

Application de l'analyse factorielle à l'étude de la croissance en hauteur des arbres : exemple du Pin maritime

B. LEMOINE

Avec la collaboration technique de A. SARTOLOU

I.N.R.A.,

*Laboratoire de sylviculture et d'écologie de la pinède landaise,
Domaine de l'Hermitage, Pierroton, 33610 Cestas*

Résumé

On cherche à obtenir pour les peuplements de Pin maritime :

— d'une part une description complète et structurée de la variabilité de la croissance ;
— d'autre part un mode de construction des courbes de croissance par l'utilisation de paramètres non corrélés entre eux.

Cette étude est entreprise dans un double but d'explication des phénomènes (causes génécologiques) et de prévision (Tables de production) de la croissance.

On définit ainsi le phénomène de croissance étudié :

« La croissance d'un arbre, ou d'un peuplement, est l'ensemble des accroissements qui se succèdent pendant tout ou partie de la vie de cet arbre, ou de ce peuplement. »

Le matériel étudié est composé d'échantillons d'arbres dominants à la coupe rase. La technique de mesure est l'analyse de tiges.

L'outil mathématique utilisé est l'analyse en composantes principales. Les variables initiales sont les accroissements quinquennaux successifs X_i .

1. - *Matrice de corrélations entre les variables initiales*

Cette matrice présente une structure très particulière. Elle ressemble par certains points à une matrice de corrélations induite — on suppose que la corrélation entre 2 accroissements X_i et X_{i+q} séparés par $5q$ années doit décroître constamment lorsque q augmente, ce qui est vrai si q est assez petit — et en diffère par d'autres — si q est assez grand la corrélation observée est de 0,5 quel que soit q —. On suppose que la corrélation observée est en fait la combinaison de 2 corrélations, l'une constante, traduisant certains effets du milieu, l'autre constamment décroissante lorsque q augmente, traduisant la diminution avec le temps du contenu d'une mémoire de l'arbre.

2. - *L'analyse en composantes principales Y_j*

On est amené à retenir trois composantes principales expliquant 90 p. 100 de la variabilité totale.

La première composante Y_1 s'explique en grande partie par le niveau d'alimentation en eau. Elle donne naissance à des courbes-types d'accroissements superposées pendant toute la vie des peuplements.

Les autres composantes Y_2 et Y_3 n'ont pas encore reçu d'explication biologique. Elles donnent naissance à des courbes-types d'accroissements qui se croisent — respectivement à 25 ans, et à 12 ans et 40 ans —. On envisage l'effet de divers facteurs, soit environnementaux, soit génétiques, mais entrant en interaction avec un facteur physiologique : « le vieillissement de l'arbre serait le fait non seulement de l'âge, mais aussi de l'intensité des croissances antérieures ». Ces modalités particulières de croissance auraient pu être mises en œuvre par la sylviculture.

3. - Conclusion

Les résultats sont de deux ordres :

- les résultats applicables concernant le mode d'action du facteur eau et la construction optimisée de Tables de production ;
- les résultats d'ordre prospectif incitant à poursuivre l'effet théorique et expérimental.

1. - But du travail

On désire obtenir une *description* complète et structurée de la variabilité des courbes de croissance en hauteur du Pin maritime quand on passe *d'un peuplement à un autre*.

On cherche à construire des courbes de croissances caractérisées par des *paramètres non corrélés* entre eux.

L'intérêt de ce genre de paramètres est double :

— au plan de *l'explication de la croissance* : les modes d'actions des divers facteurs génécologiques sur les courbes de croissance peuvent être différents ;

— au plan de la *prévision* (Tables de production) : selon l'importance de la variabilité des paramètres et la précision recherchée on pourra choisir, soit de construire un *faisceau complexe* à plus d'un paramètre, soit *d'optimiser* la construction d'un *faisceau simple* en retenant *objectivement* le seul paramètre dont la variabilité conditionne la plus grande part de la variabilité des courbes de croissance ; le résidu statistique obtenu dans le second cas est non seulement chiffré, mais bien caractérisé quant à son mode d'action sur l'erreur de prévision.

Cet objectif n'est pas atteint par l'emploi de *fonctions de croissance* du genre de celle de MITSCHERLICH (cf. PRODAN, 1968) : les variations des paramètres obtenues par le calcul sont corrélées entre elles. La méthode biométrique essayée se voudra analytique ; le choix du *degré d'analyse* devra être pertinent ce qui nous amène à définir le *phénomène de croissance* étudié :

« La croissance d'un arbre, ou d'un peuplement, est l'ensemble des accroissements qui se succèdent pendant tout ou partie de la vie de cet arbre, ou de ce peuplement. »

Les *caractères biométriques* retenus sont donc les *accroissements successifs* ; l'étude de la variabilité devra tenir compte de faits d'origine biologique et qui sont l'importance des *corrélations entre caractères*.

Une étude directe de l'ensemble des *états successifs*, ou hauteurs ou encore croissances cumulées, *fausserait la perspective* : on aurait bien évidemment affaire à des « corrélations de redondance » plus qu'à des corrélations biologiques.

2. - Matériel et méthodes

2.1. - L'analyse des données

Chaque *accroissement courant* de la vie de l'arbre ou du peuplement est considéré comme une *variable*.

L'ensemble des données, n observations \times i variables, est traité par *l'analyse multivariée*.

Cette méthode a déjà été appliquée à d'autres problèmes de croissances par :

— BARON (†) & MILLIER (travaux non publiés) aux croissances cumulées du lapin d'une part, de la poule d'autre part ;

— BUIS (1966) à la croissance de l'hypocotyle de lupin ;

— BIROT & JOANNES (1973) aux accroissements hebdomadaires d'individus appartenant à 11 provenances de Douglas.

Nous choisissons ici *l'analyse en composantes principales* pour deux raisons :

— elle est apte selon nous à mettre en évidence tout à la fois les caractères de *vigueur générale* et de forme des *courbes de croissance* ;

— elle suffit à traiter le problème posé puisque nous n'avons affaire ici qu'à un *seul niveau de variabilité* (entre peuplements) ; dans le cas contraire (variabilité entre individus d'un même peuplement et variabilité entre peuplements) il est nécessaire de recourir à l'analyse discriminante.

L'outil mathématique choisi nous paraît apte à remplir les objectifs fixés, en particulier :

— les paramètres non corrélés vont naître spontanément sous forme de composantes ;

— la construction des types de croissance est possible par l'emploi des régressions des variables sur les composantes.

Les variables étudiées sont les accroissements quinquennaux en hauteur h :

$$X_i = h_{5i} - h_{5(i-1)}$$

où 5_i et $5_{(i-1)}$ sont deux âges A distants de 5 ans.

Les composantes principales obtenues sont nommées Y_j .

