

ANALYSE EN COMPOSANTES PRINCIPALES DES PROPRIÉTÉS TECHNOLOGIQUES DES BOIS MALGACHES

F. CAILLIEZ et P. GUENEAU

Centre technique forestier tropical
45 bis, avenue de la Belle Gabrielle
94 - Nogent-sur-Marne

RÉSUMÉ

Cette étude a été menée avec deux objectifs :

— trouver dans l'arsenal des méthodes d'analyse de données un outil approprié à la description technologique des bois,

— utiliser cet outil pour établir un catalogue des vocations industrielles des bois malgaches.

L'outil employé, l'analyse en composantes principales, a été appliqué sur 9 caractères physiques et mécaniques mesurés dans :

— 414 arbres malgaches, appartenant à 73 genres et 125 espèces botaniques, provenant de 22 stations écologiques réparties dans les 6 provinces de Madagascar,

— 151 arbres africains, servant de jalons, appartenant à 16 genres et 35 espèces botaniques, provenant de 7 pays d'Afrique Équatoriale et Occidentale.

Les deux premières composantes principales trouvées expliquent 83,9 % de la variation totale de cette population et s'interprètent physiquement de façon claire : la première (*solidité*) est une composante de taille, la seconde (*sensibilité*) une composante de forme.

La représentation des 9 caractères initiaux par ces deux nouvelles variables s'avère fructueuse et amène des conclusions concrètes en permettant de déceler des similitudes ou des différences entre bois, ainsi que des observations synthétiques sur la dispersion interne des espèces.

Ces conclusions ressortent beaucoup plus clairement que par la considération séparée des variables initiales entre lesquelles de nombreuses corrélations, d'ailleurs chiffrées par l'analyse, masquent souvent le sens exact de la fluctuation des valeurs numériques.

Sur les conclusions de cette analyse, un catalogue des vocations industrielles des bois malgaches a été établi et publié par ailleurs.

INTRODUCTION

De nombreux pays tropicaux tirent un revenu élevé de leur capital forestier en faisant entrer un volume croissant de leurs bois dans les industries modernes de transformation : certaines espèces comme l'Okoumé, se sont fait une place importante sur les marchés où leur réputation est maintenant solidement établie.

Par sa situation géographique, Madagascar n'a participé que très peu, jusqu'à présent, à cette tendance, le début du développement dans ce domaine étant paralysé par le manque d'information sur les aptitudes potentielles des bois locaux aux emplois industriels. Alors que les utilisations artisanales traditionnelles reposent sur des connaissances empiriques suffisantes, les vocations des différents bois pour des industries modernes telles que le placage, la charpente collée ne sont pas connues.

Le nombre élevé des bois présumés intéressants (une centaine d'essences⁽¹⁾) et la multiplicité des usages possibles interdisent de procéder exhaustivement à l'essai de tous les bois pour tous les emplois en vraie grandeur, sur des machines de type industriel, ces essais réels étant très coûteux.

Les seules données objectives disponibles proviennent des mesures des propriétés physiques et mécaniques, par des méthodes normalisées, en laboratoire : seul un tri préalable des vocations des différents bois, opéré sur ces données de laboratoire peut orienter l'exécution ultérieure d'essais réels de fabrication (placages, etc.), essais devenant alors envisageables puisque leurs résultats auront des chances raisonnables d'être positifs.

C'est une méthode de tri préalable, sur les données physiques et mécaniques recueillies en laboratoire, qui est présentée ici.

I. — LES DONNÉES TECHNOLOGIQUES — MÉTHODE D'ACQUISITION MÉTHODE D'ANALYSE

1.1. — *Échantillonnage*

Les caractères physiques et mécaniques (densité, élasticité, résistance au choc, ... etc.) sont mesurés sur un certain nombre d'individus (individu = arbre) de chaque espèce.

