETT a proposé en 1978 d'itérer la méthode de PAPADAKIS jusqu'à convergence. Selon certains auteurs, cette méthode permet dans certains cas de réduire la variance résiduelle jusqu'à 50% (AZAIS, 1990). En 1978, BARTLETT a signalé à son tour que le gain d'efficacité dans le cas d'une analyse de la variance d'un dispositif en blocs complets par la méthode des plus proches voisins peut être appréciable quand le nombre de traitements est grand et que la méthode est itérée.

Par ailleurs, en 1976, PEARCE a testé différentes combinaisons des plus proches voisins pour corriger les données observées (AASTVEIT 1983 ; PEARCE 1976). Il a montré que l'efficacité des tests augmentait d'une façon remarquable surtout lorsqu'il y avait un gradient irrégulier du terrain. En 1991, BAIRD et MEAD ont aussi comparé l'efficacité de deux modèles des plus proches voisins par rapport aux deux analyses classiques des dispositifs en blocs aléatoires complets et des dispositifs en blocs aléatoires incomplets sur des données simulées. Ils ont montré qu'en présence d'hétérogénéité, les analyses des modèles des plus proches voisins sont plus efficaces que les analyses classiques. La seule exception correspond au cas où cette hétérogénéité est éliminée totalement par la disposition des blocs. Ils ont encore montré que lorsque l'analyse en blocs incomplets présentait un gain par rapport aux blocs complets, les modèles des plus proches voisins offrent un gain proportionnellement plus important.

Enfin, également à partir des données simulées, PICHOT a observé en 1993 que, plus le gradient de terrain est élevé, plus les modèles des plus proches voisins contribuent à la diminution du carré moyen résiduel par rapport aux analyses classiques en blocs.

Si tous les auteurs s'accordent sur le fait que les techniques d'analyses spatiales contribuent d'une façon significative en présence d'hétérogénéité de terrain à la réduction du carré moyen résiduel et par conséquent à l'augmentation de l'efficacité des analyses, il reste néanmoins le problème des données manquantes et celui des parcelles de bordures. En effet, les parcelles de bordures comme celles ayant une ou plusieurs données manquantes à côté, ne disposent pas du nombre de parcelles voisines nécessaires à certains modèles. La technique adoptée dans ce cas consiste à prendre comme rendement de toute parcelle manquante le rendement moyen de celles qui existent et qui entourent la parcelle dont on désire estimer le rendement (GOUMARI, 1985). Quant au nombre de degrés de liberté à retrancher du facteur résiduel, BARTLETT (1938), dans KEMPTON et HOWES (1981), recommande pour obtenir une estimation conservatrice du carré moyen résiduel ajusté, de soustraire deux degrés de liberté dans le cas d'un modèle à une dimension avec un seul coefficient de régression. En 1976, PEARCE et MOORE suggèrent également d'enlever trois degrés de liberté dans le cas d'un modèle avec deux coefficients de régression. PEARCE (1998) propose quant à lui de retirer deux degrés de liberté par covariable. GOUMARI a proposé en 1990, en l'absence d'une formulation mathématique en la matière, de réduire le nombre de degrés de liberté d'autant d'unités qu'il y a de parcelles prises en considération dans le

modèle des plus proches voisins utilisé. Il reste à signaler que pour des grands essais tels que le nôtre, où le nombre de degrés de liberté du facteur résiduel est très élevé, l'effet de cette correction sera minimale voire négligeable.

B. MATERIEL ET METHODES

1. Matériel expérimental

Les données qui ont servi de base pour cette étude sont issues d'un essai de comparaison d'une trentaine de provenances de pin d'Alep, d'origine tunisienne et étrangères, installées au cours de la campagne de plantation 1964 (KHOUJA, 1997). Le dispositif expérimental adopté étant un dispositif en blocs aléatoires complets avec 33 répétitions. Les unités expérimentales sont des parcelles mono arbres plantées à une distance de 3 m de chaque côté. Des mesures de la hauteur totale des arbres à l'âge de 29 ans ont été effectuées au centimètre près sur l'ensemble des individus du dispositif.

Contrairement au diamètre qui semble dépendre, à côté de la fertilité du sol, d'autres facteurs tels que la densité de plantation, la compétition entre arbres ainsi que la conduite sylvicole, la hauteur des arbres d'un peuplement semble être beaucoup plus liée à la fertilité du sol. D'ailleurs, la hauteur dominante des peuplements a été toujours considérée par les forestiers comme une mesure de l'indice de fertilité (SGHAIER et al., 2001; SGHAIER et PALM, 2002).

