

## Déterminisme génétique de la croissance en hauteur du Pin maritime (*Pinus pinaster Ait*)

### II. - Comportement interannuel Interaction génotype $\times$ année

A. KREMER

*I.N.R.A., Laboratoire d'Amélioration des Arbres forestiers,  
Pierroton, F 33610 Cestas*

#### Résumé

La croissance en hauteur du Pin maritime au cours de plusieurs années a été décrite par le modèle linéaire généralement utilisé pour l'étude de l'interaction génotype  $\times$  milieu.

L'accroissement annuel a été mesuré durant six années successives dans un test de descendance de 100 familles de demi-frères issues de pollinisation libre en forêt. Ce test est situé en landes. Les mêmes mesures ont été faites au cours de quatre années successives dans un test comprenant les mêmes familles et situé en dunes.

Trois paramètres de comportement interannuel ont été définis au niveau familial :

- l'interactivité ( $W_i$ ) dont les composantes sont :
  - la réponse de la famille aux variations écologiques annuelles ( $b_i$ ) ;
  - la stabilité autour de cette réponse ( $d_i$ ).

Dans le test situé en landes, le modèle linéaire n'explique qu'une faible partie de l'interaction famille  $\times$  année. Par contre, dans le test situé en dunes, l'hétérogénéité des droites de régression représente une majeure partie de l'interaction. Les trois paramètres définis précédemment ont été estimés au niveau familial. L'utilisation du modèle à des fins de prédiction a été discuté. Les paramètres  $b_i$  et  $d_i$  sont ensuite estimés au niveau individuel : leur héritabilité reste faible dans chaque milieu (.10 à .20).

Enfin, une analyse multivariée fournit une explication biologique de l'interaction  $F \times A$  :

- dans le test situé en landes le polycyclisme est corrélé à l'instabilité des familles ;
- dans le test situé en dunes le polycyclisme est lié à la réponse des familles.

#### Mots-clés

*Croissance en hauteur, stabilité, interaction famille  $\times$  année polycyclisme*

#### 1. - Introduction

L'utilisation en sélection d'un caractère cumulatif, tel que la croissance en hauteur chez les arbres forestiers, soulève un problème méthodologique de définition du ou des meilleur(s) critère(s) prédicteur(s) de ce caractère ; jusqu'à présent, le choix s'est porté sur le cumul des accroissements d'un nombre maximum d'années au-delà des accroissements soumis aux effets maternels, sans prendre en considération la

régularité des pousses annuelles, ou plus généralement le comportement interannuel de la croissance. C'est sur ce dernier aspect que porte la présente étude.

Une précédente étude (KREMER, 1981) s'est attachée à déterminer l'impact du polycyclisme sur la croissance en hauteur. Certains résultats laissaient présager de son rôle dans l'interaction génotype  $\times$  année. En effet, si dans de bonnes années l'avantage des génotypes polycycliques était certain, il en était différemment lors d'années sèches où génotypes polycycliques et monocycliques réalisaient des performances équivalentes.

Sur le plan biométrique, le comportement interannuel peut être interprété en termes d'additivité ou d'interaction entre les effets génétiques et ceux environnementaux de l'année dans le déterminisme de la croissance : aussi les modèles employés sont ceux utilisés dans ce type d'analyse (FINLAY-WILKINSON, 1963, EBERHART-RUSSEL, 1966).

L'hypothèse principale inhérente à ce modèle est la relation linéaire existant entre la croissance de l'arbre et les variations écologiques annuelles, estimées par un indice biologique. Le modèle, outre la description biométrique de l'interaction  $G \times A$  qu'il fournit, permet de prédire la réponse d'un génotype aux conditions climatiques annuelles et de chiffrer l'incertitude de cette prédiction. Il précise si cette réponse est due à un effet systématique (pente de la droite de régression) ou à un effet aléatoire (déviation autour de la droite). Dans le cas où la première catégorie d'effets est en cause, la réponse est prévisible, que la participation à l'interaction  $G \times A$  soit importante ou non. Dans le cas où les effets aléatoires sont importants, le modèle ne fournit guère d'informations et le sélectionneur ne peut retenir que la contribution de chaque génotype à l'interaction (interactivité). Trois paramètres utiles en sélection sont ainsi définis par le modèle, l'interactivité du génotype et ses composantes : la réponse aux variations annuelles, et la stabilité des performances autour de cette réponse. Ces paramètres ont souvent été dotés d'une signification biologique, respectivement écovalence, stabilité ou adaptabilité souvent confondus en un seul terme : stabilité.

Le modèle étant parfaitement symétrique, le même raisonnement peut être fait au niveau des années.

Il n'y a plus aucune raison pour ne pas le faire au niveau des combinaisons génotype  $\times$  année. Cette écriture très complète du modèle linéaire permet d'obtenir une meilleure structuration de l'interaction  $G \times A$  (MOAV, WOHLFARTH, 1974 ; MOAV, WOHLFARTH, HULATA, 1975 ; WRIGHT, 1971 ; WRIGHT, 1976 ; HINKELMANN, 1974). Cependant l'interprétation au niveau des milieux sera faite dans une étude ultérieure.

Outre la structuration de l'interaction et l'estimation des paramètres de stabilité au niveau des génotypes, la présente étude se propose à partir d'un dispositif bista-tionnel :

- de vérifier la valeur de prédiction du modèle et son utilisation à une telle fin ;
- de déterminer le contrôle génétique des paramètres du comportement interannuel de la croissance ;
- de trouver, à partir d'une analyse multivariable une interprétation biologique de ces paramètres.

## 2. - Matériel

Le dispositif expérimental utilisé pour cette étude est un test de descendance bistationnel (2.44.3.1. et 2.44.3.2.) planté en automne 1965 et répété en landes et en dunes. Schématiquement, la dune correspond à une frange littorale large de quelques kilomètres, caractérisée par un relief assez chahuté. Au plan pédologique, le sol est squelettique et ne présente aucun horizon bien marqué. Enfin, le plan d'eau se situe en général à un niveau très profond.

A l'opposé, la lande est uniformément plate. La seule hétérogénéité au plan écologique est due à la profondeur du plan d'eau qui reste toutefois assez superficiel. Les sols sont des podzols qui se différencient à mesure que le niveau de la nappe descend en été.

Le dispositif comprend 100 familles de demi-frères issus de pollinisation libre en forêt toutes communes aux deux tests dont les caractéristiques sont explicitées dans une précédente étude (KREMER, 1981).

Les mesures ont porté sur un échantillon systématique de 4 plants par parcelle unitaire (P.U.) à Soustons et de 3 plants par P.U. à Saint-Alban. Compte tenu des arbres morts ou non mesurables, les effectifs par famille ne sont pas rigoureusement constants. C'est la raison pour laquelle dans l'analyse de variance, la moyenne harmonique des effectifs a été utilisée dans le calcul des carrés moyens (GOSLEE, LUCAS, 1965).

Les caractères mesurés ont été les suivants :

*Saint-Alban* : Hauteurs totales successives de la 7<sup>e</sup> à la 12<sup>e</sup> année de végétation.

*Soustons* : Hauteurs totales successives de la 9<sup>e</sup> à la 12<sup>e</sup> année de végétation.

En outre, pour chaque pousse annuelle, le niveau du verticille primaire par rapport au sol a été relevé. En effet, la croissance en hauteur annuelle chez le pin maritime est occasionnellement caractérisée par plusieurs cycles successifs de croissance (polycyclisme). Après l'âge de 3 ou 4 ans, le nombre de cycles est d'un ou deux, très rarement trois. Le premier est appelé accroissement primaire, les seconds et troisièmes (quand ils s'expriment) sont cumulés et appelés accroissement secondaire. On dispose ainsi facilement des 6 accroissements totaux, primaires et secondaires à Saint-Alban et des 4 accroissements totaux, primaires et secondaires à Soustons.