Plus précisément, pour chaque espèce un plan d'échantillonnage est établi, tenant compte de son intérêt, présumé par sa fréquence, de ses utilisations locales connues le cas échéant, et de sa répartition géographique : le plan d'échantillonnage prévoit le nombre d'arbres à récolter dans chaque zone de forêt, zones elles-mêmes fonctions du réseau de communication existant ou en projet et des localisations des marchés de consommation des bois (marchés intérieurs : communications routières ou ferrées; marchés extérieurs : évacuation vers un port).

Un arbre-échantillon fournit finalement une pièce de bois qui est transportée au laboratoire et débitée en éprouvettes servant aux divers essais. Chaque caractère est mesuré sur deux éprouvettes provenant du même arbre, la moyenne arithmétique des résultats par caractère étant seule retenue dans la suite pour décrire un arbre. Ce faible nombre d'éprouvettes (deux) par arbre a été adopté en même temps que l'on a augmenté le nombre d'arbres échantillonnés par espèce : cette combinaison apporte une quantité d'information et une précision sur la moyenne d'espèce nettement supérieures à celles que donne, à dépense égale, l'essai d'un faible nombre de pieds échantillonnés chacun à dix éprouvettes. Dans les deux cas d'ailleurs, les éprouvettes sont toujours tirées du bois parfait (hors cœur et hors aubier), c'est-à-dire dans le bois effectivement et exclusivement utilisé par l'industrie.

Ici nous disposons de 565 arbres⁽²⁾ décrits chacun par une vingtaine de caractères chiffrés (moyennes d'arbre); un tiers environ de cet effectif a été testé sur dix éprouvettes par arbre et deux tiers sur deux éprouvettes par arbre.

(1) Le terme d'essence forestière désigne ici un groupe d'espèces botaniques que les pratiques commerciales actuelles rapprochent ou confondent sous un même nom vernaculaire.

(2) Dont 151 sont des arbres africains d'espèces commerciales éprouvées, introduits dans l'analyse comme jalons pour servir de points de comparaison avec les bois malgaches.

1.2. — Choix et mesure des propriétés retenues

Certains des caractères mesurés en laboratoire sont plus importants que d'autres :

— ou ils ont une influence déterminante pour l'emploi (exemple : la résistance à la flexion pour l'emploi en charpente),

— ou leur intérêt est fondamental et ils sont systématiquement mesurés et fournis par tous les laboratoires du monde, à cause de leur généralité (exemple : la densité).

D'autres sont moins utiles à considérer :

— ou leur protocole de mesure est mal uniformisé d'un pays à l'autre (exemple : résistance au cisaillement),

— ou ils présentent une « redondance » avec d'autres caractères mesurés par ailleurs, avec lesquels ils sont très liés pour des raisons physiques profondes ou par le protocole de leur mesure (exemple : sur chaque éprouvette d'essai, le point de saturation S n'est pas mesuré mais calculé par le rapport B/v des deux caractères B et v qui eux, sont effectivement mesurés. La valeur numérique de S donne un renseignement directement utilisable par le praticien, mais n'apporte aucune information supplémentaire.

Des caractères cités par SALLENAVE (1955), nous avons retenu et mesuré par les protocoles qu'il indique :

- 1 — (N) *Dureté* (méthode de Chalais-Meudon) : représente l'importance de l'empreinte laissée dans le bois par un solide que l'on y enfonce sous un effort donné (à 12 % d'humidité).
- 2 — (D) *Densité* : poids spécifique du bois (à 12 % d'humidité).
- 3 — (B) *Rétractibilité volumétrique totale* : pourcentage de diminution de volume du bois entre l'état saturé et l'état anhydre, rapporté à l'état anhydre.
- 4 — (v) *Coefficient de rétractibilité volumétrique* : pourcentage de diminution de volume du bois pour une variation de 1 % de l'humidité interne au voisinage de l'humidité normale de 12 %.
- 5 — (T + R) *Somme des rétractibilités linéaires totales, tangentielle et radiale* : chacune d'elles est le pourcentage de diminution de longueur du bois entre l'état saturé et l'état anhydre, l'une dans la direction tangentielle, l'autre dans la direction radiale ⁽¹⁾.
- 6 — (C) *Résistance à la rupture en compression de fil* (à 12 % d'humidité).
- 7 — (F) *Résistance à la rupture en flexion statique* (à 12 % d'humidité).
- 8 — (E) *Module d'élasticité apparent en flexion statique* (à 12 % d'humidité).
- 9 — (K) *Coefficient de résistance en flexion dynamique* (choc) (à 12 % d'humidité).