2. Méthodes comparées

2.1. Absence de contrôle d'hétérogénéité

Une méthode d'analyse supposera l'implantation d'un dispositif complètement aléatoire (CA) qui ne permet aucun contrôle de l'hétérogénéité du terrain. Dans ce cas, la hauteur observée d'un arbre situé dans une parcelle de coordonnées l,c d'un champ d'essai de p lignes ($l = 1, \dots, p$) et q colonnes ($c = 1, \dots, q$) peut être représentée par :

$$Y_{i(l,c)} = m + a_i + X_{l,c} \quad (1)$$

avec :

$Y_{i(l,c)}$: valeur individuelle de la provenance i située au point de coordonnées l,c

m : moyenne générale

a_i : effet dû à la provenance i

$X_{l,c}$: résidus de moyenne nulle et de même écart type σ

2.2. Contrôle global de l'hétérogénéité

L'analyse de la variance pour le dispositif en blocs aléatoires complets (BAC) tel qu'il a été mis en place sera aussi réalisée. Le facteur bloc est un facteur subsidiaire, introduit dans le but de contrôler globalement l'hétérogénéité du terrain. Les différentes parcelles du dispositif sont groupées en r blocs ou répétitions complètes ($j = 1, \dots, r$). Le modèle théorique d'analyse de la variance s'écrit de la façon suivante :

$$Y_{ij} = m + a_i + B_j + X_{ij} \quad (2)$$

avec :

Y_{ij} : valeur observée d'un individu de la provenance i située dans le bloc j

m : moyenne générale

a_i : effet dû à la provenance i

B_j : effet dû au bloc j

X_{ij} : résidus de moyenne nulle et de même écart type σ

2.3. Contrôle local de l'hétérogénéité

Quand le terrain est homogène, les quantités $X_{l,c}$ du modèle 1 représentent en effet les résidus aléatoires, supposés indépendants, de distribution normale, de moyenne nulle et de même écart-type σ . Par contre, quand le terrain est hétérogène, les mêmes quantités $X_{l,c}$ renferment en plus des résidus considérés comme aléatoires, les erreurs dues aux variations locales du terrain. Ces variations constituent l'apport ou le gradient du terrain au point de coordonnées l,c . Le principe des modèles des plus proches voisins consiste donc à estimer cette contribution locale du terrain et à l'éliminer par la suite de la valeur observée.

Les modèles d'analyses spatiales testés dans cette étude sont tous basés sur le principe d'estimation itérative de la performance propre de chaque parcelle en fonction de celles des parcelles voisines (modèles des plus proches voisins - MPPV).

1° A partir du modèle (1), on procède au calcul du résidu propre de chaque parcelle de la façon suivante :

$$X_{l,c} = Y_{i(l,c)} - m - a_i \quad (3)$$

2° On calcule la quantité $(\hat{X}_{l,c})$ par estimation à l'aide des modèles des plus proches voisins. Cette quantité représente le gradient ou la performance propre du terrain estimé au point de coordonnées l,c .

3° Après avoir retranché de chaque valeur observée ($Y_{i(l,c)}$) le gradient parcellaire estimé correspondant ($\hat{X}_{l,c}$), l'analyse de la variance est réalisée selon le modèle :

$$Y'_{i(l,c)} = m + a_i + (X_{l,c} - \hat{X}_{l,c}) \quad (4)$$

Les quantités $(X_{l,c} - \hat{X}_{l,c})$ sont des résidus dans lesquels il ne subsiste plus aucune part à expliquer (GOUMARI, 1985). Ils doivent donc être aléatoires et de moyenne nulle.

Pour améliorer les résultats de l'ajustement, différents auteurs suggèrent d'itérer la méthode. En cas d'itération, les quantités $Y'_{i(l,c)}$ seront utilisées à la place de $Y_{i(l,c)}$ dans le cycle suivant, et ainsi de suite jusqu'à la n^{ième} itération (GOUMARI, 1985). L'élimination de la performance parcellaire ($\hat{X}_{l,c}$) estimée se fait donc d'une façon progressive à l'aide d'une procédure itérative.