2.2. - Méthodes dendrométriques

On a procédé à deux types d'analyse de tiges :

— une analyse classique avec lecture des âges sur rondelles (R) prises tous les 2 m dans le bas et le milieu de l'arbre et tous les mètres dans le haut ; les *accroissements quinquennaux* successifs sont obtenus par interpolation linéaire sur la courbe de croissance expérimentale ;

— une analyse « économe » sur des arbres dont on connaissait la hauteur à 20 et 30 ans prise au dendromètre : les hauteurs à 5 et 10 ans ont été obtenues par carottage (K) ; on a vérifié que cette méthode était, d'une part sans biais, d'autre part suffisamment précise pour la détermination des moyennes des accroissements de quelques arbres.

2.3. - *Echantillons*

Les peuplements étudiés sont au nombre de 40 dont :

— 33 peuplements nés entre 1897 et 1929 (R) : on obtient pour chacun d'eux une *courbe* représentant la croissance moyenne de 10 à 15 arbres *dominants* à la coupe rase ;

— 7 peuplements nés en 1937 ou 1947 (K) : on obtient pour chacun d'eux une *courbe* représentant la croissance moyenne de 3 arbres *dominants* au moment de la mesure.

2.4.- *Situation écologique des peuplements*

Les 40 peuplements étudiés se répartissent en :

— 11 stations de *dunes* (D) ;

— 29 stations de *landes* (L.).

Les stations de landes se répartissent en 4 classes de *pédoclimat* appréhendées par la *végétation* (cf. LEMOINE, 1969) :

— 5 stations de lande humide (Lh) ;

— 6 stations de lande mésophile humide (Lmh) ;

— 6 stations de lande mésophile sèche (Lms) ;

— 8 stations de lande sèche (Ls).

Les études écophysiologicals (accroissements mensuels en circonférence) montrent que l'ensemble des peuplements de dunes est plus sensible à une sécheresse que l'ensemble des peuplements de landes (cf. LEMOINE, 1979). On peut donc généraliser l'échelle pédoclimatique de la façon suivante :

—	<i>Alimentation en eau</i>				+
	<i>Landes</i>				
	Ls	Lms	Lmh	Lh	
<i>Dunes</i>					

2.5. - *Calculs effectués*

On a procédé à 3 analyses en composantes principales (soit « ACP ») :

— ACP 01 : étude de la croissance de 0 à 70 ans sur 17 observations *peuplements* ; d'où 14 variables X_i ;

— ACP 02 : étude de la croissance de 0 à 50 ans sur 33 observations *peuplements* ; d'où 10 variables X_i ;

— ACP 03 : étude de la croissance de 0 à 30 ans sur 40 observations *peuplements* ; d'où 4 variables X_i .

3. - **Résultats**3.1. - *Corrélations entre les accroissements*

Les matrices de corrélations entre les accroissements quinquennaux X_i sont données dans le tableau 1.

TABLEAU 1

Matrice des coefficients de corrélation entre les accroissements quinquennaux
 Matrix of correlations between five years increments

$$X_i = h_{5i, \text{ans}} - h_{5(i-1), \text{ans}}$$

A) ACP 01 : landes + dunes jusqu'à 70 ans
 17 observations ; chacune = moyenne des dominants d'une placette

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1													
2	,70	1												
3	,60	,94	1											
4	,45	,70	,79	1										
5	,43	,62	,75	,90	1									
6	,17	,33	,47	,58	,67	1								
7	,23	,36	,52	,64	,75	,97	1							
8	,26	,38	,55	,59	,64	,90	,92	1						
9	,32	,31	,47	,47	,65	,79	,84	,91	1					
10	,46	,48	,58	,58	,70	,74	,81	,85	,91	1				
11	,28	,39	,50	,69	,74	,63	,71	,73	,73	,85	1			
12	,35	,33	,46	,56	,68	,55	,62	,70	,76	,81	,93	1		
13	,32	,35	,44	,64	,66	,55	,58	,57	,57	,68	,90	,88	1	
14	,52	,59	,64	,76	,76	,57	,61	,65	,65	,75	,90	,89	,93	1

Seuils de signification : 5 % * ,48.
 1 % ** ,61.

B) ACP 02 : landes + dunes jusqu'à 50 ans
33 observations ; chacune = moyenne des dominants d'une placette

i	1	2	3	4	5	6	7	8	9	10
1	1									
2	,74	1								
3	,53	,89	1							
4	,46	,75	,87	1						
5	,39	,64	,78	,86	1					
6	,26	,45	,58	,65	,78	1				
7	,30	,47	,61	,70	,72	,89	1			
8	,31	,47	,60	,63	,64	,74	,86	1		
9	,40	,38	,41	,38	,38	,48	,65	,81	1	
10	,53	,52	,51	,48	,40	,45	,65	,75	,92	1

Seuils de signification : 5 % * ,35.
1 % ** ,45.

C) ACP 03 : 40 observations ; chacune = moyenne des dominants d'une placette

i	1	2	3	4	5	6
1	1					
2	,77	1				
3	,31	,68	1			
4						
5	,34	,51	,64	1		
6						

Seuils de signification : 5 % * ,31
1 % ** ,40

On remarque que la corrélation entre deux accroissements X_i et X_{i+q} est influencée par le temps, soit $5q$ ans, qui sépare ces accroissements. Par exemple (ACP 01) la corrélation entre X_i et X_{i+1} (diagonale supérieure de la matrice) est comprise entre 0,67 et 0,93, la corrélation entre X_i et X_{i+9} (9^e diagonale en partant du haut de la matrice) est comprise entre 0,39 et 0,76.

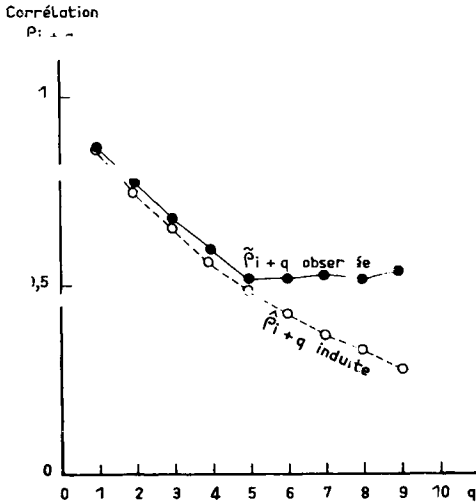


FIG. 1

Corrélations moyennes, observée et induite, entre les accroissements X_i et X_{i+q} distants de 5 ans. Evolution avec q (ACP 01)

Observed and induced average correlations between 5 years distant increments X_i and X_{i+q} . Following q evolution (ACP 01)

Afin de préciser la structure de cette matrice nous procédons empiriquement en effectuant les moyennes des coefficients de chaque diagonale ; nous obtenons ainsi pour chaque différence d'âge $5q$ ans la corrélation \tilde{Q}_q entre les accroissements X_i et X_{i+q} (quel que soit l'âge 5_i), ainsi :

$$\tilde{Q}_1 = \frac{1}{13} [,70 + ,94 + \dots + ,93] = ,869$$

$$\tilde{Q}_2 = \frac{1}{12} [,60 + ,70 + \dots + ,89] = ,775$$

$$\tilde{Q}_9 = \frac{1}{5} [,46 + ,39 + \dots + ,76] = ,542$$

Ces résultats sont reportés sur la figure n° 1.