Chaque arbre-échantillon peut donc être représenté par un point dans l'espace à 9 dimensions, où ses coordonnées sont les 9 caractères mesurés sur lui (moyennes d'éprouvettes).

La disposition géométrique des points dans cet espace à 9 dimensions ne peut pas être visualisée concrètement. Pourtant elle apporterait une vue extrêmement synthétique et fructueuse de l'ensemble des résultats de tous les individus de toutes les espèces étudiées :

— des points voisins indiqueraient des bois très semblables à tous points de vue à la fois,
— un groupement d'arbres de la même espèce en un nuage dense indiquerait une bonne homogénéité de l'espèce, à tous les points de vue à la fois,

— la position géométrique d'un tel nuage dans l'espace à 9 dimensions renseignerait immédiatement sur la vocation industrielle présumée du bois (exemple : zone des fortes duretés et des faibles rétractibilités = zone des bois à parquets),

— la distance des nuages de deux espèces différentes renseignerait sur la différence globale entre deux bois. Une faible distance permettrait éventuellement de préconiser ou d'autoriser des groupements commerciaux, des mélanges de divers bois pour le même emploi,

— dans une espèce à nuage très étendu, des scissions possibles apparaîtraient, correspondant par exemple à des origines géographiques différentes et pourraient permettre de diviser l'espèce en deux sous-groupes à vocations différentes.

On s'est donc fixé pour but de saisir d'un seul coup d'œil toute l'information contenue dans le tableau des $565 \times 9 = 5\,085$ données, ou, ce qui revient au même, toute l'information contenue dans le nuage des 565 points de l'espace à 9 dimensions.

⁽¹⁾ Cette somme ne diffère théoriquement de B que par l'intervention du retrait longitudinal et de termes d'ordre supérieur et égal à deux : on a surtout retenu cette variable linéaire (T + R) pour la comparer à B mesurée volumétriquement, et non pour le minime supplément d'information qu'elle apporte.

1.3. — *Méthode d'analyse* (1)

La méthode la mieux à même de représenter concrètement ce nuage de points est l'*analyse en composantes principales* qui consiste à projeter le nuage dans un sous-espace de dimension plus faible, donc représentable graphiquement, de façon que le nuage projeté ressemble le plus possible au nuage de départ; plus précisément en désignant par P_1, P_2, \dots, P_{565} les 565 points (arbres) du nuage à 9 dimensions et par Q_1, Q_2, \dots, Q_{565} leurs projections sur un espace de dimension $p < 9$, on retient parmi les sous-espaces de dimension p , celui qui rend minimum la somme des carrés des distances des points P_i à leurs projections Q_i ($i = 1, \dots, 565$).

— Supposons que cette somme soit déjà très faible pour $p = 2$: on représentera alors le nuage dans un plan avec une bonne approximation;

— s'il faut $p = 3$ pour que cette somme soit très faible, on représentera le nuage dans l'espace à 3 dimensions, c'est-à-dire en pratique dans trois plans;

— et ainsi de suite.

Dans le cas présent, la représentation à *deux dimensions* (Y_1, Y_2) est déjà très bonne, c'est ce qui a autorisé une très large utilisation des observations faites dans ce plan (Y_1, Y_2), observations dont nous donnons plus bas plusieurs exemples (chapitre IV).

Soulignons l'un des intérêts de la méthode : elle n'utilise aucune hypothèse sur la forme des distributions des 9 caractères mesurés en laboratoire, en particulier sur leur normalité (contrairement aux techniques de statistique classique, telles que l'analyse de variance); elle entre à ce titre dans le groupe des méthodes d'« Analyse des données » et sous ce nom on comprend mieux son rôle d'intermédiaire entre l'expérimentateur et le résultat de ses essais.