La somme des quantités $(\hat{X}_{l,c})$ estimées par le modèle au cours des n itérations ($j = 1, \dots, n$), constitue le gradient ou l'effet total du terrain. Cette somme qui correspond également à la différence

entre les valeurs observées $Y_{i(l,c)}$ et les valeurs ajustées $Y'_{i(l,c)}$ obtenues à la $n^{\text{ième}}$ itération peut être désignée par $T_{l,c}$ telle que :

$$T_{l,c} = \sum_{j=1}^n \left(\hat{X}_{l,c} \right)_j = Y_{i(l,c)} - Y'_{i(l,c)}$$

Trois types des voisins les plus proches ont été testés dans cette étude. Le tableau 1 présente les positions et les coordonnées des voisins utilisés dans les différents modèles testés.

Tableau 1 - Positions et coordonnées des voisins utilisés dans les modèles testés.

Types	Positions des voisins	Coordonnées	Appellation																		
1	<table style="border-collapse: collapse; margin: auto;"> <tr><td></td><td>X</td><td></td></tr> <tr><td>X</td><td>Y</td><td>X</td></tr> <tr><td></td><td>X</td><td></td></tr> </table>		X		X	Y	X		X		<table style="border-collapse: collapse; margin: auto;"> <tr><td></td><td>$l-1,c$</td><td></td></tr> <tr><td>$l,c-1$</td><td>l,c</td><td>$l,c+1$</td></tr> <tr><td></td><td>$l+1,c$</td><td></td></tr> </table>		$l-1,c$		$l,c-1$	l,c	$l,c+1$		$l+1,c$		Voisins
	X																				
X	Y	X																			
	X																				
	$l-1,c$																				
$l,c-1$	l,c	$l,c+1$																			
	$l+1,c$																				
2	<table style="border-collapse: collapse; margin: auto;"> <tr><td>X</td><td></td><td>X</td></tr> <tr><td></td><td>Y</td><td></td></tr> <tr><td>X</td><td></td><td>X</td></tr> </table>	X		X		Y		X		X	<table style="border-collapse: collapse; margin: auto;"> <tr><td>$l-1,c-1$</td><td></td><td>$l-1,c+1$</td></tr> <tr><td></td><td>l,c</td><td></td></tr> <tr><td>$l+1,c-1$</td><td></td><td>$l+1,c+1$</td></tr> </table>	$l-1,c-1$		$l-1,c+1$		l,c		$l+1,c-1$		$l+1,c+1$	Coins
X		X																			
	Y																				
X		X																			
$l-1,c-1$		$l-1,c+1$																			
	l,c																				
$l+1,c-1$		$l+1,c+1$																			
3	<table style="border-collapse: collapse; margin: auto;"> <tr><td>X</td><td>X</td><td>X</td></tr> <tr><td>X</td><td>Y</td><td>X</td></tr> <tr><td>X</td><td>X</td><td>X</td></tr> </table>	X	X	X	X	Y	X	X	X	X	<table style="border-collapse: collapse; margin: auto;"> <tr><td>$l-1,c-1$</td><td>$l-1,c$</td><td>$l-1,c+1$</td></tr> <tr><td>$l,c-1$</td><td>l,c</td><td>$l,c+1$</td></tr> <tr><td>$l+1,c-1$</td><td>$l+1,c$</td><td>$l+1,c+1$</td></tr> </table>	$l-1,c-1$	$l-1,c$	$l-1,c+1$	$l,c-1$	l,c	$l,c+1$	$l+1,c-1$	$l+1,c$	$l+1,c+1$	Voisins et Coins
X	X	X																			
X	Y	X																			
X	X	X																			
$l-1,c-1$	$l-1,c$	$l-1,c+1$																			
$l,c-1$	l,c	$l,c+1$																			
$l+1,c-1$	$l+1,c$	$l+1,c+1$																			

Y : Parcelle dont la performance est à estimer.

X : Parcelle voisine à prendre en compte pour l'estimation.

Pour chaque type de voisins, deux modèles ont été construits, le premier avec un seul coefficient de régression et le deuxième avec deux coefficients. On obtient ainsi six modèles des plus proches voisins dont la formulation mathématique est comme suit :