On remarque ainsi que la corrélation \tilde{Q}_q :

- décroît régulièrement de $q = 1$ à $q = 5$;
- reste pratiquement constante et voisine de 0,526 pour $q > 5$.

Autrement dit la corrélation \tilde{Q}_q entre deux accroissements X_i et X_{i+q} dépend de la différence d'âge entre ces deux accroissements quand celle-ci n'excède pas $5q = 25$ ans ; pour des valeurs supérieures à 25 ans cette corrélation est constante.

On rapproche cette corrélation observée \tilde{Q}_q d'une *corrélation induite* \hat{Q}_q à partir de la corrélation \tilde{Q}_1 en envisageant que la corrélation \hat{Q}_q diminue régulièrement quand q augmente — c'est-à-dire quand l'éloignement dans le temps des deux accroissements, ou différence d'âge, augmente — et tend vers 0, quand q tend vers l' ∞ . On procède empiriquement de la façon suivante :

$$\begin{aligned}\hat{Q}_1 &= \tilde{Q}_1^1 \\ \hat{Q}_2 &= \tilde{Q}_1^2 \\ \hat{Q}_3 &= \tilde{Q}_1^3 \\ \hat{Q}_q &= \tilde{Q}_1^q\end{aligned}$$

On observe ainsi sur la figure 1 que les valeurs de \hat{Q}_q et \tilde{Q}_q :

- sont très voisines pour $q \leq 5$;
- suivent des lois différentes quand q s'accroît à partir de la valeur $q = 5$.

Autrement dit la matrice de corrélation observée ressemble par certains points à une matrice de corrélations induites et en diffère par certains autres.

3.2. - Structure de la variabilité

Le tableau 2 donne, rangés par ordre décroissant, les pourcentages d'explication de la variation totale par chacune des composantes Y_j .

On remarque que la première composante explique les 2/3 de la variation et qu'il suffit de 2 ou 3 composantes pour expliquer 90 p. 100 de cette variation.

TABLEAU 2

Valeurs propres des composantes Y_j en p. 100 de la variation totale
Eigen values of principal components Y_j in p. 100 of total variation

ACP \ j	1	2	3	4	5	etc.
01	66,4	13,6	8,3	5,2	2,2	
02	63,9	14,2	11,8	4,0	1,9	
03	66,1	21,2	10,1	2,6		

Nota : 66,4, etc. : composantes retenues (90 % cumulés de la variance totale).

3.3. - Situation des observations dans l'espace des composantes

On étudie les projections des observations sur les plans des composantes Y_1 , Y_2 et Y_1 , Y_3 .

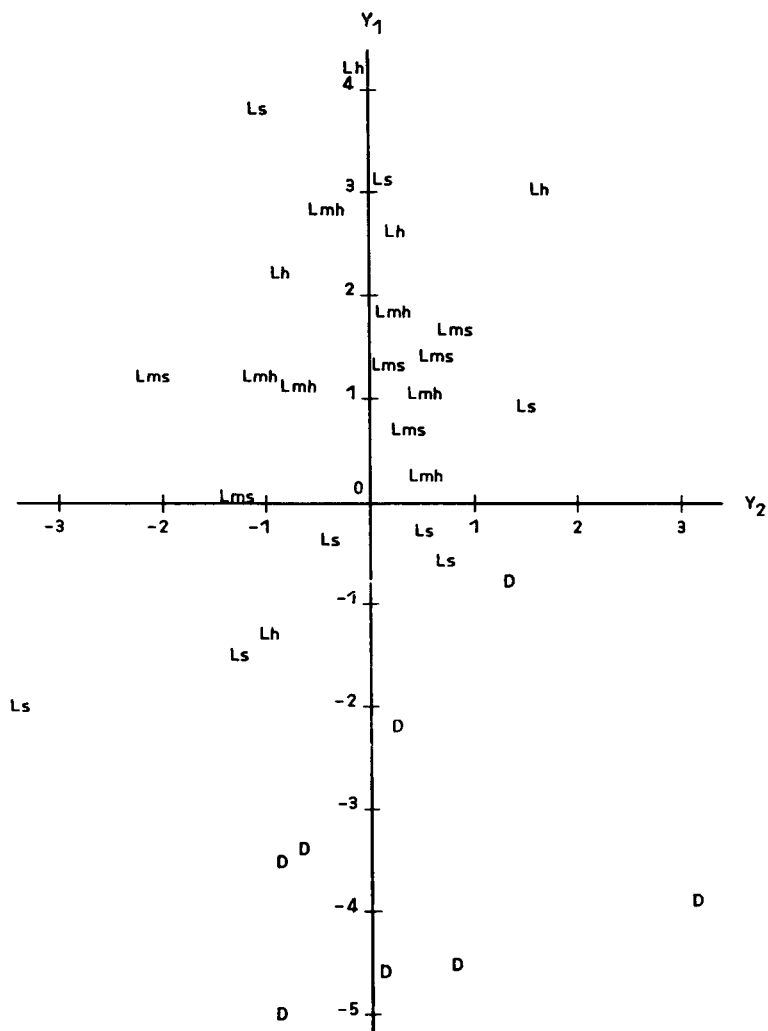


FIG. 2 — ACP 02

Projections des observations sur le plan des composantes Y_1 , Y_2 ,
Observations projections upon principal components Y_1 , Y_2 plan

Légende : Lh : landes humides,
Lmh : landes mésophiles humides,
Lms : landes mésophiles sèches,
Ls : landes sèches.
D : dunes.

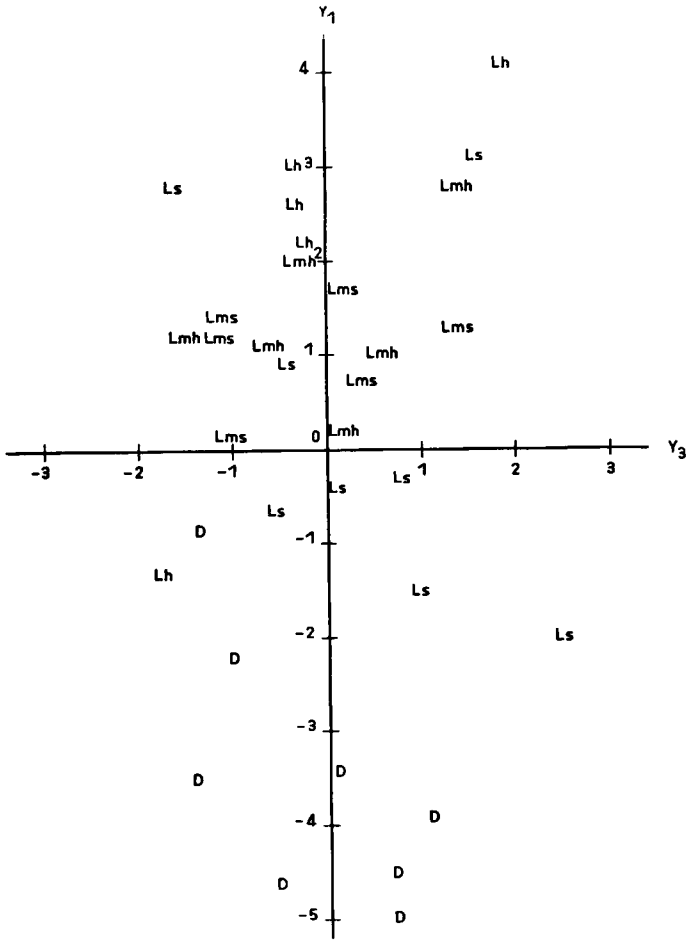


FIG. 2 bis — ACP 02

*Projections des observations sur le plan des composantes Y_1, Y_3 ,
Observations projections upon principal components Y_1, Y_3 plan*

- Légende :** Lh : landes humides,
 Lmh : landes mésophiles humides,
 Lms : landes mésophiles sèches,
 Ls : landes sèches,
 D : dunes.

