

## Distinction des blés durs algériens (*Triticum durum* Desf. ) par l'utilisation de la méthode de la distance de Mahalanobis

Dekhili M., Aggoun A.

Institut de Biologie, Laboratoire de Génétique et Statistique, Université FerhatT Abbes de Sétif, 19000, Algérie.

### Résumé

*La distance généralisée de Mahalanobis  $D^2$  est utilisée comme critère de distinction entre cinq populations de blé dur décrites par quatre variables quantitatives. L'expérimentation a été menée sur deux années (1993-94 et 1994-95) dans l'exploitation agricole Dehal Nouari située au nord de Sétif (Algérie). Les résultats (de cette analyse) indiquent que parmi les quatre variables étudiées, la longueur du grain et de l'épi constituent les variables les plus importantes ou les plus discriminantes. Les populations Oued-Zenati 368 et Bidi 17, constituent le seul couple non distinct ou similaire. Cette ressemblance s'explique davantage par leur généalogie commune. Les autres populations sont, au contraire, distinctes l'une de l'autre.*

**Mots clés :** Distance généralisée, distance critique, similarité, distinction, profil de distance de Mahalanobis, populations algériennes de blé dur

### Abstract : Distinctness among Algerian durum wheat using the Generalised Mahalanobis Distance

*The multivariate Mahalanobis Generalised  $D^2$  statistic is being used to distinguish between five algerian durum wheat populations and four measured characters. Trials were carried out for two years at Dehal Nouari farm, Sétif (Algeria). The results indicated that grain length and ear length were the most important characters in distinctness testing. Oued-Zenati 368 and Bidi 17 constitute a non distinct pair. Their complete similarity is well explained by their*

*common genealogy. On the other side, Mohamed Ben Bachir, Hedba 3 and Ziban are well distinct from each other.*

**Keys words :** Generalized distance, critical value, similarity, distinctness, profiles of distances, algerian durum wheat populations

**ملخص :** تفريق القمح الصلب الجزائري عن طريقة إستعمال مسافة Mahalanobis

دخيلي م. و عقون ع.

استعملت مسافة Mahalanobis للتفريق بين خمس مجموعات من القمح الصلب و ذلك من خلال دراسة أربعة صفات كمية. أجريت الدراسة على مدى سنتين (94/93 و 95/94) بمزرعة دحال النواري المتواجدة بشمال ولاية سطيف (الجزائري).

أظهرت نتائج هذا التحليل أن من بين الصفات الأربعة المدروسة كان طول البذرة و طول السنبله من الصفات الأكثر تميزا بين المجموعات.

تكون المجموعتان Bidi17 و Oued Zenati 368 ثنائيا متشابهة تماما، و يمكن تفسير هذا التشابه من خلال انحدارهما من نفس الأصل الوراثي، بينما المجموعات الأخرى كانت غير متشابهة مع بعضها البعض.

**الكلمات المفتاحية :** المسافة المعممة، المسافة الحرجة، التشابه، التفريق، مظهر مسافة Mahalanobis. المجموعات الجزائرية للقمح الصلب

## Introduction

Depuis que les blés durs cultivés en Algérie ont été inventoriés (Erroux, 1958 ; Laumont et Erroux, 1961), les populations locales de blé ont été délaissées par les organismes spécialisés et les agriculteurs au profit de variétés introduites massivement, causant ainsi une régression significative de la grande diversité antérieure (Erroux, 1958). De nouvelles études descriptives de l'état actuel de ces populations, devraient permettre d'évaluer les diversifications qu'elles ont subi, de dégager un protocole de contrôle des semences, d'élaborer une protection variétale et des travaux d'amélioration génétique plus effectifs.

A cet égard, l'Union internationale pour la Protection des Obtentions Végétales (UPOV), a élaboré des principes directeurs pour caractériser la Distinction, l'Homogénéité et la Stabilité (UPOV, 1988).

La présente étude menée dans la région céréalière de Sétif (Algérie) vise à identifier (1) les caractères les plus distinctifs et (2) les populations les plus similaires. A cet effet, la distance généralisée de Mahalanobis (Mahalanobis, 1936) a été utilisée comme critère de distinction.

## Matériel et méthodes

Le matériel végétal étudié a été fourni par la station de l'Institut Technique des Grandes Cultures (ITGC) de Sétif. Il se compose de cinq populations locales de blé dur (*Triticum durum* Desf.) de génération zéro (G0). Les populations étudiées sont : Oued Zenati 368 (OZ), Hedba 3 (HD), Mohamed Ben Bachir (MBB), Bidi 17 (BD) et Ziban (ZB).