1. $X_{l,c} = b\left(\frac{1}{4}X_{l-1,c} + \frac{1}{4}X_{l+1,c} + \frac{1}{4}X_{l,c-1} + \frac{1}{4}X_{l,c+1}\right) + Z_{l,c}$
2. $X_{l,c} = b_1\left(\frac{1}{2}X_{l-1,c} + \frac{1}{2}X_{l+1,c}\right) + b_2\left(\frac{1}{2}X_{l,c-1} + \frac{1}{2}X_{l,c+1}\right) + Z_{l,c}$
3. $X_{l,c} = b\left(\frac{1}{4}X_{l-1,c-1} + \frac{1}{4}X_{l+1,c+1} + \frac{1}{4}X_{l+1,c-1} + \frac{1}{4}X_{l-1,c+1}\right) + Z_{l,c}$
4. $X_{l,c} = b_1\left(\frac{1}{2}X_{l-1,c-1} + \frac{1}{2}X_{l+1,c+1}\right) + b_2\left(\frac{1}{2}X_{l+1,c-1} + \frac{1}{2}X_{l-1,c+1}\right) + Z_{l,c}$
5. $X_{l,c} = b\left(\frac{1}{8}X_{l-1,c} + \frac{1}{8}X_{l+1,c} + \frac{1}{8}X_{l,c-1} + \frac{1}{8}X_{l,c+1} + \frac{1}{8}X_{l-1,c-1} + \frac{1}{8}X_{l+1,c+1} + \frac{1}{8}X_{l+1,c-1} + \frac{1}{8}X_{l-1,c+1}\right) + Z_{l,c}$
6. $X_{l,c} = b_1\left(\frac{1}{4}X_{l-1,c} + \frac{1}{4}X_{l+1,c} + \frac{1}{4}X_{l,c-1} + \frac{1}{4}X_{l,c+1}\right) + b_2\left(\frac{1}{4}X_{l-1,c-1} + \frac{1}{4}X_{l+1,c+1} + \frac{1}{4}X_{l+1,c-1} + \frac{1}{4}X_{l-1,c+1}\right) + Z_{l,c}$

L'application de l'ensemble de ces modèles est assurée par un programme informatique mis au point spécialement pour cette étude. En cas de donnée manquante, la valeur de la parcelle concernée est remplacée par la moyenne des autres parcelles voisines.

Les six modèles des plus proches voisins définis ci-dessus sont appliqués sur les mesures de la hauteur pour estimer la performance de chaque parcelle en fonction de celles de ses voisines grâce à un processus d'estimation itérative. Une analyse de la variance est par la suite réalisée sur les hauteurs ajustées à l'effet terrain.

2.4. Conditions d'application et critères de comparaison

Pour les différentes analyses de la variance, la normalité des résidus est vérifiée par le calcul des coefficients de dissymétrie G_1 de FISHER et d'aplatissement β_2 de PEARSON.

Pour l'analyse de la variance à deux critères de classification (BAC), le modèle linéaire correspondant (3) est par construction additif. Cela signifie que la valeur observée est le résultat de l'addition d'un effet provenance et d'un effet bloc, avec en plus, un effet moyen du champ (GOUET et PHILIPPEAU, 1997). Il faut donc vérifier, en plus de la normalité des résidus, l'absence d'interaction entre les provenances et les blocs, c'est-à-dire que les différences entre les provenances sont les mêmes dans tous les blocs. Comme le dispositif adopté est mono arbre, c'est-à-dire qu'on ne dispose pas d'une vraie résiduelle, le test de signification de l'interaction entre les provenances et les blocs est réalisé grâce au test d'additivité de TUKEY (DAGNELIE, 1998).

Les comparaisons des trois approches d'analyse effectuées et des six modèles des plus proches voisins testés se font principalement sur la base de :

- l'évolution de l'écart-type résiduel (ETR),
- l'évolution de la valeur calculée de la variable F de SNEDECOR correspondant au facteur «provenance» (F_{prov}),
- l'évolution des deux indices de normalité des résidus, à savoir le paramètre de dissymétrie de FISHER (G_1) et le paramètre d'aplatissement de PEARSON (β_2).

3. RESULTATS

Afin de déterminer le nombre nécessaire d'itérations à effectuer, nous avons suivi pour l'ensemble des six modèles testés, l'évolution de l'écart-type résiduel et de la valeur observée de la variable F de Snedecor, résultants de l'analyse de la variance à un critère de classification de la hauteur totale des arbres. Le tableau 2 présente les résultats de cette évolution en fonction du nombre d'itérations.

Tableau 2 - Evolution en fonction du nombre d'itérations, de l'écart-type résiduel et de la valeur F de SNEDECOR obtenus à l'aide des six modèles testés.