Les figures 2 et 2 bis et 3 permettent de dégager un point commun aux 3 analyses : l'axe principal « croissance » s'identifie assez bien avec l'échelle pédoclimatique. Il existe cependant quelques « bavures ».

— un peuplement de lande humide (Lh) à $Y_1 < 0$: l'explication se trouve en l'excès d'eau car le site est submergé en hiver ;

— deux peuplements de lande sèche (Ls) à $3 < Y_1 < 4$, pour lesquels on ne trouve pas d'explication.

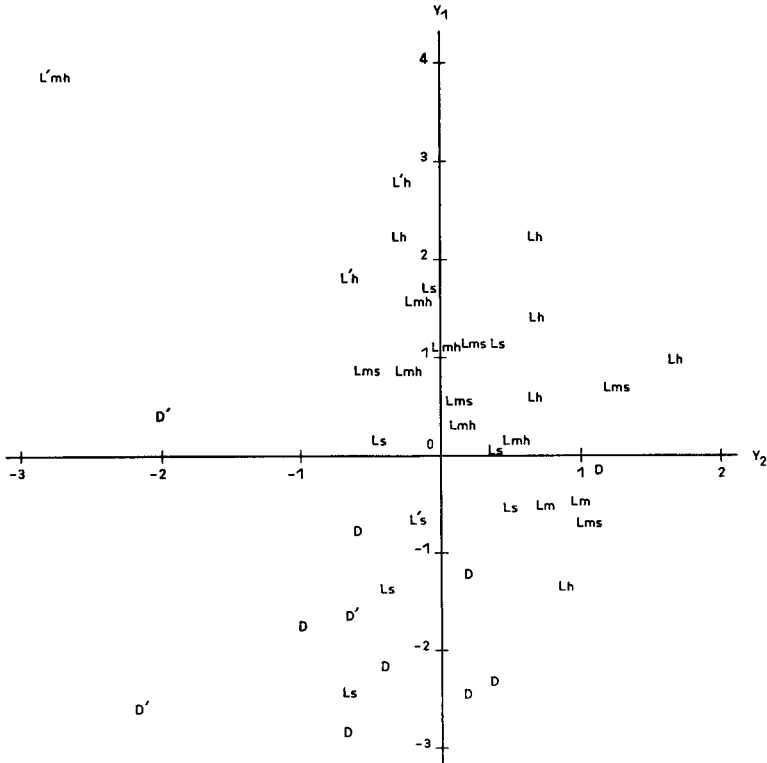


FIG 3 — ACP 03

*Projections des observations sur le plan des composantes Y_1, Y_2
Observations projections upon principal components Y_1, Y_2 plan*

- Légende :
- Lh et L'h : landes humides,
 - Lmh et L'mh : landes mésophiles humides,
 - Lms : landes mésophiles sèches,
 - Ls et L's : landes sèches,
 - D et D' : dunes,
 - L et D : analyse de tiges par rondelles des placettes régénérées entre 1897 et 1929,
 - L' et D' : analyse de tiges par carottes et dendromètres des placettes régénérées en 1939 et 1947.

Le deuxième axe Y_2 de l'ACP 03 (cf. figure 3) permet de distinguer :

- d'une part le groupe des 33 peuplements nés entre 1897 et 1929 ;
- d'autre part le groupe des 7 peuplements nés en 1937 et 1947.

Les autres axes Y_2 et Y_3 des ACP 01 et ACP 02 n'ont aucun rapport avec les groupes constitués selon les critères de pédoclimat et d'année de naissance.

3.4. - Typologie des croissances

On examine les corrélations (cf. tableau 3) et les régressions (cf. tableau 4) des variables X_i en fonction des composantes Y_j .

TABLEAU 3

Coefficients de corrélation entre variable X_i et composantes Y_j
Correlation coefficients between variable X_i and principal components Y_j

A) ACP 01

j	i	1	2	3	4	5	6	7	8	9	10
1		,59	,78	,86	,86	,84	,87	,87	,86	,72	,77
2		,32	,49	,40	,33	,25	-,13	-,27	-,40	-,57	-,41
3		,60	,28	,00	-,18	-,35	-,45	-,29	-,10	,32	,43

B) ACP 02

j	i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1		,50	,63	,75	,82	,88	,80	,86	,86	,84	,91	,90	,85	,81	,90
2		,62	,71	,56	,33	,17	-,35	-,32	-,32	-,35	-,17	-,19	-,19	-,10	,11
3		,02	-,17	-,23	-,03	-,06	-,40	-,35	-,30	-,18	-,04	,34	,41	,51	,37

C) ACP 03

i	1	2	3	4	5	6
1	,74	,92	,82		,76	
2	-,63	-,24	-,42		,46	

TABLEAU 4

Coefficient des régressions $\hat{X}_i = \beta_0 + \beta_1 \cdot Y_1 + \beta_2 \cdot Y_2 + \beta_3 \cdot Y_3$
des variables X_i sur les composantes Y_j

Coefficients of the regressions $\hat{X}_i = \beta_0 + \beta_1 \cdot Y_1 + \beta_2 \cdot Y_2 + \beta_3 \cdot Y_3$
(variables X_i upon principal components Y_j)

A) ACP 01

β_j	i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
β_0		2,01	2,72	2,94	2,88	2,68	2,29	1,87	1,54	1,25	1,05	,92	,87	,78	,73
β_1		,052	,105	,132	,131	,109	,103	,110	,085	,072	,062	,059	,062	,051	,047
β_2		,143	,264	,218	,117	,046	-,099	-,089	-,070	-,066	-,025	-,027	-,031	-,014	,012
β_3		,005	-,082	-,113	-,013	-,019	-,143	-,127	-,083	-,042	-,009	,062	,084	,089	,055

B) ACP 02

β_j	i	1	2	3	4	5	6	7	8	9	10
β_0		1,95	2,65	2,98	2,91	2,65	2,27	1,93	1,60	1,28	1,09
β_1		,073	,161	,190	,167	,133	,111	,124	,105	,080	,073
β_2		,084	,213	,188	,137	,083	-,038	-,083	-,103	-,135	-,082
β_3		,172	,136	-,001	-,081	-,127	-,145	-,096	-,026	,083	,095

C) ACP 03

β_j	i	1	2	3	4	5	6
β_0		2,07	2,72	2,87		2,45	
β_1		,217	,333	,286		,189	
β_2		-,325	-,155	,260		,188	