L'essai a été mené durant l'année 1993-94 et l'année 1994-95, dans l'exploitation Dehal Nouari, localisée au Nord de Sétif (altitude de 1200 m, avec un sol pauvre en calcaire, de texture argileuse de couleur noire).

Le dispositif expérimental préconisé par l'UPOV (1988) a été utilisé, le précédent cultural est une jachère travaillée. Lors de l'expérimentation, le désherbage a été réalisé manuellement, avec deux apports d'azote en début du tallage et en début de la montaison.

Les variables prises en considération sont respectivement : la hauteur de la plante (P en cm), la longueur de l'épi (LE en cm), la longueur du grain (LGR en mm) et le poids des grains par épi (PGE en gr). Les notations des variables ont été réalisées sur le terrain pour la hauteur de la plante, et en laboratoire pour les autres variables.

Afin de décider de la distinction entre deux populations, la distance généralisée de Mahalanobis (1936) a été utilisée. Cette distance généralisée  $D^2$  constitue un critère de proximité qui tient compte de la dispersion (matrice de variance-covariance) entre les éléments. L'avantage de ce critère par rapport à la distance euclidienne dans la représentation spatiale des éléments, est mis en évidence par Dagnelie et Merckx (1991).

La mesure de la distance  $D^2$  est définie par :

$$D^2 = \mathbf{d}' \mathbf{R}^{-1} \mathbf{d}$$

où  $\mathbf{d}$  est le vecteur formé des différences entre les moyennes standardisées des variables sur le couple de variétés soumis à la comparaison ; d'où si les moyennes d'une population pour  $p$  variables sont  $(x_1, x_2, \dots, x_p)$  et les moyennes d'une autre population pour la même variable sont  $(y_1, y_2, \dots, y_p)$ , alors :

$$\mathbf{d} = \begin{bmatrix} x_1 - y_1 / s_1 \\ x_2 - y_2 / s_2 \\ " \\ " \\ " \\ x_p - y_p / s_p \end{bmatrix}$$

$\mathbf{d}'$  est le vecteur transposé de  $\mathbf{d}$ ,  $\mathbf{R}^{-1}$  désigne l'inverse de la matrice des coefficients de corrélation entre les variables.

La méthode de calcul de la distance  $D^2$  entre deux variétés, nécessite au préalable les analyses de variance et covariance pour toute une série d'années et de variétés (Weatherup, 1994). Pour les variables  $i$  et  $j$ , la présente table des carrés moyens et de leurs produits est obtenue pour  $m$  années et  $n$  variétés :

Source de variation	d.d.l	CMii	PCMij	CMjj
Années	$m - 1$	Aii	Aij	Ajj
Variétés	$n - 1$	Vii	Vij	Vjj
Résiduelle	$(m - 1)(n - 1)$	Rii	Rij	Rjj

Les erreurs standard  $s_i$  et  $s_j$  intra années et variétés pour les variables  $i$  et  $j$  respectivement, sont données par :

$$s_i = \sqrt{R_{ii}/(m-1)(n-1)}$$

Les valeurs des distances observées doivent être comparées à d'autres valeurs de  $D^2$ , dites valeurs critiques. Ces dernières sont nécessaires pour la distinction ; elles peuvent être généralisée à partir de la statistique  $T^2$  donnée par Hotelling (1931) :

$$D^2 \text{ critique} = \frac{2 p (m-1)(n-1)}{m (mn-m-n-p+2)} F$$

où  $F$  est le ratio de Fisher évalué au seuil de 1 % pour les degrés de libertés  $p$  ( nombre de variables) et  $mn-m-n-p+2$ .

## Résultats et discussion

Les résultats sont présentés sous forme de profils (tableau 1). Pour chacun d'eux, sont données les distances observées successivement pour le premier trait (le plus important); et celles qui concernent deux, trois et quatre variables. Un second profil de distances critiques est également fourni.

Pour être significative et pouvoir ainsi déclarer deux populations comme étant distinctes pour un nombre déterminé de variables, la distance critique doit être supérieure à la distance observée; dans le cas contraire, le couple est déclaré non distinct ou similaire.