Itér.	M1		M2		M3		M4		M5		M6	
	ETR	F_{prov}	ETR	F_{prov}	ETR	F_{prov}	ETR	F_{prov}	ETR	F_{prov}	ETR	F_{prov}
0 ⁽¹⁾	119,98	5,96	119,98	5,96	119,98	5,96	119,98	5,96	119,98	5,96	119,98	5,96
1	93,76	9,74	93,56	9,64	98,65	8,09	98,27	8,10	91,88	9,59	91,87	9,76
2	80,45	12,83	80,47	13,01	87,14	11,45	86,55	11,58	81,29	13,08	79,90	13,44
3	73,61	15,41	74,00	15,23	80,90	12,50	80,19	12,61	74,98	14,78	72,12	16,03
4	69,48	16,93	70,16	17,03	77,11	14,74	76,34	14,96	71,63	16,90	67,71	18,71
5	66,83	18,43	67,84	18,06	74,74	15,01	73,95	15,19	69,28	17,55	64,43	20,27
10	61,71	21,11	63,87	20,48	70,66	17,55	69,78	17,87	64,74	20,69	57,05	26,27
15	60,67	21,87	63,23	20,83	70,05	17,67	69,07	18,04	63,83	21,11	54,81	28,39
20	60,43	21,97	63,09	20,97	69,95	17,82	68,91	18,23	63,63	21,34	54,14	29,23
50	60,36	22,03	63,02	21,01	69,93	17,82	68,83	18,25	63,58	21,35	53,61	29,71

(1) Itération 0 : analyse de la variance réalisée sur les données brutes.

En regardant l'évolution des deux paramètres suivis (tableau 2), nous constatons que la réduction de l'écart-type résiduel et l'augmentation de la valeur F de SNEDECOR se stabilisent vers la 15^{ème} itération et ceci pour l'ensemble des six modèles testés. Aussi, nous avons considéré dans cette étude 15 itérations.

A l'itération 0, les coefficients de corrélation de Pearson entre les résidus des différentes parcelles et la moyenne des résidus des parcelles voisines pour les trois types de voisins pris en considération dans cette étude sont de 0,621, de 0,565 et de 0,640 pour respectivement les modèles M1 (voisins), M3 (coins) et M5 (voisins et

coins). A la 15^{ème} itération, les mêmes corrélations deviennent 0,049, 0,031 et 0,050. La figure 1 présente l'évolution en valeur absolue du coefficient de corrélation en fonction du nombre d'itérations pour le modèle 5. Cette figure montre comment la dépendance entre les résidus des parcelles voisines diminue d'une itération à une autre et ceci par l'élimination progressive de l'effet terrain en tout point du dispositif.