## 3. - Méthodes

### 3.1. Analyse globale : choix de paramètres de stabilité

La performance phénotypique  $X_{ijh}$ , d'un individu  $h$  est le résultat d'un effet  $g_i$  propre à la population, d'un effet  $e_j$  de l'année et d'un effet  $\varepsilon_{ijh}$  dû à l'individu  $h$ .

$$(1) \quad X_{ijh} = \mu + g_i + e_j + \varepsilon_{ijh}$$

$$\begin{array}{ll} g_i \sim N(0, \sigma_g^2) & i = 1 \dots \dots n \\ e_j \sim N(0, \sigma_e^2) & j = 1 \dots \dots k \\ \varepsilon_{ijh} \sim N(0, \sigma_\varepsilon^2) & h = 1 \dots \dots p \end{array}$$

L'additivité de  $g$  et de  $e$  est rarement vérifiée : force est d'introduire un terme d'interaction  $(ge)_{ij}$ .

$$(2) \quad X_{ijh} = \mu + g_i + e_j + (ge)_{ij} + \varepsilon_{ijh} \quad (ge)_{ij} \sim N(0, \sigma_{ge}^2)$$

Parler de stabilité ou d'instabilité de  $g$  ou de  $e$  c'est attribuer à l'un de ces effets tout ou partie de  $(ge)$ . D'où la décomposition faite par FINLAY & WILKINSON (1963), suivie par celle d'EBERHART & RUSSEL (1966).

$$(ge)_{ij} = \beta_i e_j + r_{ij}$$

$$(ge)_{ij} = \alpha_j g_i + s_{ij}$$

Le modèle (2) s'écrit :

$$(3) \quad X_{ijh} = \mu + g_i + e_j(1 + \beta_i) + r_{ij} + \varepsilon_{ijh}$$

$$(4) \quad X_{ijh} = \mu + e_j + g_i(1 + \alpha_j) + s_{ij} + \varepsilon_{ijh}$$

Dans ces deux modèles, l'interaction au niveau de la SCE est partitionnée en deux termes :

— le premier traduit l'hétérogénéité des pentes des droites de régression  $\beta_i$  (ou  $\alpha_j$ ) ;

— le second est dû à la somme des déviations par rapport à ces droites ( $r_{ij}$  ou  $s_{ij}$ ).

$\beta_i$  et  $r_{ij}$  ou  $\alpha_j$  et  $s_{ij}$  sont utilisés comme paramètres de stabilité (FINLAY, WILKINSON, 1963 ; BUCIO, ALANIS, 1966 ; EBERHART, RUSSEL, 1966 ; MOAV, WOHLFARTH, 1974).

Les modèles (3) et (4) sont en partie redondants : en effet les parts de l'interaction due à l'hétérogénéité des  $\beta_i$  et des  $\alpha_j$  ne sont pas indépendantes, on peut dire qu'une composante de l'interaction peut être expliquée par la régression ( $\gamma$ ) sur les effets génotype ( $g_i$ ) et année ( $e_j$ ) simultanément.

$$(ge)_{ij} = \gamma g_i e_j + \beta'_i e_j + \alpha'_j g_i + t_{ij}$$

(WRIGHT, 1971 ; MOAV, WOHLFARTH, 1974 ; MOAV, WOHLFARTH, HULATA, 1975 ; HINKELMANN, 1974 ; WRIGHT, 1976 ; MANDEL, 1969).

$$(5) \quad X_{ijh} = \mu + g_i + e_j + \gamma g_i e_j + \beta'_i e_j + \alpha'_j g_i + t_{ij} + \varepsilon_{ijh}$$

$\gamma$  est le coefficient de régression de  $\beta_i$  sur  $g_i$  ou de  $\alpha_j$  sur  $e_j$

$\beta'_i$  et  $\alpha'_j$  sont les coefficients régressions résiduelles sur  $e_j$  et  $g_i$

Les significations de  $\gamma$ ,  $\beta'_i$  et  $\alpha'_j$  sont les suivantes :

—  $\gamma$  est un effet d'échelle dû à la corrélation entre la variation génotypique pour une année donnée et le niveau moyen de cette année.  $\gamma$  représente ainsi la « fausse » interaction non liée aux seuls changements de classements de populations (DICKERSON, 1952 ; ROBERTSON, 1959 ; BURDON, 1978).

Graphiquement, cela signifie que lorsque tous les  $\beta'_i$  et  $\alpha'_j$  sont nuls, toutes les droites des modèles (3) et (4) sont concourantes en un point donné.

— Compte tenu de ce qui vient d'être dit,  $\beta'_i$  et  $\alpha'_j$  sont des coefficients de stabilité plus explicites que  $\beta_i$  et  $\alpha_j$ .

### 3.2. Estimation des paramètres de stabilité

#### 3.2.1. Paramètres se référant au modèle (2)

$$(2) \quad x_{ij} = \mu + g_i + e_j + (ge)_{ij} + \bar{\varepsilon}_{ij}.$$

En se plaçant au niveau des moyennes  $x_{ij}$ , SHUKLA (1972) a proposé la décomposition suivante des termes résiduels (interaction + erreur) :

$$\delta_{ij} = (ge)_{ij} + \bar{\varepsilon}_{ij}.$$

$$\sigma_{\delta_i}^2 = \sigma_{(ge)_i}^2 + \frac{\sigma_{\varepsilon}^2}{p}$$

$\sigma_{\delta_i}^2$  est appelé *variance de stabilité*.

Un estimateur non biaisé de  $\sigma_{\delta_i}^2$  est (HINKELMANN, 1974)

$$\hat{\sigma}_{\delta_i}^2 = \frac{1}{(n-1)(n-2)(k-1)} \left[ n(n-1)W_i - \sum_i W_i \right]$$

$$\text{Avec : } W_i = \sum_j W_{ij} = \sum_j (x_{ij} - x_{i..} - x_{.j.} + x_{...})^2$$

$W_i$  correspond à l'écovalence de WRICKE (1962). Ce terme est appelé dans le texte *interactivité* du génotype  $i$  car il représente la participation globale de ce génotype à la SCE du terme interaction.

La variance de stabilité peut être testée par un test F (SHUKLA, 1972) :

$$F = \frac{p \sigma_{\delta_i}^2}{CM(\varepsilon)} \quad (k-1) \text{ et } nk(p-1) \text{ dl}$$

CM( $\varepsilon$ ) : carré moyen résiduel

#### 3.2.2. Paramètres se référant au modèle (3)

$$(3) \quad X_{ijh} = \mu + g_i + e_j (1 + \beta_i) + r_{ij} + \varepsilon_{ijh}$$

Soit  $b_i$  l'estimation de  $(1 + \beta_i)$ .

L'interactivité  $W_i$  peut également s'écrire :

$$W_i = \sum_j [(x_{ij} - x_{i..}) + b_i(x_{.j.} - x_{...}) - b_i(x_{.j.} - x_{...}) - (x_{.j.} - x_{...})]^2$$

En développant cette expression

$$W_i = \underbrace{\sum_j (b_i - 1)^2 (x_{i,j} - x_{i...})^2}_{\substack{\text{Contribution due à l'écart} \\ \text{par rapport à 1} \\ \text{du coefficient} \\ \text{de régression}}} + \underbrace{\sum_j [(x_{i,j} - x_{i...}) - b_i (x_{i,j} - x_{i...})]^2}_{\sum r_{ij}^2}$$

On retrouve ici la contribution des deux paramètres de stabilité d'EBERHART & RUSSELL dans l'interactivité du génotype  $i$ .

— le premier ( $b_i$ ) est appelé *réponse* du génotype aux variations annuelles ;

— le second ( $d_i = \sum r_{ij}^2 / (k - 2)$ ) exprimé en variance représente la *stabilité* autour de cette réponse.