### Couple n° 5

Ce couple est constitué de Oued-Zenati 368 et de Bidi 17. Selon les distances présentées (tableau 1), Oued-Zenati 368 et de Bidi 17 sont non distinctes dès l'introduction de la première variable (longueur de l'épi, LE). Les distances  $D^2$  observées restent inférieures aux valeurs de la  $D^2$  critique de la première à la dernière variable. La forte ressemblance de ces deux populations s'explique par leur généalogie commune (Laumont et Erroux, 1961). En l'absence de

programmes de sélection, elles ont réussi à maintenir la même constitution génétique à travers les années, résultat qui a été aussi confirmé par Dekhili et al (1997) selon les données ordinales.

### Couples n° 1, 2, 7 et 9

Ces couples sont constitués de Ziban - Oued Zenati 368, Ziban - Bidi 17, de Oued Zenati 368 - Mohamed Ben Bachir et de Mohamed Ben Bachir - Bidi 17. Selon les résultats de ces profils (tableau 1), on remarque une conformité dans la hiérarchisation des caractères discriminants selon leur importance. La première variable la plus discriminante entre ces différents couples, est la longueur du grain (LGR), la seconde variable importante est la longueur de l'épi (LE). L'observation des profils pour ces couples (tableau 1), indique que la distance observée augmente dès l'introduction du premier (LGR) et second (LE) caractère, puis diminue au troisième (PGE) et quatrième (HP) trait, pour enfin devenir non significative ( $P < 0.05$ ). La différence initiale observée s'est diluée avec l'addition de différences de moindre importance. Dans ce cas, la variable la moins discriminante ou la moins importante est représentée par la hauteur de la plante (tableau 2).

### Couples n° 3 et 6

Ces deux couples sont constitués successivement de Ziban - Hedba 3 et de Oued Zenati 368 - Hedba 3 (tableau 1). Selon les profils de ces couples, on observe également une hiérarchisation des caractères d'après leur degré de discrimination à savoir en premier la longueur de l'épi, en second la hauteur de la plante et en dernier la longueur du grain. Pour ces couples, la variable PGE ne contribue nullement à la discrimination ( $P < 0.05$ ).

### Couples 4, 8 et 10

Ces couples forment un groupe atypique (tableau 1). En effet, l'ordre d'importance des variables discriminantes varie d'un couple à un autre (tableau 1). La longueur de l'épi constitue la variable la plus importante pour les couples 8 et 10, pour être reléguée en seconde position dans le couple 4.

La contribution de chaque variable à la  $D^2$  (excepté pour le premier important trait), dépend de l'observation de larges et négatives corrélations (intra-années et variétés) d'une variable avec les autres caractères (Cochran, 1962). Selon les résultats du tableau 1, les couples 1, 2, 7, 9 et 10 ont leur second important trait (LE pour 1, 2, 7 et 9 ; LGR pour 10) négativement corrélé avec le premier caractère (Tableau 3), avec une valeur de - 0.90. La même remarque peut être faite pour les couples 3, 6, 4 et 8, où l'on observe une corrélation négative entre le premier (LE) et le second (HP) caractère. Ceci explique la forte association selon les profils entre LGR - LE et LE - HP. L'observation de la corrélation positive de PG avec les autres variables est à l'origine de sa faible participation à la  $D^2$ . Le présent résultat corrobore ceux de Weatherup (1994).

## Conclusion

L'utilisation de la distance généralisée de Mahalanobis  $D^2$ , comme critère de distinction entre les populations de blé dur cultivées en Algérie, nous a permis de répondre aux deux précédentes questions.

En premier, la longueur du grain (4 cas sur 10) et la longueur de l'épi (4 cas sur 10) constituent les variables les plus discriminantes ou les plus importantes dans la distinction.

En second, la similitude la plus évidente entre les populations est représentée par Oued-Zenati 368 et Bidi 17. Ce couple se différencie nettement des trois autres populations, qui sont elles-mêmes distinctes l'une de l'autre. L'avantage dans l'utilisation de la distance généralisée  $D^2$  réside dans la possibilité de pouvoir distinguer deux populations avec seulement deux variables. Cette distinction permet également d'aboutir à des explications des différences observées selon le profil. Par exemple, pour les profils 3, 8, 6 et 10, la supériorité de Hedba 3 en hauteur et en longueur de l'épi est bien mise en évidence par rapport aux autres populations. Le présent critère de proximité peut être donc avantageusement utilisé, cependant nous suggérons à l'avenir de considérer un nombre plus important de variétés et de variables.