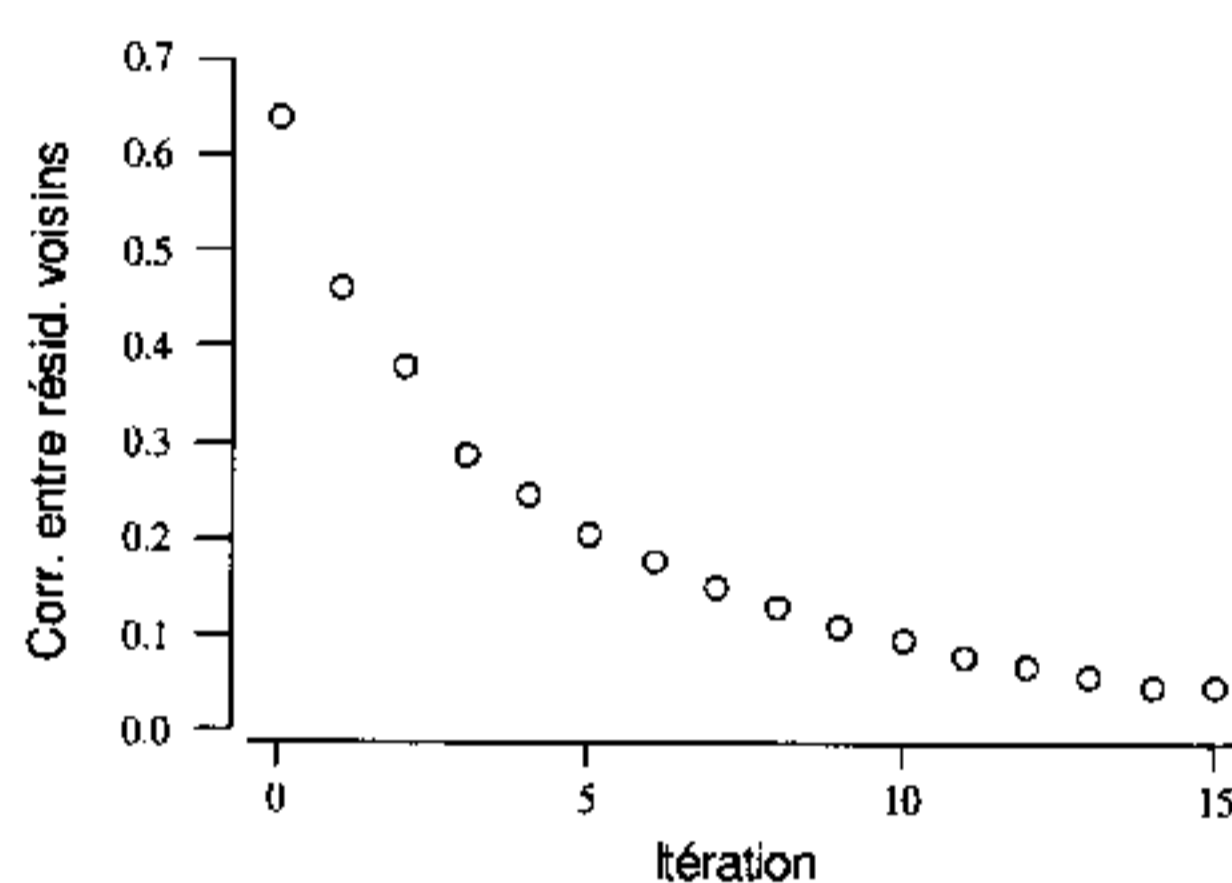


Figure 1 - Évolution en valeur absolue du coefficient de corrélation en fonction du nombre d'itérations entre les résidus des différentes parcelles et la moyenne des résidus des parcelles voisines pour le modèle 5 (voisins et coins).

Le tableau 3 présente les résultats relatifs aux différentes analyses réalisées sur les données des mesures de la hauteur des arbres. Il s'agit des valeurs des coefficients de dissymétrie G_1 de FISHER et d'aplatissement β_2 de PEARSON relatives aux résidus, de l'écart type résiduel, du coefficient de variation résiduel en pourcentage, de la valeur F de SNEDECOR relative au facteur provenance et/ou bloc selon le cas et enfin, de l'efficacité relative des différentes analyses par rapport à l'analyse sans contrôle d'hétérogénéité (rapport des carrés moyens résiduels). Ces analyses ont été effectuées d'une part sans contrôle d'hétérogénéité (dispositif complètement aléatoire - CA) et, d'autre part avec un contrôle global d'hétérogénéité (dispositif en blocs aléatoires complets - BAC) et avec un contrôle local de l'hétérogénéité basé sur l'analyse des données ajustées à l'aide des MPPV.

Pour vérifier si la structure en blocs doit être prise en considération après ajustement des données à l'aide des MPPV, nous avons réalisé également dans le cas du contrôle local de l'hétérogénéité l'analyse de la variance à deux critères de classification.

Tableau 3 – Résultats des analyses statistiques sans contrôle et avec contrôle global et local de l'hétérogénéité du terrain.

Type d'analyse	G ₁	β ₂	ETR	F _{prov.}	F _{bloc}	ER% (/CA)
- Sans contrôle (CA)	0,192	2,798	119,982	5,96***	-	-
- Contrôle global (BAC)	0,011	3,090	100,292	8,55***	13,67***	143
- Contrôle local (MPPV)						
CA						
M1	-0,115	3,179	60,665	21,87***	-	391
M2	-0,096	3,351	63,229	20,83***	-	360
M3	-0,004	3,087	70,049	17,67***	-	293
M4	0,010	3,118	69,071	18,04***	-	302
M5	-0,152	3,272	63,826	21,11***	-	353
M6	-0,156	3,218	54,811	28,37***	-	479
BAC						
M1	-0,116	3,172	61,436	21,37***	0,27 ^{NS}	381
M2	-0,111	3,355	63,914	20,45***	0,37 ^{NS}	352
M3	-0,027	3,163	70,383	17,55***	0,72 ^{NS}	291
M4	-0,013	3,175	69,483	17,87***	0,65 ^{NS}	298
M5	-0,144	3,269	64,632	20,63***	0,27 ^{NS}	346
M6	-0,155	3,211	55,507	27,72***	0,27 ^{NS}	467

NB : *** effet significatif au seuil $\alpha = 0,001$.