Chaque variable X_i est ajustée aux composantes Y_1, Y_2, Y_3 par une régression multiple qui lui est propre. L'estimation de chaque coefficient de régression $\beta_{i,j}$ est excessivement simple puisque les composantes Y_j sont non corrélées entre elles (cf. TOMASSONE, 1966 ; mais ici les variables expliquées sont les variables initiales) :

$$\beta_{i,j} = \frac{\sum X_i \cdot Y_j}{\sum Y_j^2} = r_{i,j} \cdot \sqrt{\frac{\sigma^2 X_i}{\sigma^2 Y_j}}$$

où $\sigma^2 Y_j$ est la valeur propre correspondant à la composante Y_j . Nous avons ainsi :

$$\hat{X}_1 = \beta_{1,0} + \beta_{1,1} \cdot Y_1 + \beta_{1,2} \cdot Y_2 + \beta_{1,3} \cdot Y_3$$

$$\hat{X}_2 = \beta_{2,0} + \beta_{2,1} \cdot Y_1 + \beta_{2,2} \cdot Y_2 + \beta_{2,3} \cdot Y_3$$

$$\hat{X}_i = \beta_{i,0} + \beta_{i,1} \cdot Y_1 + \beta_{i,2} \cdot Y_2 + \beta_{i,3} \cdot Y_3$$

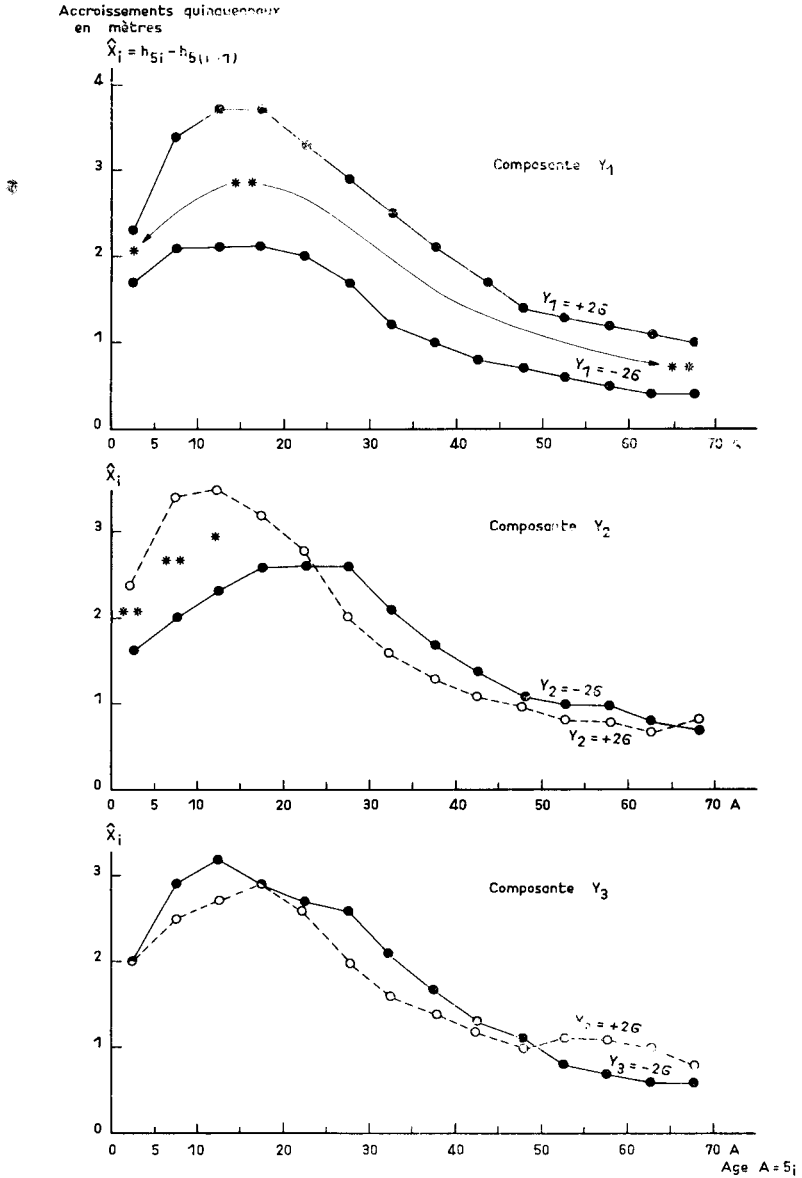


FIG. 4 — ACP 01

Courbes de croissance extrêmes résultant des ajustements des variables X_i à la composante Y_j
 Coefficients de corrélation entre X_i et Y_j :

- * : $0,48 \leq |r| < 0,61$.
- ** : $0,61 \leq |r|$.

Growth extreme curves resulting from variables X_i upon principal component Y_j fit
 Correlation coefficients between X_i and Y_j :

- * : $0,48 \leq |r| < 0,61$.
- ** : $0,61 \leq |r|$.

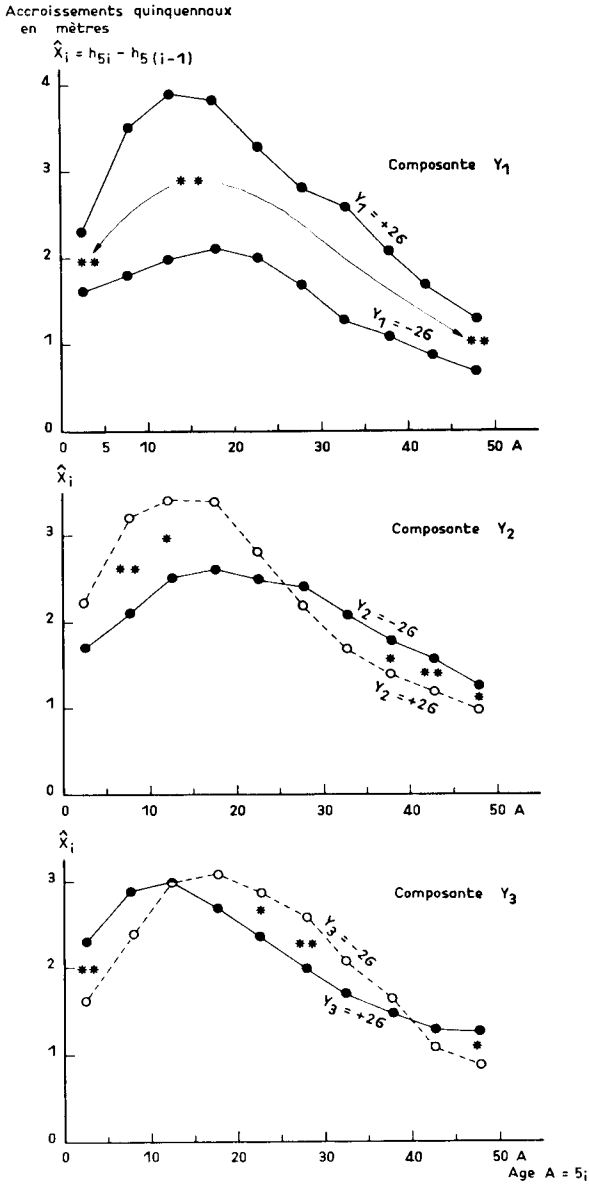


FIG. 5 — ACP 02

Courbes de croissance extrêmes résultant des ajustements des variables X_i à la composante Y_j
 Coefficients de corrélation entre X_i et Y_j :

* : $0,35 \leq |r| < 0,45$.