L'analyse en composantes principales fournit en somme un *outil optique* qui permet de voir dans un plan une représentation de l'espace à 9 dimensions inaccessible directement à nos sens, moyennant, c'est inévitable, une certaine distorsion, qui est seulement rendue la plus faible possible.

On a songé à d'autres méthodes d'analyse multidimensionnelle, telles que l'analyse discriminante, la classification automatique, l'analyse canonique.

— *L'analyse discriminante* : un de ses buts est de voir dans quelle mesure des sous-populations connues a priori, par exemple des espèces botaniques, diffèrent sur le plan des caractères mesurés, ici les propriétés technologiques. La composition de notre échantillonnage (espèces botaniques très nombreuses, origines géographiques très nombreuses) faisait que les sous-populations que l'on souhaitait comparer étaient en nombre très élevé et composées chacune de peu d'individus, ce qui aurait imposé d'effectuer au préalable des regroupements qui, au stade initial, ne pouvaient être que subjectifs.

Une telle analyse a été effectuée sur une partie du fichier, le groupe des Tavolo, mais le découpage choisi (identité botanique, abstraction faite des provenances) est critiquable car dans ce groupe constitué d'espèces ou de races très voisines du point de vue de la systématique botanique, la notion de provenance peut s'avérer primordiale.

— *La classification automatique* : elle a pour but de former des sous-populations d'individus « qui se ressemblent »; l'empirisme, inévitable dans ces méthodes, nous a fait préférer l'analyse en composantes principales qui permet, pourvu que le plan principal soit une bonne image de la réalité, de classer « à vue ».

— *L'analyse canonique* : elle a pour but d'examiner les liens entre deux ensembles de caractères, ici par exemple le groupe des propriétés physiques et celui des propriétés mécaniques. Elle sera très utile dans une étude ultérieure où le centre d'intérêt sera, moins la description des bois, que l'analyse critique des propriétés mesurées.

II. — RÉSULTATS

Le programme fournit successivement :

— le tableau des données initiales arbre par arbre, caractère par caractère,

— une nomenclature en clair rappelant l'identité de chaque arbre, ses adresses dans les fichiers des laboratoires, le lieu de récolte codifié, et le nombre d'éprouvettes utilisées pour le décrire,

— une description de la distribution de chacune des 9 variables dans la population

(1) Son exposé théorique fait l'objet de l'annexe I.

totale : histogramme (vingt classes d'égale amplitude), moyenne, écart-type, coefficient de variation, observations minimum et maximum,

— la matrice de covariance et la matrice de corrélation des variables initiales, les valeurs propres et vecteurs propres de cette dernière, la contribution de chaque axe principal à l'inertie totale,

— les valeurs des premières composantes principales pour chacun des 565 arbres,

— les corrélations des variables initiales avec les deux premières composantes principales,

— la position de l'ancien repère (caractères initiaux) par rapport au nouveau (axes principaux),

2.1. — Histogrammes

Exemple : distribution de la densité (variable 2) : Fig. 1.