** effet significatif au seuil $\alpha = 0,01$

^{NS} effet non significatif.

Afin de tester l'additivité du modèle en blocs aléatoires complets (BAC) sur les données brutes, nous avons eu recours au test d'additivité de TUKEY. Pour réaliser ce test qui s'applique uniquement

aux données équilibrées, nous avons éliminé du dispositif une dizaine de provenances qui ont enregistré au total une vingtaine d'arbres morts. La valeur du test F_{add} de TUKEY est égale à 4,60 ce qui donne avec 1 et 607 degrés de liberté une valeur $P < 0,032$. Ce résultat indique que le modèle d'analyse de la variance à deux critères utilisé n'est pas additif, c'est-à-dire la présence d'une interaction provenance x bloc. Cette interaction significative indique que l'essai manifeste un comportement particulier :

- ou bien les blocs sont tellement différents entre eux que les écarts entre les provenances ne sont pas constants : certaines provenances s'extériorisent beaucoup mieux dans les bonnes conditions et moins bien en mauvaises conditions ou le contraire,

- ou bien les blocs sont mal disposés sur le terrain et les écarts entre les provenances sont influencés par la répartition de ces provenances à l'intérieur de chaque bloc.

En ce qui concerne la normalité des résidus, l'effectif total des arbres ayant fait l'objet des mesures est de 970 individus. En se référant aux tables de PEARSON et HARTLEY (1966), nous constatons que, pour un champ de 1.000 parcelles, des valeurs de G_1 et β_2 comprises dans les intervalles $- 0,18$ à $+ 0,18$ et $2,68$ à $3,41$ permettent de supposer que les observations se distribuent selon une loi normale. Le tableau 3 montre que les résidus issus des analyses de la variance des six modèles des plus proches voisins utilisés se distribuent selon une loi normale.

Le même tableau montre également que l'efficacité relative d'une analyse avec contrôle global de l'hétérogénéité par rapport à une analyse sans contrôle de l'hétérogénéité est de l'ordre 143%. Celles des analyses avec contrôle local de l'hétérogénéité à un et à deux critères varient de 290 à 480% environ. Il apparaît également du même tableau que la prise en compte de la structure en blocs après ajustement des données à l'aide des MPPV n'entraîne pas une augmentation de l'efficacité relative. En supposant vérifiée l'additivité du modèle d'analyse de la variance à deux critères, le test de l'effet bloc qui était très hautement significatif dans le cas de l'analyse de la variance sur les données brutes avec une valeur F de Snedecor de 13,67 devient non significatif avec des valeurs inférieures à l'unité pour les mêmes analyses sur les données ajustées et ceci pour les six modèles des plus proches voisins testés. Ces chiffres permettent donc de conclure qu'après ajustement à l'aide des MPPV il n'est plus justifié de garder la structure en blocs.

En ce qui concerne les modèles des plus proches voisins étudiés, les résultats obtenus (tableaux 2 et 3) montrent une nette supériorité du modèle 6 et ceci pour l'ensemble des critères de comparaison retenus. Il s'agit du modèle qui prend en considération pour l'estimation de la performance propre de chaque parcelle celles des 8 parcelles voisines avec deux coefficients de régression, un pour les voisins et un pour les coins.

4. DISCUSSION

Les différents modèles des plus proches voisins utilisés pour l'analyse des données relatives à notre dispositif ont permis de réduire le carré moyen résiduel de 66 à 79%. Ces modèles qui ont été utilisés pour déterminer la performance des parcelles en fonction de celle de leurs voisines diffèrent par le nombre et la position des voisins. Les résultats relatifs à l'efficacité des MPPV par comparaison aux analyses classiques, obtenus dans cette étude, dépassent ceux obtenus par AZAIS *et al.* (1990) qui affirment que la méthode des plus proches voisins permet dans certains cas de réduire la variance résiduelle jusqu'à 50%. Le maximum de réduction du carré moyen résiduel étant obtenu à l'aide du modèle 6 (voisins et coins) qui fait intervenir les 8 voisins avec deux coefficients de régression pour la détermination de la performance du terrain en tout point du dispositif. Ce dernier modèle a permis également d'améliorer la normalité de la distribution des résidus d'une façon très remarquable par comparaison aux analyses classiques.

Les résultats obtenus dans le cadre de cette étude montrent également qu'il n'est plus justifié de prendre en considération la structure en blocs après ajustement des données à l'aide des modèles des plus proches voisins et ceci pour l'ensemble des modèles testés.