Les contributions à l'interaction exprimée par  $b_i$  et  $d_i$  peuvent être testées de la manière suivante :

Pour  $b_i$

$$t = \frac{(b_i - 1) \sqrt{\sum_j (x_{i,j} - x_{i...})^2}}{\sqrt{d_i}} \quad [k - 2] \text{ dl.}$$

Pour  $d_i$

$$F = \frac{p \cdot d_i}{\text{CM}(\varepsilon)} \quad [k - 2 \text{ et } nk(p - 1) \text{ dl}]$$

### 3.23. Paramètres se référant au modèle (5)

$$(5) \quad X_{ijh} = \mu + g_i + e_j + \gamma g_i e_j + \beta'_i e_j + \alpha'_j g_i + t_{ij} + \varepsilon_{ijh}$$

Pour une population  $i$  donnée :

$$X_{ijh} = \mu + g_i + (1 + \beta'_i + \gamma g_i) e_j + r_{ij} + \varepsilon_{ijh}$$

Un estimateur  $d$  de  $\gamma$  est le coefficient de régression de  $\beta_i$  sur  $g_i$

$$d = \frac{\sum_i (x_{i..} - x_{...}) b_i}{\sum_i (x_{i..} - x_{...})^2}$$

Un estimateur  $b'_i$  de  $(\beta'_i + 1)$  :

$$b'_i = b_i - d (x_{i..} - x_{...})$$

Les analyses de variances et les estimations des paramètres de stabilité sont faites à partir du programme FLOYD écrit en BASIC pour P 6060 OLIVETTI.

### 3.3. Limites d'application du modèle

Dans les modèles expliqués dans le paragraphe 3.1., les effets principaux sont supposés aléatoires.

Cette hypothèse peut se justifier pour les raisons suivantes :

#### *Effet famille*

Les arbres mères  $F_0$  dont sont issues les familles de demi-frères, ont été sélectionnés phénotypiquement en forêt landaise sur la base de critères de vigueur et de forme. Par la suite, on a pu montrer que cette sélection s'était avérée inefficace pour la vigueur, caractère utilisé dans la présente étude. En outre, d'autres études utilisant les gènes marqueurs tels que les terpènes ont montré que sur l'ensemble du massif la population était homogène (en équilibre panmictique). On peut donc penser que les 100 arbres mères représentent un échantillon aléatoire de la population landaise.

#### *Effet année*

Cet effet est la résultante de deux types de facteurs : l'âge physiologique de l'arbre et les effets pédoclimatiques de l'année. La période couverte par l'étude se situe dans une zone homogène de la courbe de croissance de l'arbre (7<sup>e</sup> à la 12<sup>e</sup> année) : dans cette tranche d'âge l'accroissement courant annuel est pratiquement constant et donc indépendant de l'année. Les facteurs climatiques jouent donc un rôle prépondérant dans l'effet global de l'année, et leur distribution est quasiment aléatoire au cours des années.

Enfin, l'étude de l'interaction génotype  $\times$  année sur des plantes pérennes par une analyse de variance univariable soulève le problème de la non-indépendance des termes résiduels.

Les mesures étant effectuées chaque année sur les mêmes individus il vient que :

$$\text{cov}(\varepsilon_{ijh}, \varepsilon_{ij'h}) \neq 0$$

Entre les années  $j$  et  $j'$  existe une autocorrélation qui peut être exprimée par l'importance de  $K_{j,j'}$

$$K_{j,j'} = \frac{\text{cov}(\varepsilon_{ijh}, \varepsilon_{ij'h})}{\sigma_{ijh} \times \sigma_{ij'h}}$$

Plus généralement le terme  $\varepsilon_{ijh}$  peut être décomposé en un effet individu (aléatoire) et un effet interaction individu  $\times$  année.

$$\varepsilon_{ijh} = v_{ih} + w_{ijh}$$

Par définition même du coefficient de corrélation intra classe, un estimateur du coefficient d'autocorrélation  $K$  est alors :

$$K = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_w^2}$$

Les valeurs de  $K$  à Saint-Alban et à Soustons sont respectivement les suivantes : 0,04 ( $\approx 0$ ) et 0,14. Ces valeurs étant relativement faibles voire nulles, l'hypothèse de la non-dépendance des termes résiduels est vérifiée dans le cas présent et le modèle devient applicable.

## 4. - Résultats

## 4.1. Comportement interannuel

4.1.1. Interaction génotype  $\times$  année

TABLEAU 1  
Analyse de variance  
Analysis of variance

Source de variation . . . . .	Saint-Alban		Soustons	
	d. l.	Carrés moyens	d. l.	Carrés moyens
Famille . . . . .	99	1 753 ***	99	1 303 ***
Année . . . . .	5	256 582 ***	3	259 194 ***
Interaction F $\times$ A . . . . .	495	541 ***	297	364 ***
Joint régression . . . . .	1	383 NS	1	29 NS
Résiduelle . . . . .	494	542 ***	296	365 ***
Régression (années) . . . .	98	632 *	98	509 ***
Régression (familles) . . .	4	1 951 ***	2	303 NS
Déviations . . . . .	392	505 ***	196	293 ***
Erreur . . . . .	14 577	195	9 179	209

d = .0051

d = .0017

\*\*\* : test F significatif au seuil de 1 %

\*\* : test F significatif au seuil de 5 %

\* : test F significatif au seuil de 10 %

Les résultats de l'analyse de variance (modèle 5) se trouvent dans le tableau 1. Les principales conclusions en sont les suivantes :

— L'effet dû à la régression simultanée sur les années et les familles ( $\gamma$ ) est non significatif dans les deux cas. Il est donc possible d'utiliser et d'interpréter les modèles (3) et (4) indépendamment l'un de l'autre. Par contre, la prédiction de combinaisons famille  $\times$  année non testées est inefficace.

— Une bonne partie de l'interaction peut être expliquée par les régressions résiduelles sur les familles ( $\alpha'$ ) et les années ( $\beta'$ ). A Saint-Alban, la part de l'interaction due à l'hétérogénéité entre les droites de régression sur les années ajoutée à l'hétérogénéité entre les droites de régression sur les familles est de 26 p. 100, à Soustons elle est de 47 p. 100. Ceci permet, dans une certaine mesure de prédire une part de l'interaction et donc de limiter ses effets dans un schéma de sélection.

Quand  $\gamma$  est négligeable, la structuration de l'interaction peut être faite de deux manières :

- soit par la régression sur les années ;
- soit par la régression sur les familles.



Ceci est assez bien illustré par l'exemple étudié :

— A Saint-Alban, l'hétérogénéité des droites de régression sur les années est peu significative (à 10 p. 100) alors que celle des droites de régression sur les génotypes l'est au seuil de 1 p. 100.

— A Soustons, la situation est inverse.

Il faut toutefois remarquer que les régressions sur les familles n'expliquent qu'une petite part de l'interaction (3 p. 100 à Saint-Alban et 1 p. 100 à Soustons). Il reste que dans cet exemple précis, les décisions du sélectionneur seront totalement différentes dans l'un ou l'autre cas :

— A Saint-Alban, les effets de l'interaction seront diminués par le choix d'une année optimale pour la sélection fait à l'aide de paramètres appropriés.

— A Soustons, réduire l'interaction c'est effectuer une sélection pour les familles les plus stables.

L'effet de  $\gamma$  étant non significatif, les coefficients de stabilité seront estimés à partir du modèle (3) et non à partir du modèle (5).

#### 4.12. *Interactivité des familles*

L'écovalence de WRICKE réalise une décomposition de l'interaction en participations dues à chaque famille, qui représente sa responsabilité dans l'interaction. L'analyse de variance a montré qu'une partie de l'interactivité est due à un effet systématique (droite de régression  $\neq 1$ ) et une autre à des effets aléatoires (déviations autour de la droite).