**Tableau 1.** Valeurs de  $D^2$  observées et critiques selon les traits entre les couples de variétés

1 - ZB-OZ		2 - ZB-BD		3 - ZB-HD	
Traits	$D^2$	Traits	$D^2$	Traits	$D^2$
LGR	36.44**	LGR	41.56**	LE	118.97**
LE	480.05***	LE	238.19**	HP	1372.43***
PGE	4404.21**	PGE	2262.62**	LGR	2551.95**
HP	12510.59*	HP	6670.57*	PGE	36676.82*
4 - ZB-MBB		5 - OZ-BD		6 - OZ-HD	
Traits	$D^2$	Traits	$D^2$	Traits	$D^2$
HP	42.59**	LE	11.50*	LE	53.30**
LE	192.77**	LGR	49.72*	HP	439.78***
PGE	206.76*	PGE	361.23*	LGR	599.10**
LGR	669.99 ns	HP	920.22 ns	PGE	6406.07*
7 - OZ-MBB		8 - BD-HD		9 - BD-MBB	
Traits	$D^2$	Traits	$D^2$	Traits	$D^2$
LGR	45.80**	LE	114.50***	LGR	52.60**
LE	459.20***	HP	732.70***	LE	228.6**
PGE	4058.40**	LGR	1087.50**	PGE	2021.6**
HP	17458.80*	PGE	12162.40*	HP	10607.20*

**10 - HD-MBB**

Traits	D2
LE	97.20***
LGR	1236.80***
PGE	13705.40***
HP	44383.20*

ns: non significatif (  $P > 0.05$  ), \* :  $P < 0.05$  ,

\*\*,  $P < 0.01$  , \*\*\*,  $P < 0.005$ , D<sup>2</sup>: distance observée

Valeurs critiques

1 variable	21.20
2 variables	83.21
3 variables	595.02
4 variables	90000

Abbévations : OZ= Oued-Zénati 368 ; ZB= Ziban ; BD= Bidi 17; HD = Hedba 3 ; MBB = Mohamed Ben Bachir.

**Tableau 2.** Moyennes par population pour chaque variable, leur erreur standard ( e.s. ) et coefficients de variation ( c.v.)

Traits / Populations	HP cm	LE cm	PGEgr	LGRmm
Ziban	73.08	6.91	2.13	7.53
O.Zenati 368	90.50	7.63	2.30	8.60
Bidi 17	90.70	6.95	2.14	8.70
Hedba 3	104.6	9.10	2.02	8.40
M.B.Bachir	94.19	7.12	2.11	7.40
e.s.	2.29	0.142	0.087	0.124
c.v	3.57	2.67	5.75	2.17

**Tableau 3.** Matrice des corrélations intra-années et variétés

	HP	LE	LGR	PGE
HP	1	-0.84	0.76	-0.44
LE		1	-0.90	0.34
LGR			1	0.08
PGE				1

## Remerciements

Les auteurs remercient particulièrement Mr le Professeur Ouknider M. pour ses précieux conseils qu'il a aimablement prodigué lors de la réalisation de ce travail.

## Références

- Cochran W.G. (1962). On the performance of the lineardiscriminant function. Bulltin of the InternationalStatistical Institute. 39, 435 - 447.
- Dagnelie P., Merckx A. (1991). Using Generalized Distances in Classification of Groups. *Biometrics Journal* 33, 6, 683-95
- Dekhili M., Kharfallah N., Aggoun A., Harkati B. (1998). Distinction et homogénéité de cinq populations algériennes de blé dur (*Triticum durum* Desf.), *Cahiers Agricultures* 7, 67-71.
- Hotelling H.O. (1931). The generalisation of Student's ratio. *Annals of Mathematic Statistics* 2, 360-78.
- Laumont P., Erroux J. (1961). Inventaire des blés durs rencontrés et cultivés en Algérie. Alger: Imprimeries " La typo-litho ", 95 p.
- Mahalanobis P.C.( 1936 ). On the generalized distance in statistics. *Proceedings of National Institute of Science, India* 12, 49-55.
- UPOV 1988 ) : Principes directeurs pour la conduite de l'examen des caractères distinctifs de l'homogénéité et de la stabilité. Blé dur (*Triticum durum* Desf.). TG/120/3, Geneva, 1-15.
- Weatherup S.T.C.(1994). Use of Mahalanobis distance to measure varietal distinctness. *Plant Varieties and seeds* 7, 107-19.