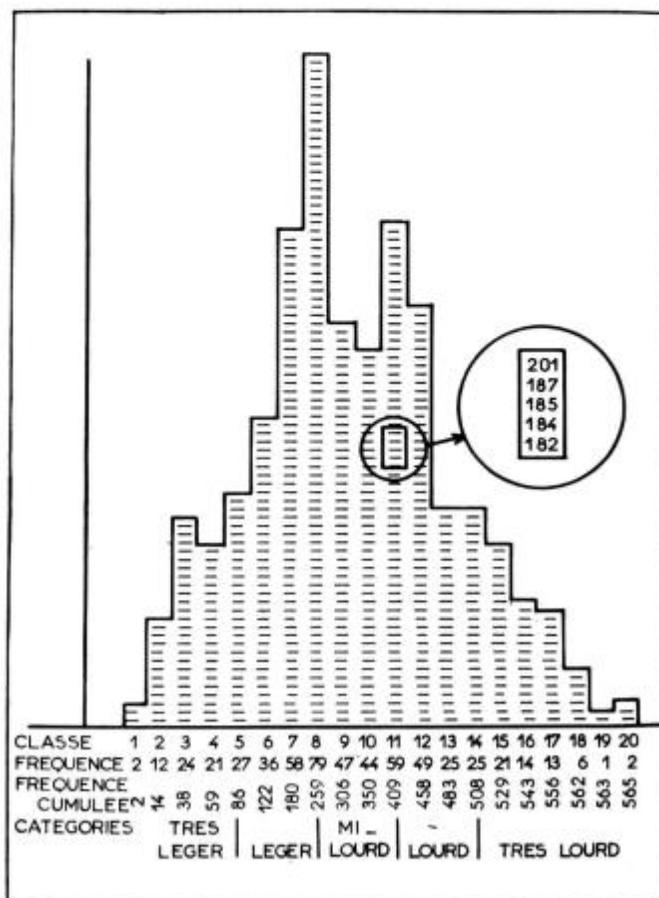


FIG. 1. — Histogramme de la densité

FIG. 1. — Histogram of density

- la qualité de la représentation de chaque arbre sur les axes et plans principaux,
- le graphique à grande échelle représentant la projection du nuage initial de 565 points sur le plan des deux premiers axes principaux (obtenu au traceur Benson).

L'intérêt du programme est que les arbres ne perdent pas leur identité (n° de nomenclature du fichier).

Les possibilités d'utilisation des histogrammes sont multiples, par exemple :

- ils donnent la forme de la distribution (assez symétrique pour la densité — fig. 1),
- ils localisent immédiatement l'ensemble des résultats par rapport à un système préétabli de catégories dont le maniement est déjà couramment utilisé par les praticiens. Par exemple, en ce qui concerne la densité (fig. 1), on constate que le découpage empirique en vigueur depuis longtemps est bien adapté.

2.2. — Matrice de corrélation des caractères initiaux

Analyse I (390 arbres malgaches seulement)

	N	D	B	ν	T + R	C	F	E	K
N	1,00	0,84	0,34	0,47	0,35	0,78	0,70	0,70	0,55
D	0,84	1,00	0,61	0,60	0,60	0,90	0,85	0,80	0,67
B	0,34	0,61	1,00	0,71	0,86	0,63	0,63	0,63	0,56
ν	0,47	0,60	0,71	1,00	0,59	0,61	0,62	0,58	0,54
T + R	0,35	0,60	0,86	0,59	1,00	0,63	0,63	0,62	0,55
C	0,78	0,90	0,63	0,61	0,63	1,00	0,88	0,81	0,71
F	0,70	0,85	0,63	0,62	0,63	0,88	1,00	0,85	0,75
E	0,70	0,80	0,63	0,58	0,62	0,81	0,85	1,00	0,68
K	0,55	0,67	0,56	0,54	0,55	0,71	0,75	0,68	1,00

Analyse II (414 arbres malgaches + 151 arbres africains)

	N	D	B	ν	T + R	C	F	E	K
N	1,00	0,85	0,35	0,49	0,36	0,80	0,76	0,72	0,57
D	0,85	1,00	0,59	0,63	0,61	0,91	0,90	0,82	0,71
B	0,35	0,59	1,00	0,72	0,88	0,57	0,61	0,62	0,54
ν	0,49	0,63	0,72	1,00	0,60	0,61	0,63	0,62	0,53
T + R	0,36	0,61	0,88	0,60	1,00	0,60	0,63	0,62	0,56
C	0,80	0,91	0,57	0,61	0,60	1,00	0,92	0,83	0,73
F	0,76	0,90	0,61	0,63	0,63	0,92	1,00	0,88	0,79
E	0,72	0,82	0,62	0,62	0,62	0,83	0,88	1,00	0,70
K	0,57	0,71	0,54	0,53	0,56	0,73	0,79	0,70	1,00

Les coefficients de corrélation totale, bien que ne caractérisant que les liaisons linéaires, sont déjà éloquentes : ils font apparaître, d'une part le groupement des variables de résistance mécanique (C, F, E, K) avec dureté (N) et densité (D) et d'autre part le groupement des variables de rétractibilité (B, ν , T + R). L'existence de ces liaisons était prévisible, de par la nature physique des caractères mesurés, mais leurs intensités sont précisées et peuvent être comparées entre elles.