Ce terme renferme une double signification :

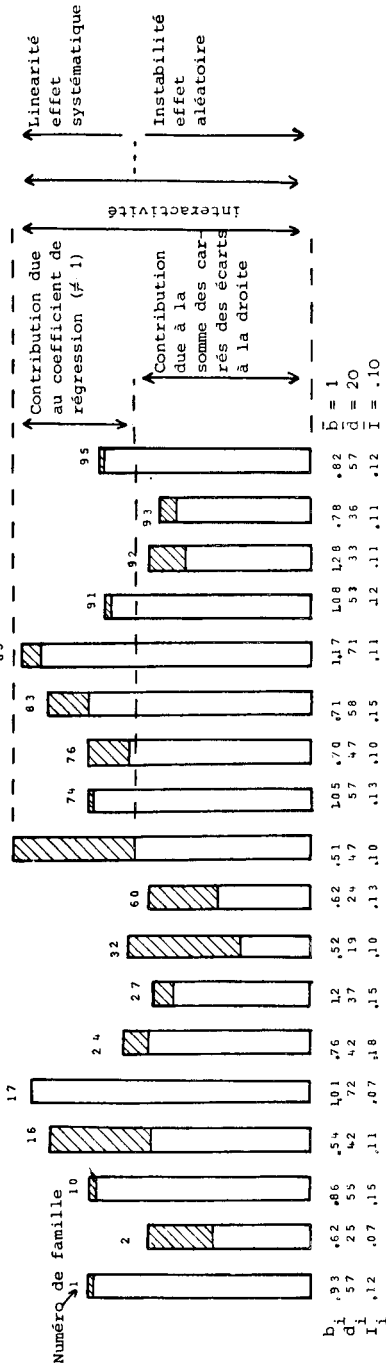
- la réponse à des variations annuelles (en l'occurrence linéarité de la réponse) ;
- la stabilité des réactions autour de la réponse.

Ceci est illustré dans les figures 1(a) et 2(a). Sur ces schémas sont représentées les interactivités des familles pour lesquelles les variances de stabilité sont significativement différentes de zéro (au seuil de 5 p. 100). Dans les deux cas (Saint-Alban et Soustons), le nombre de ces familles est important (1/5 de l'effectif).

A Saint-Alban, elles représentent 40 p. 100 de la SCE du terme interaction, à Soustons 50 p. 100.

Pour chaque famille, l'interactivité est partitionnée en somme des carrés due à la régression et somme de carrés due aux déviations (cf. 3.22.). En moyenne, au niveau des familles, on retrouve les résultats déjà donnés par l'analyse de variance. A Soustons, une grande partie de l'écovalence des familles les plus interactives peut être augmentée par la possibilité de prédiction des performances de ces familles (26, 31, 36, 37, 51, 52, 55, 75, 87, 88, 97, 99). Ceci n'est vrai que pour un nombre très réduit de familles à Saint-Alban (32, 62).

(a) Familles à interactivité élevée  
(a) Families with high interactivity



(b) Familles à réponse significative  
(b) Families with significant response

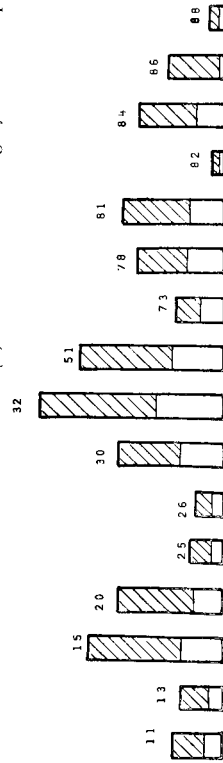


FIG. 1

Interactivité, réponse, stabilité  
Saint-Alban 2.44.3.1.

Interactivity, response, and stability of different families

$b_i$  : coefficient de régression (voir 3.22)

$d_i$  : somme des carrés des écarts à la droite de régression (voir 3.22)

$I_i$  : indice du polycyclisme ( $\Sigma$  pousses secondaires /  $\Sigma$  pousses totales) (voir 4.32)

#### 4.13. Réponse des familles

Les droites de régression propres à chaque famille sont intéressantes à double titre :

- elles précisent la nature de leur participation à l'interaction ( $b_i \neq 1$ ) ;
- elles permettent de prédire leur réponse dans d'autres années.

Dans les figures 1(b) et 2(b), figurent les écovalences des familles dont la contribution à l'interaction due au coefficient de régression est significative au seuil de 10 p. 100 (voir 3.22.). A Saint-Alban, quand cette contribution est significative, elle est inférieure à la moyenne. Le fait que seule la famille 32 soit commune avec le sous-ensemble de familles présentes dans la figure 1(a) est une autre illustration de la faible contribution des coefficients de régression à l'interaction.

Pour Saint-Alban, la faible contribution globale de l'hétérogénéité des pentes des droites de régression peut s'expliquer de deux manières :

- les familles les plus interactives (écovalence la plus faible) le sont surtout par leurs écarts aux droites de régression (figure 1a) ;
- les familles ayant une pente de droite de régression très différente de 1 ont une contribution totale à l'interaction faible (figure 1b).

A Soustons, la situation est différente. La moitié de l'effectif des familles présentes dans la figure 2(b) est commune à celui de la figure 2(a). Il reste cependant que les familles de l'autre moitié n'ont qu'une interactivité très faible.

Même si, à Saint-Alban, les pentes des droites de régression n'expliquent qu'une faible partie de l'interaction, les réponses des génotypes sont bien interprétées par le modèle linéaire. En effet, les familles dont les coefficients de régression ne sont pas significativement  $\neq$  de 0 [test t à  $(k - 2)$  dl] sont peu nombreuses (9 à Saint-Alban). A Soustons, elles sont au nombre de 17. La signification d'un tel test, peut être discutée quand la variable expliquée et la variable explicative ne sont pas strictement indépendantes : dans l'indice biologique du milieu  $j$  (variable explicative) intervient la moyenne de la famille  $i$  (variable expliquée). Dans le cas présent, où le nombre de familles est élevé, les deux variables peuvent être considérées comme indépendantes.

#### 4.14. Stabilité des familles

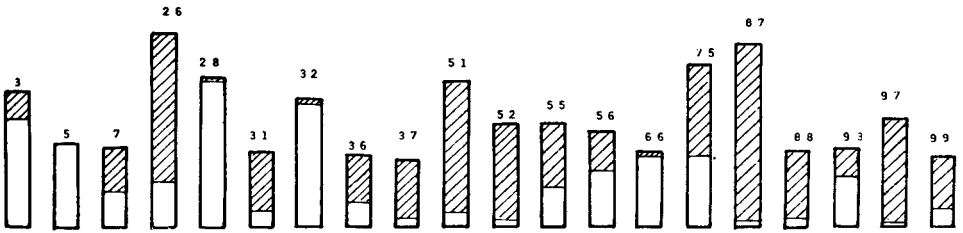
L'instabilité d'un génotype donné peut être définie comme l'interactivité de ce génotype avec l'année purgée de l'effet de linéarité : il s'agit en fait des déviations autour de la droite de régression, par définition imprévisible (cf. figure 1a). Un test F (cf. 3.22.) permet d'identifier les familles :

- A Saint-Alban : les familles particulièrement instables sont la 1, 10, 17, 62, 74, 76, 83, 85, 91, 95. Elles sont toutes représentées sur la figure 1(a).
- A Soustons, ce sont les familles 3, 5, 28, 32, 56, 66, 72 75. Seule la famille 72 ne fait pas partie de la figure 2(a).

Une représentation globale pour l'ensemble des familles de la stabilité, de la réponse des familles et de leurs performances est donnée dans les figures 3 et 4. Les familles se trouvent sur la droite passant par le centre et le point dont l'abscisse et l'ordonnée sont respectivement les écarts centrés réduits (ECR) de la somme des carrés des écarts à la droite de régression ( $d_i$ ) et l'ECR de la pente de la droite de régression.

(a) Familles à interactivité élevée

(a) Families with high interactivity



(b) Familles à réponse significative

(b) Families with significant response

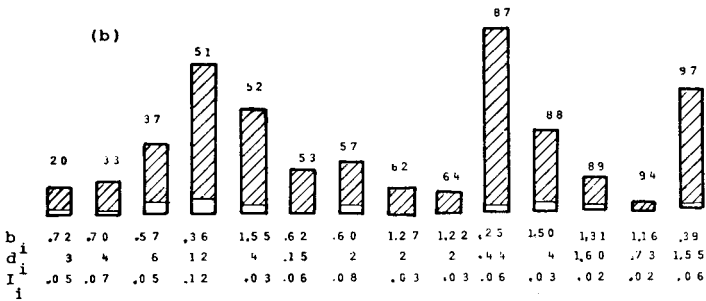


FIG. 2

Interactivité, réponse, stabilité de différentes familles  
Soustons 2.44.3.2.