** : $0,45 \leq |r|$.

Growth extreme curves resulting from variables X_i upon principal component Y_j fit
 Correlation coefficients between X_i and Y_j :

* : $0,35 \leq |r| < 0,45$.

** : $0,45 \leq |r|$.

On détermine 2 types extrêmes de croissance ayant trait à la variation d'une composante en utilisant alternativement les valeurs extrêmes $+2\sigma$ et -2σ de cette dernière et en annulant les valeurs des deux autres composantes. Les 2 séquences des \hat{X}_i selon l'âge permettent de construire les 2 types de courbes désirés.

A une exception près, la composante Y_3 de l'ACP 01 qui semble caractériser la croissance finale, et dès que l'on utilise un échantillonnage suffisant de peuplements, on constate les faits suivants (cf. figures 4, 5 et 6) :

— la présence (corrélation ou régression) d'une composante dans une ou plusieurs variables X_i pendant une période de la vie du peuplement *n'exclut pas* la présence d'une autre composante dans cette ou ces variables pendant la même période ;

— il existe 2 modes de passage d'un type extrême de croissance à l'autre :

- composante Y_1 : les variables successives X_i sont toutes corrélées positivement avec elle ; il en résulte 2 courbes de croissance constamment superposées ;

- composante Y_2 ou Y_3 : les premières variables X_i sont corrélées positivement (ou négativement) avec elle, les dernières négativement (ou positivement) ; il en résulte 2 courbes de croissance qui se croisent.

Accroissements quinquennaux
en mètres

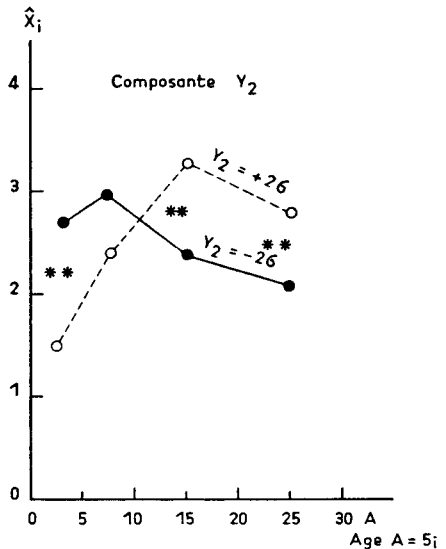
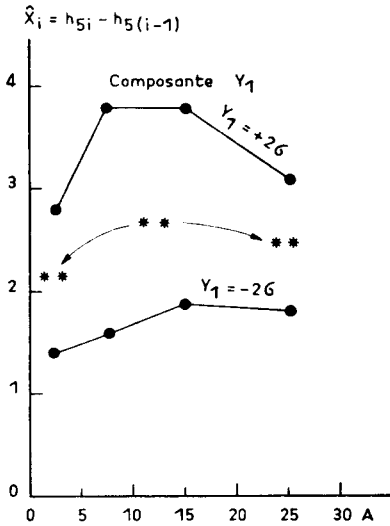


FIG. 6 — ACP 03

Courbes de croissance extrêmes résultant des ajustements des variables X_i à la composante Y_j
Coefficients de corrélation entre X_i et Y_j :

* : $0,31 \leq |r| < 0,40$.

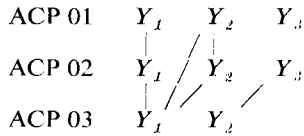
** : $0,40 \leq |r|$.

Growth extreme curves resulting from variables X_i upon principal component Y_j fit
Correlation coefficients between X_i and Y_j :

* : $0,31 \leq |r| < 0,40$.

** : $0,40 \leq |r|$.

La parenté des composantes quand on passe d'une analyse à l'autre est illustrée par le schéma suivant :



le passage à l'ACP 03, où on étudie la croissance de 0 à 30 ans seulement, regroupe 2 composantes en une seule ; ceci s'explique aisément par le fait que pour la composante Y_2 des ACP 01 et ACP 02 les 2 courbes types sont superposées jusqu'à 25 ans, d'où une ressemblance avec l'action de la composante Y_1 de ces deux premières analyses.

Les répercussions de cette typologie sur les *courbes de croissances cumulées* $h = f(A)$ sont les suivantes (ACP 02) :

— composante Y_1 : les rapports successifs des hauteurs des 2 types extrêmes de croissances restent compris dans une fourchette assez étroite (entre 1,71 et 1,82) de 10 à 50 ans : ces 2 courbes de croissance présentent donc un caractère très accusé d'*affinité orthogonale* ;

— composantes Y_2 et Y_3 : les *hauteurs à 50 ans sont voisines, voire identiques*, d'un type extrême de croissance à l'autre ; ainsi pour la variation de Y_2 : 22,5 m et 20,6 m et pour la variation de Y_3 : 21,1 m et 21,2 m.

4. - Discussion et conclusion

L'analyse en *composantes principales* d'un tableau de données à n éléments et i variables peut être vue comme une méthode d'analyse d'une *matrice de corrélations*. Il n'en reste cependant pas moins souhaitable d'examiner directement la structure de telles matrices comme nous l'avons fait au § 3.1. Notre interprétation a trait à ces *deux stades du traitement des données*.

4.1. - La matrice de corrélation

La figure 1 montre qu'elle possède une structure très particulière : elle ressemble par certains points à une matrice de corrélation induite, et en diffère par certains autres. On peut alors se poser la question suivante : les corrélations observées ne sont-elles pas la *combinaison de deux corrélations* d'origines biologiques différentes ?

La première de ces corrélations reflétant la liaison entre deux accroissements X_i et X_{i+q} , séparés par $5q$ années, *serait constante*, c'est-à-dire indépendante de q . C'est ainsi que peuvent se présenter les modalités d'action de certaines caractéristiques du milieu comme la fertilité des sols (richesse minérale, alimentation en eau) : leur action sur la croissance s'exerçant *pendant toute la vie du peuplement* il en résulte pour deux accroissements le même degré de ressemblance quels que soient leurs âges respectifs.

La deuxième de ces corrélations serait *d'autant plus faible* que la différence d'âge entre les deux accroissements serait plus grande. Tout se passe comme si le peuplement d'arbres était pourvu d'une « *mémoire* » enregistrant certaines actions et que le contenu de cette mémoire allait en *s'amenuisant* au fur et à mesure que se déroule le temps ($5q$) entre le premier accroissement considéré X_i (quel que soit l'âge $5q$) et un accroissement ultérieur X_{i+q} . Nous ne sommes pas en mesure actuellement de proposer d'hypothèse concernant la nature des facteurs concernés.