De plus, l'utilisation des modèles des plus proches voisins permet d'estimer le gradient du terrain et d'établir par conséquent la carte de la fertilité du dispositif expérimental. En se basant sur la valeur du terrain estimée en tout point du dispositif et pour autant que la plage d'hétérogénéité soit suffisamment étendue, l'essai pourrait être divisé en sous-essais et le comportement des provenances soumises à des conditions locales différentes pourrait être évalué. Le recours à de telles techniques d'analyses constitue donc une solution fort intéressante pour résoudre le problème d'un grand nombre d'essais, installés depuis des décennies sur des terrains forestiers le plus souvent accidentés, et qui sont conçus selon des dispositifs qui ne permettent en général qu'un contrôle fort limité de l'hétérogénéité du terrain.

BIBLIOGRAPHIE

- AASTVEIT H., 1983 - On the effect of correlation between plots in randomized block experiments. *Biom. J.*, 25, No. 2, 129-153.
- AZAIS J. M., DENIS J. B., DHORNE T., KOBILINSKY A., 1990 - Neighbour analysis of experiments : a review of the different approaches. *Biom. Parxim.*, 30, 15-39.
- BAIRD D., MEAD R., 1991 - The empirical efficiency and validity of two neighbour models. *Biometrics*, 47, 1473-1487.
- BARTLETT M. S., 1978 - Nearest neighbour models in the analysis of Field experiments. *J. R. Statist. Scoc. B*, 40, No. 2, 147-174.
- DAGNELIE P., 1987 - La méthode de Papadakis en expérimentation agronomique : considérations historiques et bibliographiques. *Biom. Praxi.*, 27, 49-64.
- DAGNELIE P., 1998 - Statistique théorique et appliquée. Tome 2 : inférence statistique à une et à deux dimensions. Bruxelles, De Boeck, 660 P.
- GOUET J. P., PHILIPPEAU G., 1997 - Analyse de la variance. Comment interpréter les résultats ? Collection Stat – ITCF, 47 p.
- GOUMARI A., 1985 - Contribution à l'étude et à la maîtrise de l'hétérogénéité des expériences en champ. Thèse de doctorat. Faculté Universitaire des Sciences agronomiques de Gembloux., 150 p.
- GOUMARI A., 1990 - Analyse comparative des résultats d'essais en champ selon les techniques des blocs aléatoires complets, des lattices et des plus proches voisins. *Biom. Praxim.*, 30, 91-105.
- KEMPTON R. A., HOWES C. W., 1981 - The use of neighbour ing plot values in the analysis of variety trials. *Appl. Statist.*, 30, No. 1, 59-70.
- KHOUJA M. L., 1997 - Variabilité géographique du pin d'Alep en Tunisie. Perspective d'amélioration de la productivité et de la qualité physique du bois. Thèse de doctorat. Université catholique de Louvain. Faculté des Sciences agronomiques. 181 p.
- PAPADAKIS J. S., 1937 - Méthode statistique pour les expériences sur champ. Institut d'amélioration des plantes à Salonique (Grèce). Bulletin scientifique, no. 23.

- PEARCE S. C., 1976 - An examination of Fairfield Smith's law of environmental variation. *J. Agric. Sci.*, 87, 21-24.
- PEARCE S. C., 1998 - Field experimentation on rough land : the method of Papadakis reconsidered. *J. Agric. Sci.*, 131, 1-11.
- PEARCE S. C., MOORE C. S., 1976 - Reduction of experimental error in perennial crops, using adjustment by neighbouring plots. *Expl. Agric.*, 12, 267-272.
- PEARSON E. S., HARTLEY H. O., 1966 - *Biometrika tables for statisticians* (Vol. I). University Press, Cambridge, 264 p.
- PICHOT C., 1993 - Analyse de dispositifs par approches itératives prenant en compte les performances des plus proches voisins. *Agronomie*, 13, 109-119.
- SGHAIER T. et PALM R., 2002 - Répartition des arbres et des volumes par classes de grosseur dans les peuplements de pin d'Alep (*Pinus halepensis* Mill.) en Tunisie. *Ann. For. Sci.* 59, 293-300.
- SGHAIER T., AMMARI Y., GARCHI S. et KHALDI A., 2001 - Croissance en hauteur dominante et classes de fertilité du pin d'Alep. *Les annales de l'INRGREF*, numéro spécial, 45-53.