Interactivity, response and stability of different families

La distance au centre représente le niveau moyen des performances sur l'ensemble des années, les moyennes étant centrées par rapport à la moyenne la plus basse.

Sur ces deux figures, on peut noter qu'une majorité de points se situent dans les deux cadrans de gauche (supérieur et inférieur). Dans la plupart des cas le comportement des familles est donc prévisible. La forte interaction au niveau global ne s'explique donc que par un faible nombre de familles très interactives.

D'autre part, dans chaque station, la vigueur est indépendante de la stabilité et de la réponse des familles : les coefficients de corrélation entre stabilité, réponse, interactivité et la vigueur ne sont pas significatifs (voir tableau 2).

TABLEAU 2

*Coefficients de corrélations entre interactivité, réponse, stabilité et vigueur*  
*Correlation coefficients between interactivity, response, stability and vigor*

Soustons

Saint-Alban

	Soustons			Saint-Alban			
$w_i$ .....				$w_i$	$b_i$	$d_i$	$x_i$
$b_i$ .....	-.24 **				-.26 ***	.92 ***	.05
$d_i$ .....	.61 ***	.06				-.22 **	.07
$x_i$ .....	-.13	.01	-.04				.06

\*\*\* : test t significatif au seuil de 1 %

\*\* : test t significatif au seuil de 5 %

$w_i$  : écovalence

$b_i$  : pente de la droite de régression

$d_i$  : somme des carrés des écarts à la droite de régression exprimée en variance

$x_i$  : moyenne des accroissements sur les k années

$w_i$  : ecovalence

$b_i$  : slope of regression line

$d_i$  : sum of squares of deviations from regression line expressed in variance

$x_i$  : mean of shoot increments on the k years values

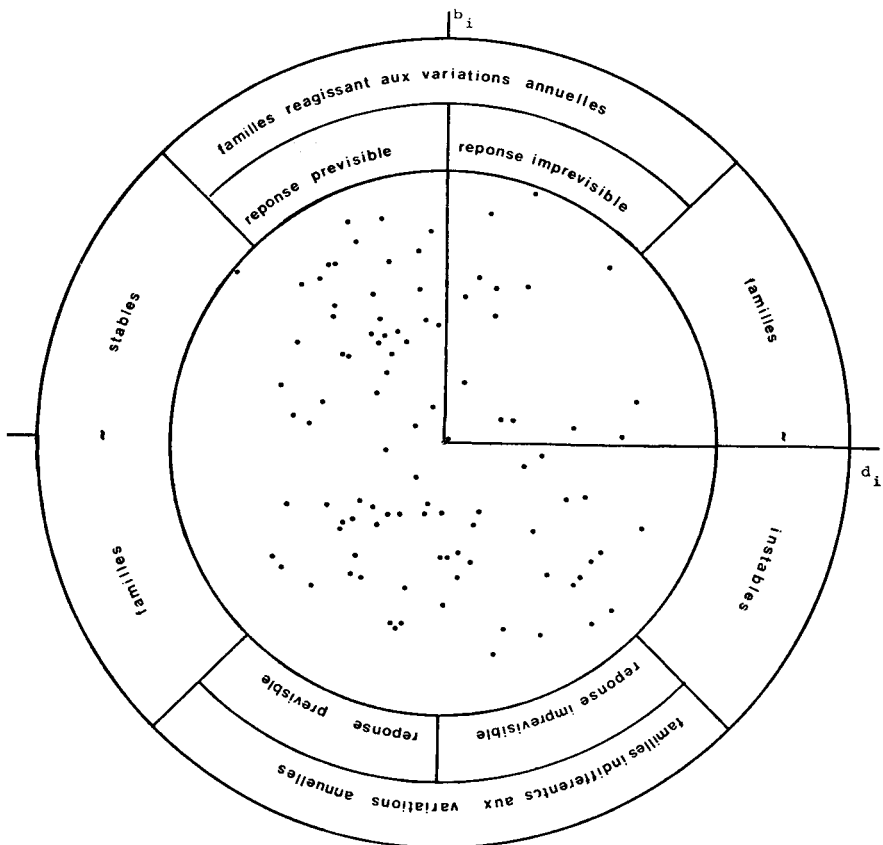


FIG. 3 :  $b_i = .17$  (en écart type);  $d_i = .22$ ;  $x_i = .80$

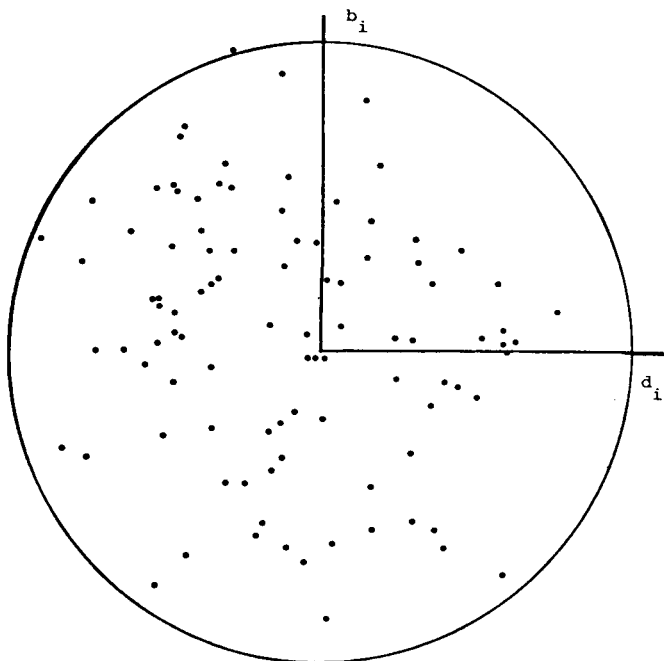


FIG. 4 :  $b_i = .17$  (en écart type);  $d_i = .15$ ;  $x_i = .77$

4.2. Interaction famille  $\times$  station

Les coefficients de corrélation entre les moyennes familiales des différents caractères dans les deux milieux, sont les suivants :

$$\begin{array}{ll} w_i = -.04 & d_i = -.02 \\ b_i = -.24^{**} & x_i = .33^{***} \end{array}$$

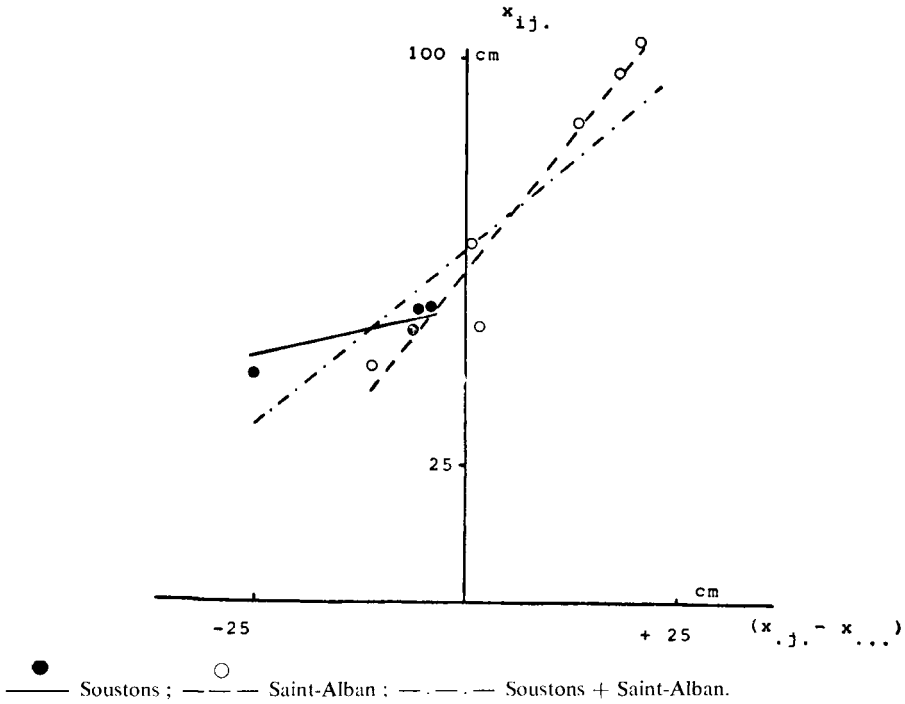


FIG. 5

*Droites de régression de la famille 87*  
*Regression lines of family 87*

Si la vigueur n'est que peu affectée par l'interaction génotype  $\times$  milieu, il n'en va pas de même pour le paramètre réponse aux variations annuelles. Bien que le coefficient de corrélation entre les moyennes familiales observées dans les deux milieux soit relativement faible il n'en est pas moins significatif. Sa valeur négative dénote un comportement différent dans les deux stations. Ce résultat abonde dans le sens de ceux d'une précédente étude (KREMER, 1981) qui montrait qu'à des conditions annuelles différentes (sèches ou humides) correspondait un comportement différent des polycycliques et des monocycliques.