Il ne faut cependant pas accorder dès ce stade trop de foi aux seules valeurs numériques des coefficients de corrélation car il est nécessaire d'observer en même temps le nuage de points dans les plans des caractères initiaux pris deux à deux, plusieurs configurations de nuage pouvant fournir la même valeur du coefficient de corrélation.

2.3. — Composantes principales

2.31. Contribution des premières composantes à l'explication de la variation.

Analyse I

Composante	Contribution (%)	Contribution cumulée (%)
1	70,5	70,5
2	11,7	82,2
3	5,1	87,3
4	4,6	91,9

Analyse II

Composante	Contribution (%)	Contribution cumulée (%)
1	71,5	71,5
2	12,4	83,9

2.32. Corrélations entre les premières composantes principales et les caractères initiaux.

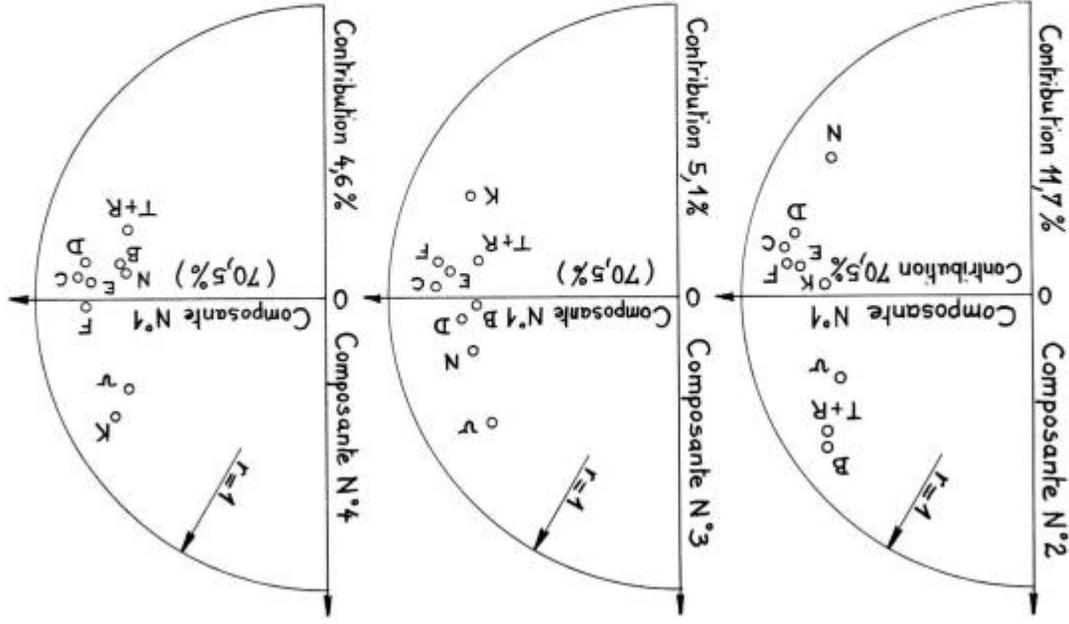
Analyse^{*}

Composante n°	Caractères initiaux								
	N	D	B	v	T + R	C	F	E	K
1	0,77	0,92	0,78	0,75	0,77	0,93	0,92	0,89	0,80
2	-0,53	-0,24	0,55	0,30	0,51	-0,18	-0,13	-0,12	-0,04
3	0,21	0,07	0,03	0,48	-0,14	-0,01	-0,12	-0,09	-0,38
4	-0,10	-0,13	-0,11	0,33	-0,25	-0,08	0,02	-0,07	0,44

Ces corrélations sont représentées sur la figure 2.