Le seul examen de la matrice de corrélation appelle donc des développements de recherches aussi bien au plan théorique que concret.

4.2. - *L'analyse en composantes principales*

L'analyse multivariée effectuée ici sur l'ensemble des accroissements successifs de la vie ou d'une fraction de la vie du peuplement met en œuvre principalement *trois composantes* significativement corrélées avec tout ou partie des *variables initiales*.

4.21. - *Composante Y_1 ou axe principal des croissances*

On a pu en trouver une large part d'explication par le *facteur alimentation en eau*. On craint qu'il soit difficile d'affiner l'explication de cette composante par ce facteur car la caractérisation de multiples sites à l'aide de profils hydriques est trop coûteuse ; de même la confrontation du *facteur minéral* avec cette composante (ou les autres) est impossible car dans les sols des Landes de Gascogne la teneur en certains éléments minéraux (P et K) est du même ordre de grandeur que l'erreur d'analyse chimique. Les *courbes d'accroissements paramétrées* par cette composante (cf. figures 4 et 5) sont bien distinctes pendant toute la vie des peuplements. Le facteur alimentation en eau *agit donc en permanence et avec la même intensité* tout au long du processus de croissance. On rapprochera cette conclusion de celle du § 4.1 : il existerait pour une certaine corrélation entre deux accroissements X_i et X_{i+q} une seule valeur pour toutes les différences d'âges $5q$ entre ces deux accroissements ; la valeur unique de cette corrélation serait due à la permanence et à la constance de l'effet de certaines caractéristiques du milieu.

On retiendra enfin sur un plan plus pratique que l'échelle pédoclimatique donne naissance à un faisceau de courbes cumulées $h = f(A)$ se déduisant les unes des autres par affinité orthogonale.

4.22. - *Autres composantes Y_2 et Y_3*

Elles ont une part notable (cf. ACP 02), le 1/4 à elles deux, dans la description de la variabilité de la croissance. Leur *origine biologique*, si tant est qu'il en existe une, n'est *pas identifiée*.

La manifestation de ces deux composantes à travers la seule analyse biométrique ne peut donc amener à conclure sur cette recherche.

L'intérêt de ce résultat de l'analyse ne peut donc résider que dans la possibilité de formuler des *hypothèses de travail*.

L'une ou l'autre de ces deux composantes génère des *types extrêmes de courbes de croissance* (accroissements) *qui se croisent* (une fois pour la variation de Y_2 , deux fois pour la variation de Y_3). Les facteurs biologiques envisagés par l'hypothèse devront être reconnus, au plan théorique, comme capables d'engendrer par leurs variations ces modifications très particulières du processus de croissance. Ces facteurs peuvent être soit *environnementaux*, soit *génétiques*. Leurs variations peuvent avoir pris naissance soit *dans l'espace* (à différentes échelles), soit *dans le temps*.

4.221. - Les facteurs environnementaux

Une *modification possible du climat* « normal » au début du siècle et pendant deux ou trois décennies serait parfaitement capable de donner naissance à la composante Y_2 de l'ACP 02 : des peuplements observés nés entre 1897 et 1929 les plus jeunes auraient subi cette modification climatique entre 0 et 25 ans, les plus vieux entre 25 et 50 ans ; on vérifie aisément que les courbes de croissance des plus récents et des plus anciens se croiseraient à 25 ans. Le modèle d'action du climat et de la station, puisqu'ayant trait au même facteur, l'eau, serait cumulatif, donc du type le plus simple. Or il n'existe pas de liaison entre la variation de Y_2 et la date de naissance du peuplement : nous ne retiendrons pas cette hypothèse.

Parmi les *facteurs édaphiques* nécessaires à la croissance on peut en concevoir un, ou plusieurs, qui serait, selon la station, plus ou moins *rapidement accessible* au peuplement d'arbres par le processus de *croissance racinaire* ; cette modalité serait peut-être générale ou simplement particulière aux *podzols* des Landes de Gascogne. Si on se place par exemple au niveau de la composante Y_2 de l'ACP 02, les variations stationnelles de ce type de facteur occasionneraient des *croissances plus ou moins fortes* de 0 à 20 ans avec comme extrêmes les deux courbes-types de la figure 5 : dès que le facteur serait rendu accessible en quantité suffisante dans toutes les stations il y aurait *homogénéisation des croissances* de tous les peuplements, ce qui se produirait ici à 25 ans. Mais les *courbes-types extrêmes se croisent* comme si l'action du facteur devenait négative après avoir été positive.

Il est difficile de concevoir un facteur édaphique agissant sur la croissance suivant toutes ces modalités.

L'*alternative* concernant l'origine biologique des composantes Y_2 et Y_3 est alors la suivante :

— soit il s'agit d'une origine toute autre qu'édaphique, voire plus généralement qu'environnementale ;

— soit il s'agit bien d'une action édaphique mais entrant en *interaction avec un autre phénomène*.

Ce phénomène pourrait être d'ordre *physiologique*. Si on considère encore une fois la composante Y_2 de l'ACP 02 on peut émettre l'hypothèse que les peuplements qui ont le mieux démarré auraient vieilli plus vite que les autres : *l'âge ne serait donc pas le seul paramètre de vieillissement de l'arbre*. Celui-ci vieillirait sous la *double action de l'âge et de l'intensité des croissances antérieures*. Cette conception se rapproche ainsi de celle de *l'âge organique*, ou « organic age » (cf. PRODAN, 1968) défini comme le rapport de l'âge A à l'âge A_M auquel se produit l'accroissement maximum ; les équations de croissance basée sur ce concept génèrent bien, par la

variation du paramètre $\frac{A}{A_M}$, des courbes qui se croisent. Mais on remarquera que les deux âges A_M des deux courbes-types extrêmes de la composante Y_2 ne diffèrent que de 2,5 ans. Une autre formulation de l'âge organique semble donc nécessaire.

L'interprétation du § 4.1 envisage pour le peuplement d'arbres une mémoire dont le contenu irait en s'amenuisant. Ici nous sommes amenés à concevoir une mémoire dont le contenu serait l'âge organique. Il semble donc *a priori* que ces deux interprétations doivent s'exclure l'une l'autre.

4.222. - *Les facteurs génétiques*

Si l'origine de ces deux composantes était génétique elle se manifesterait donc par l'existence simultanée dans le massif forestier landais de génotypes à « croissance rapide » et de génotypes « à croissance lente ». Si on prend pour exemple la composante Y_2 de l'ACP 02, ces divers génotypes seraient alors *le plus facilement repérables (par leur hauteur) entre 15 et 25 ans, à condition de tenir compte au préalable des effets stationnels* déjà mis en évidence (cf. composante Y_1). Cette éventuelle explication génétique devrait selon nous tenir compte du *concept de vieillissement* développé ci-dessus.