FIG. 3 : Saint-Alban

*Représentation schématique des performances, réponses et stabilités des familles*  
*Schematic graph of the performance response, and stability of all the families*

FIG. 4 : Soustons

*Représentation schématique des performances, réponses et stabilités des familles*  
*Schematic graph of the performance, response and stability of all the families*

TABLEAU 3  
*Valeur estimée (V.E.) et valeur réelle (V.R.) des accroissements*  
*Expected value (V.E.) and real values (V.R.) of shoot increments*

Famille	Régression Saint-Alban		1973 Soustons		1974 Soustons		1975 Soustons		1976 Soustons	
	b <sub>i</sub>	d <sub>i</sub>	V.E.	V.R.	V.E.	V.R.	V.E.	V.R.	V.E.	V.R.
25	1.22 ± .08	3.46	53.82 ± 3.24	57.54	78.68 ± 2.18	75.92	79.93 ± 2.15	79.58	75.08 ± 2.18	78.94
82	.87 ± .06	1.89	68.40 ± 2.39	62.83	86.13 ± 1.61	85.42	87.01 ± 1.59	85.46	86.31 ± 1.61	85.67
86	1.32 ± .06	1.80	53.52 ± 2.33	60.75	80.42 ± 1.57	82.54	81.77 ± 1.55	80.54	80.70 ± 1.57	79.33
88	.84 ± .05	1.13	72.86 ± 1.85	53.21	89.98 ± 1.25	81.25	90.83 ± 1.23	86.67	90.15 ± 1.24	84.04
96	1.03 ± .08	3.51	63.58 ± 3.26	62.63	84.57 ± 2.20	82.71	85.62 ± 2.17	86.25	84.79 ± 2.19	83.13



Une illustration de la corrélation négative du paramètre réponse entre les deux stations est donnée dans la figure 5. La famille 87 réagit plus vivement dans une station favorable (Saint-Alban) que dans une station défavorable (Soustons). Ce comportement est en fait idéal et fait apparaître la limite du modèle linéaire habituellement utilisé. Sans doute le modèle curvilinéaire permettrait-il de mieux détecter de telles familles.

Enfin, malgré la corrélation négative de  $b_i$  entre les 2 milieux, un intérêt non négligeable du modèle est de pouvoir l'utiliser à des fins de prédiction. Est-il possible de prédire les performances annuelles à Soustons en utilisant l'information issue de Saint-Alban ? Pour vérifier cette hypothèse, les accroissements annuels à Soustons des cinq familles les plus stables à Saint-Alban ( $d_i$  le plus faible) ont été estimés en utilisant leur droite de régression définie à Saint-Alban (cf. tableau 3).

Mis à part l'année 73, qui correspond à l'indice biologique ( $x_{.j} - x_{...}$ ) le plus faible, les correspondances entre valeurs estimées et valeurs réelles ne sont point trop discordantes.

#### 4.3. *Interprétation génétique et biologique du comportement interannuel de la croissance en hauteur dans deux stations différentes*

##### 4.31. *Déterminisme génétique de la réponse et de la stabilité*

Pour chaque milieu (Soustons et Saint-Alban) les coefficients  $b_i$  et  $d_i$  ont été définis non pas au niveau familial mais au niveau individuel, selon le modèle suivant :

$$X_{ijh} = \mu_{ih} + b_{ih} I_j + S_{ijh}$$

$$I_j = x_{.j} - x_{...}$$

$$d_{ih} = \frac{1}{k - 2} \sum_j S_{ijh}^2$$

Les deux coefficients  $b_{ih}$  et  $d_{ih}$  caractérisent la réponse et la stabilité de l'individu  $h$  de la famille  $i$  au cours des  $k$  années.

Dans le tableau ci-dessous se trouvent les valeurs d'héritabilité au sens strict de la réponse et de la stabilité (avec l'erreur standard). L'héritabilité est estimée par le coefficient de corrélation intra classe. L'erreur standard est estimée par les formules de BECKER.

	Réponse $b_{ih}$	Stabilité $d_{ih}$
Saint-Alban .....	.093 ± .036	.107 ± .037
Soustons .....	.202 ± .034	.025 ± .026

Dans les conditions où les paramètres de réponse et de stabilité s'expriment au mieux ( $b_{ih}$  à Soustons et  $d_{ih}$  à Saint-Alban), ils sont également sous contrôle génétique non négligeable. En ce qui concerne le paramètre réponse, l'héritabilité à Soustons est du même ordre de grandeur que celle de l'accroissement annuel lui-même.

A la lumière de ces résultats, il pourrait être envisagé d'utiliser deux composantes de la vigueur dans les schémas de sélection au moins à Soustons :

- les performances elles-mêmes (cumul des accroissements totaux sur un nombre maximum d'années);
- un indice de régularité des performances annuelles (coefficient de régression  $b_{ih}$ ).

Les résultats à Saint-Alban sont plus équivoques. Ceci tient avant tout à l'efficacité relative du modèle linéaire dans cette station. Néanmoins, les valeurs d'héritabilité des deux paramètres sont différentes de zéro et indiquent un contrôle génétique faible de leur expression.

#### 4.32. Polycyclisme, réponse et stabilité

Dans le but de trouver une interprétation biologique du comportement interannuel différent dans l'une et l'autre station une analyse en composantes principales a été faite sur différentes composantes de l'accroissement en hauteur.

- caractères de performance
  - Moyenne des accroissements annuels sur les k années ( $x_{i..}$ )
  - Indice de polycyclisme : ( $I_i$ )
  - $\Sigma$  des pousses secondaires/ $\Sigma$  pousses totales
- caractères de comportement
  - Réponse ( $b_i$ )
  - Stabilité  $d_i$

Les résultats se trouvent dans le tableau 4.

TABLEAU 4  
Corrélation entre composantes principales et caractères  
Correlations coefficients between principal components and characters

<i>Saint-Alban</i>					
	$b_i$	$d_i$	$x_i$	$I_i$	VP
$C_1$ .....	— .01	.38	.80	.84	37
$C_2$ .....	.81	— .70	.25	.09	30.49
$C_3$ .....	.58	.59	— .24	— .03	18.61
$C_4$ .....	.03	.14	.50	— .54	13.90
<i>Soustons</i>					
	$b_i$	$d_i$	$x_i$	$I_i$	VP
$C_1$ .....	.82	.16	— .07	— .81	34.23
$C_2$ .....	.10	— .68	.74	.09	25.91
$C_3$ .....	.06	.71	.66	.15	24.09
$C_4$ .....	.56	— .08	— .08	.55	15.77

$C_1, C_2, C_3, C_4$  : composante principale

$I_i$  : indice du polycyclisme : rapport de la somme des pousses secondaires sur la somme des pousses totales

VP : valeur propre en % de la variance totale

$C_1, C_2, C_3, C_4$  : principal component

$I_i$  : index of polycyclism, ratio of the sum of second shoot increments on the sum of total shoot increments

VP : eigen value expressed as percentage of total variance

Dans les deux cas, les composantes exprimant le maximum de la variation totale intègrent le polycyclisme avec une caractéristique du comportement interannuel (stabilité ou réponse).

— A Saint-Alban, la composante 1 pourrait avoir la signification suivante : vigueur, polycyclisme et instabilité.

— A Soustons : cette signification est toute différente. Réaction aux variations annuelles et monocyclisme et inversement indifférence aux variations annuelles et au polycyclisme.

L'explication pourrait être la suivante :

— Dans des conditions riches (Saint-Alban), dans lesquelles le polycyclisme s'exprime régulièrement à peu près pour l'ensemble des génotypes, l'instabilité de la croissance annuelle est essentiellement due aux variations de l'amplitude des pousses secondaires. Cette cause affecte indifféremment tous les génotypes.

— Dans des conditions pauvres (Soustons) on peut distinguer les familles à comportement monocyclique et à comportement polycyclique.

- Les premières dans les années les plus défavorables auront peu tendance à faire un second cycle. Par contre, dans les années aux conditions écologiques favorables, elles réaliseront un second cycle. Ces années plutôt rares font que ces génotypes sont appelés monocycliques. Cette interprétation explique les  $b_1$  élevées de ces génotypes (cf. figure 2b).

- Les secondes ont tendance à réaliser plus généralement un second cycle : ainsi même dans des conditions difficiles, leur aptitude au polycyclisme fait que la pousse secondaire s'exprime. Cette pousse joue en quelque sorte le rôle de tampon par rapport aux années riches. Ceci explique que les  $b_1$  des génotypes les plus polycycliques soient inférieures à 1, exprimant ainsi une relative indifférence aux variations annuelles.

D'une manière générale, à Saint-Alban, dans des conditions relativement favorables, la croissance du pin maritime garderait un tempérament juvénile jusqu'à un âge plus élevé qu'à Soustons. Ce tempérament se caractérise par la possibilité d'induire et d'allonger la même année une seconde pousse. Ce qui autorise à penser que la distinction entre individus polycycliques et individus monocycliques est moins nette qu'à Soustons.

La répartition des familles sur les graphes des composantes principales (cf. figures 6, 7) est à comparer avec les figures (1) et (2).

Les familles individualisées sur ces dernières figures se retrouvent groupées sur les premières :

— A Saint-Alban : toutes les familles à interactivité forte (fig. 1a) sont situées dans le cadran supérieur gauche.

— A Soustons : les familles à interactivité forte sont réparties en deux zones (fig. 2b) :

- cadran supérieur gauche : familles monocycliques à  $b_1 > 1$  ;
- cadran inférieur gauche : familles polycycliques à  $b_1 < 1$ .

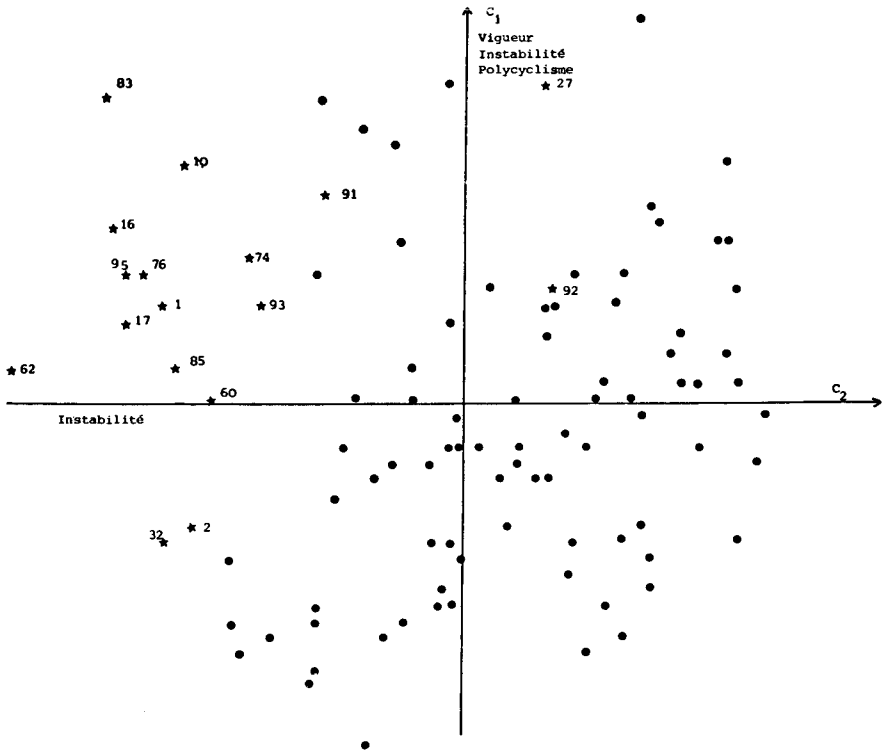


FIG. 6

*Graphe des composantes principales*

*Les familles identifiées sur ce schéma se trouvent sur la figure 1a (Saint-Alban 2.44.3.1)*

*Graph of principal components*

*Families identified on the graph are on figure 1a*

## 5. - Conclusion

Vue sous l'angle de l'interaction génotype  $\times$  année, l'étude du comportement interannuel de la croissance du pin maritime a montré qu'il obéit à des déterminismes différents selon que les milieux incriminés sont humides ou secs.

### *Saint-Alban*

L'interaction génotype  $\times$  année n'est qu'imparfaitement expliquée par le modèle linéaire, les déviations autour des droites de régression sont par ailleurs corrélées au polycyclisme : ce qui force à attribuer à la pousse secondaire la responsabilité de l'instabilité de l'accroissement annuel autour de la réponse du génotype. Enfin réponse et instabilité sont sous contrôle génétique mais leur héritabilité reste faible ( $\approx .10$ ).

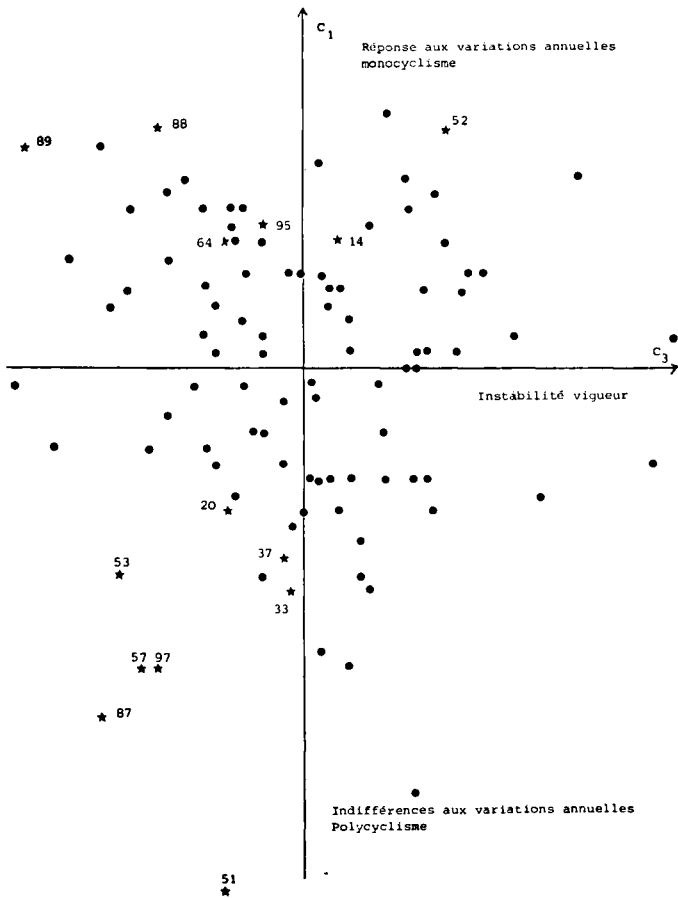


FIG. 7

*Graphe des composantes principales*

*Les familles identifiées sur ce schéma sont représentées sur la figure 2b (Soustons 2.44.3.2)*

*Graph of principal components*

*Families identified on the graph are on figure 2b*

### *Soustons*

Les familles les plus interactives sont celles dont les coefficients de régression s'écartent le plus de 1 : au niveau global, l'interaction génotype  $\times$  année est expliquée pour 50 p. 100 par l'hétérogénéité des pentes des droites de régression. L'utilisation des pentes comme critère de sélection réduirait dans des proportions importantes l'interaction  $G \times A$ . Ceci est d'autant plus vrai qu'elles sont en partie héritables ( $h^2 \simeq .20$ ). Enfin, elles sont corrélées négativement avec le polycyclisme. Celui-ci intervient donc directement dans la réponse du génotype et non dans la stabilité autour de cette réponse comme précédemment.