4.223. - *Autre perspective d'analyse génécologique*

Les composantes Y_2 et Y_3 de l'ACP 02 ont pu prendre naissance à partir de la *variabilité interindividuelle au sein des peuplements* et non de la variabilité entre peuplements, cette dernière pouvant être inexistante. L'analyse porte en effet sur des arbres, dominants à la coupe rase, dont on ne sait pas tout quant à leur position sociale antérieure au sein du peuplement. S'agit-il dans un cas d'arbres dominants de tout temps et dans un autre cas d'arbres codominants promus à la classe des dominants par le sylviculteur ou les accidents météorologiques (éclaircies sylvicoles ou naturelles, plus ou moins précoces, par le haut ou par le bas) ?

4.224. - *Exemple d'« effet chronologique »*

Les figures 3 et 6 (composante Y_2 de l'ACP 03) montrent que les peuplements de création récente, nés en 1937 et 1947, ont une croissance initiale sensiblement plus rapide que les peuplements nés entre 1897 et 1929. Tous les facteurs envisagés ci-dessus (alinéa 4.221 et 4.222), ainsi que certaines modalités particulières de leur action (alinéa 4.223) sont envisageables. On peut penser par exemple, au plan environnemental à des progrès cultureux (débroussaillages plus fréquents et dépressages plus précoces), et au plan génétique à une sélection sylvicole d'autant plus efficace que le massif landais est de création récente. On n'exclura pas non plus la possibilité que l'expérimentateur ait eu affaire dans un cas (peuplements récents) à des arbres dominants et dans un autre cas (peuplements plus anciens) à des arbres originellement codominants.

4.3. - *Applications dendrométriques*

Il s'agit essentiellement d'envisager l'utilisation des courbes-types, définies par l'analyse factorielle, dans la mise sur pied de Tables de production.

Les composantes mises en œuvre sont assimilables à des *paramètres*. Les *fonctions traditionnelles de croissance* sont remplacées par la *matrice des coefficients de régressions* des accroissements X_i sur les composantes Y_j ; la *courbe-guide* du faisceau complexe est représentée par la séquence des constantes $\beta_{i,0}$, c'est-à-dire par la suite des moyennes des accroissements de l'échantillon des peuplements.

Le choix d'un modèle de croissance en hauteur dominante pour une Table de production résulte selon nous d'un compromis entre, d'une part un certain degré de complexité nécessaire à une prévision honnête, d'autre part un certain degré de simplicité qui est le mérite de tout bon outil. L'examen du tableau 2 devrait nous aider à fixer objectivement notre choix. On remarque en effet que la variabilité de la croissance se décompose en quatre éléments (cf. ACP 02) :

- une composante Y_1 expliquant les 2/3 de la variabilité ;
- deux composantes Y_2 et Y_3 expliquant à elles deux le 1/4 de la variabilité ;
- un résidu (10 p. 100 de la variabilité) fait de composantes non significatives.

Il reste bien entendu à vérifier que la croissance en hauteur des dominants à la coupe rase est une bonne approximation de l'évolution de la hauteur dominante du peuplement.

4.4. - Conclusion générale

Les résultats de ce travail sont de deux ordres :

— les *résultats applicables* concernent, d'une part la description des modalités d'action sur la croissance du facteur édaphique alimentation en eau, d'autre part la possibilité à court terme d'optimiser la construction d'une Table de production ;

— les *résultats d'ordre prospectif* incitent à poursuivre l'effort théorique et expérimental. Les disciplines de recherche concernées sont, à l'énoncé des hypothèses, nombreuses. Nous estimons par ailleurs que le champ d'investigation doit être ouvert plus largement par l'étude de la variabilité individuelle à l'intérieur des divers types de peuplements.

Reçu pour publication en octobre 1980.

Remerciements

Nous remercions tout particulièrement :

N. DECOURT et C. MILLIER qui ont accepté de faire la critique de ce travail et nous ont suggéré des développements au niveau de l'interprétation ;

Ph. BARADAT, C. BOUSSION et A. KREMER qui nous ont aidé dans le traitement automatique des données ;

Les propriétaires et exploitants forestiers qui ont facilité notre tâche sur le terrain .

M^{me} GUILLEM, notre secrétaire, et tous ceux qui ont collaboré avec nous.

Summary

Factorial analysis application to trees height growth : maritime pine exemple

The results we are trying to get for maritime pine stands are the following :

- on one hand, an entire and structured description of growth variability ;
- on the other hand, a building up method for growth curves by using non correlated parameters.

We enter upon this study both for explaining phenomenons (genecological causes) and growth forecasting (Yield Tables).

The growth phenomenon studied is defined as herunder :

« The tree or stand growth is all the increments that follow each other during tree or stand life. »

The material studied consists of dominant trees samples at clear cutting. The measure technique is stem analysis.

The mathematical instrument used is the principal components analysis. The initial variables are successive five years increments X_i .

1. - *Matrix of correlations between initial variables*

That matrix has a very special structure. On one part, it looks like a correlations induced matrix — we assume that the correlation between $5q$ years distant 2 increments X_i and X_{i+q} has to decrease constantly with q increasing, that is true if q is little enough — and on the other part it is unlike that induced matrix — if q is big enough, the observed correlation is 0,5 whatever q may be —. We assume the observed correlation is as matter of fact two correlations combination, the first being constant and showing some surrounding effects, the second decreasing constantly with q increasing and showing, in the long run, the decrease of a tree memory content.

2. - *The principal components Y_j analysis*

We are led to retain three principal components explaining 90 p. 100 total variability.

The first component Y_1 is explained widely by a water supply standard. It gives rise to increments curves patterns that are superposed each other during the stands whole life.

The other components Y_2 and Y_3 have not yet biological explanation. They give rise to increments curves patterns which intersect each other — at 25 years, and at 12 and 40 years respectively —. We envisage the effect of various ecological or genetics factors but interacting with a physiological factor : « the tree growing old would be the result of not only age but also of previous growths intensity ». These growth special modalities would have might be realized by forestry.

3. - *Conclusion*

We distinguish two types of results :

- the applicable results concerning action modality of water factor and the best building up of yield Tables,
- the prospective results inciting us to continue the theoretical and experimental endeavour.

Références bibliographiques

- BIROT Y., JOANNES H., 1973. Variabilité juvénile de la phénologie et des allures de croissance chez le Douglas (*Pseudotsuga menziesii* Mirb. Franco). *IUFRO working party on Douglas Fir provenances*. Göttingen, sept. 2-6, 1973.
- BUIS R., 1966. Recherches factorielles sur la régulation de la croissance de l'hypocotyle de Lupin (*Lupinus albus* L.). *Ann. Physiol. Végét.*, **5** (1), 1-36.
- LEMOINE B., 1969. Le Pin maritime dans les Landes de Gascogne. Etude des relations d'Allométrie concernant le volume des peuplements, en liaison avec certaines caractéristiques de la station. *Ann. Sci. Forest.*, **26** (4), 445-473.
- LEMOINE B., 1979. Pin maritime et sécheresses dans les Landes de Gascogne. Croissances en circonférences. *Communication à l'Académie d'Agriculture de France*. Séance du 2 mai 1979.
- PRODAN M., 1968. *Forest Biometrics*. 447. *Pergamon Press*.
- TOMASSONE R., 1966. *Analyse des composantes principales et régression orthogonale*. C.N.R.F., Station de Biométrie, Nancy.