Ces résultats contribuent à la connaissance de la croissance en hauteur chez le pin maritime et à l'optimisation des méthodes de sélection utilisant ce caractère.

- Que ce soit dans les sols régulièrement alimentés en eau ou dans des milieux souffrant de sécheresse estivale, le polycyclisme intervient directement dans l'interaction  $G \times A$  en accentuant l'imprévisibilité des réponses dans le premier cas et en augmentant l'hétérogénéité des réponses dans le second cas. Dans le même ordre d'idées, on peut penser que le polycyclisme explique également une grande partie de l'interaction génotype  $\times$  milieu observé au niveau des moyennes d'accroissements entre les landes et les dunes.

- En ce qui concerne la sélection, les résultats de cette étude sont utilisables autant au niveau de la stratégie qu'à celui de la méthodologie.

### *Stratégie de sélection*

Jusqu'à présent la régularité de la croissance était intégrée dans la sélection sous l'angle des corrélations jeune-adulte par l'intermédiaire de plantations clonales des parents incriminés. On peut reprocher à cette méthode de ne tenir compte que de l'information familiale. A une échelle moindre dans le temps, cette méthode peut être optimisée par l'utilisation dans l'indice de sélection des paramètres du modèle linéaire estimés au niveau individuel. Le contrôle génétique partiel de ceux-ci laisse présager un gain génétique possible obtenu après sélection sur la base de ces paramètres. Enfin, il reste à préciser les critères à utiliser dans l'indice. En première approximation, dans le modèle linéaire, le génotype idéal est celui qui répond le plus régulièrement aux variations annuelles ( $b_i \simeq 1$ ). Le critère à utiliser dans ce cas serait  $|b_i - 1|$ . Il est toujours possible d'optimiser la définition du génotype idéal par l'utilisation d'un modèle curvilinaire : le choix se porterait dès lors sur celui relativement indifférent aux conditions écologiques des années pauvres, mais par contre réagissant bien aux années favorables.

### *Méthodologie de la sélection*

Le modèle linéaire pourrait être utilisé comme modèle de croissance. Il permettrait en particulier d'estimer les accroissements des arbres coupés en éclaircie dans les vergers à graines de familles et dans les tests de descendance. Enfin, pour autant que les milieux considérés soient relativement analogues (landes ou dunes), il n'est pas exclu d'utiliser l'équation de régression du génotype donné obtenu dans un milieu pour prédire ses performances dans un autre milieu, si toutefois une population de génotypes stables préalablement testés peut fournir un indice biologique de ce milieu. Pour ces applications de prédiction, il convient que les paramètres des droites de régression soient estimés avec le maximum de précision (retenir dans l'échantillonnage des années, celles aux conditions extrêmes).

*Reçu pour publication en août 1980.*

### Summary

#### *Genetic control of height growth in Pinus pinaster : II. - Interannual pattern, genotype $\times$ year interaction*

Shoot growth pattern of *Pinus pinaster* over several years is described by using a general linear model for genotype  $\times$  environment interaction analysis. Height growth was measured over 6 or 4 successive years on 100 half sib families replicated on two different sites (landes area and dunes area). Three parameters of height growth pattern are defined :

— *Interactivity*, which is the inverse of Wricke's Ecovalence.

This parameter includes both others :

- General *response* to annual conditions (coefficient b) ;
- *Stability*, around the response (coefficient d).

In the first site, the best one, family  $\times$  year interaction is mostly due to instability, that means deviations around the regression lines. But in the second, interaction is due to heterogeneity of the different responses (regression slopes). The tree parameters are estimated for each family. All the parameters are affected by family  $\times$  site interaction and especially the b one. Use of the linear model is discussed for prediction purposes, in each site and from one to the other site. Both parameters b and d are under low genetic control. Heritability of d is about .10 in the first site and the one of b is about .20 in the second site. Biological interpretation of family  $\times$  year interaction are discussed by using a multivariate analysis. Multimodality or polycyclism is responsible of family  $\times$  year interaction in the following way :

It affects d in the first site and b in the second one.

#### Key-words

*Height growth, stability, family  $\times$  year interaction, multimodality*

### Références bibliographiques

- BARADAT Ph., 1975. The juvenile mature genetic correlations. Their importance for understanding changes in physiology of forest trees and for selection. *I.U.F.R.O. Physiol. Genet. Conf.*, 1975, Edimburg.
- BUCIO ALANIS, 1966. Environmental and genotype environmental components of variability. I. Inbred Lines. *Heredity*, **21**, 387-397.
- BRESE E.L., 1969. The measurement and significance of genotype environment interactions in grasses. *Heredity*, **24**, 27-44.
- BURDON R., 1977. Genetic correlation as a concept for studying genotype environment interaction in forest tree breeding. *Silvae Genet.*, **26**, 5-6, 168-175.
- DICKERSON G.E., 1952. Implications of genetic-environmental interaction in animal breeding. *Anim. Prod.*, **4**, 47-64.
- FINLAY R.W., WILKINSON G.N., 1963. The analysis of adaptation in a plant breeding programme. *Aust. J. agr. Res.*, **14**, 792-754.
- EBERHART S.A., RUSSELL W.A., 1966. Stability parameters for comparing varieties. *Crop Sci.*, **6**, 36-40.
- GOSLEE D.H., LUCAS H.L., 1965. Analysis of variance of disproportionate data when interaction is present. *Biometrics*, **21** (1), 115-133.

- HINKELMANN K., 1974. Genotype-environment interaction. Aspects of statistical design analysis and interpretation. *Int. Biom. Conf.*, 1974, Constanta, Roumanie.
- KREMER A., 1981. Déterminisme génétique de la croissance en hauteur du Pin maritime (*Pinus pinaster* Ait.). I. Rôle du polycyclisme). *Ann. Sci. for.*, **38** (2).
- MANDEL J., 1969. A method for fitting empirical surfaces to physical or chemical data. *Technometrics*, **11** (3), 411-429.
- MOAV R., WOHLFARTH G.W., 1974. Magnification through competition of genetic differences in yield capacity in carp. *Heredity*, **33** (2), 18-202.
- MOAV R., HULATA G., WOHLFARTH G., 1975. Genetic differences between the chinese and european races of the common carp. I Analysis of genotype environment interactions for growth rate. *Heredity*, **34** (3), 323-340.
- PARODA R.S., HAYES J.D., 1971. An investigation of genotype  $\times$  environment interactions for rate of ear emergence in spring barley. *Heredity*, **26**, 157-175.
- ROBERTSON A., 1959. The sampling variance of the genetic correlation coefficient. *Biometrics*, **15**, 469-485.
- SHUKLA G.K., 1972. Some statistical aspects of partitioning genotype environment components of variability. *Heredity*, **29**, 237-245.
- WRICKE G., 1962. Über eine methode zur Erfassung der Oekologischen Streubreite in Feldversuchen. *Z. Pflanzenschutz*, **47**, 92-96.
- WRIGHT A.J., 1971. The analysis and prediction of some two factor interactions in grass breeding. *J. agri. Sci.*, **76**, 301-306.
- WRIGHT A.J., 1976. The significance for breeding of linear regression analysis of genotype environment interactions. *Heredity*, **37** (1), 83-93.