FIG. 2. — Représentation des caractères initiaux par leurs coefficients de corrélation avec les premières composantes principales (Analyse I)

FIG. 2. — Representation of initial characters by their correlation coefficients with the first principal components (Analysis I)



Analyse II

Composante n°	Caractères initiaux								
	N	D	B	v	T + R	C	F	E	K
1	0,79	0,93	0,76	0,76	0,76	0,92	0,95	0,90	0,81
2	-0,48	-0,21	0,60	0,31	0,53	-0,22	-0,17	-0,11	0,07
3	0,21	0,06	0,02	0,46	-0,16	-0,01	-0,10	-0,03	-0,41

Ces corrélations sont représentées sur la figure 3.

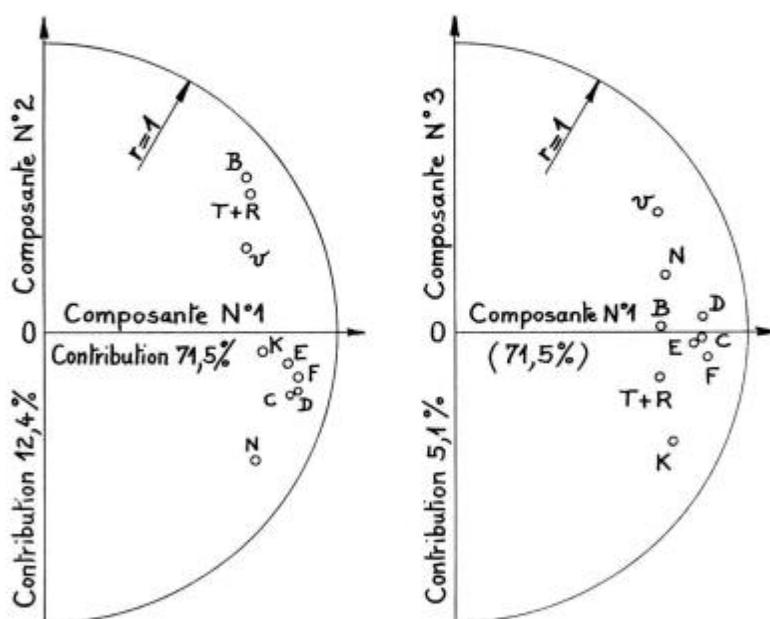


FIG. 3. — Représentation des caractères initiaux par leurs coefficients de corrélation avec les premières composantes principales (Analyse II)

FIG. 3. — Representation of initial characters by their correlation coefficients with the first principal components (Analysis II)

2.4. — Commentaire des résultats

Comme on le verra au chapitre III, l'échantillonnage des bois africains a une structure très différente de celui des bois malgaches : chaque genre botanique africain est représenté par un

nombre d'individus important, appartenant en plus très souvent à la même espèce, alors que beaucoup de bois malgaches sont représentés par deux ou trois individus seulement, et qu'à l'intérieur des genres, le nombre des espèces différentes est beaucoup plus grand.

Malgré cela les deux analyses donnent des résultats très voisins en ce qui concerne :

- la moyenne générale de chaque caractère initial,
- l'écart-type de chaque caractère initial,
- les corrélations entre variables initiales,
- les premiers vecteurs propres, c'est-à-dire le « mode de construction » des premières composantes principales à partir des caractères initiaux.

Cette unicité des résultats suggère que *l'analyse ait une portée générale*, non tributaire de l'échantillonnage utilisé. Les observations qu'elle permet ont donc valeur pour les bois feuillus tropicaux, en particulier les lois qui relient entre elles les variables initiales (corrélations, dispersions inter et intra-génériques, etc...) et aussi celles qui lient ces variables initiales aux deux premières composantes principales. Celles-ci ont donc un sens physique général pour le bois, de même que l'interprétation de diverses directions dans le premier plan principal.

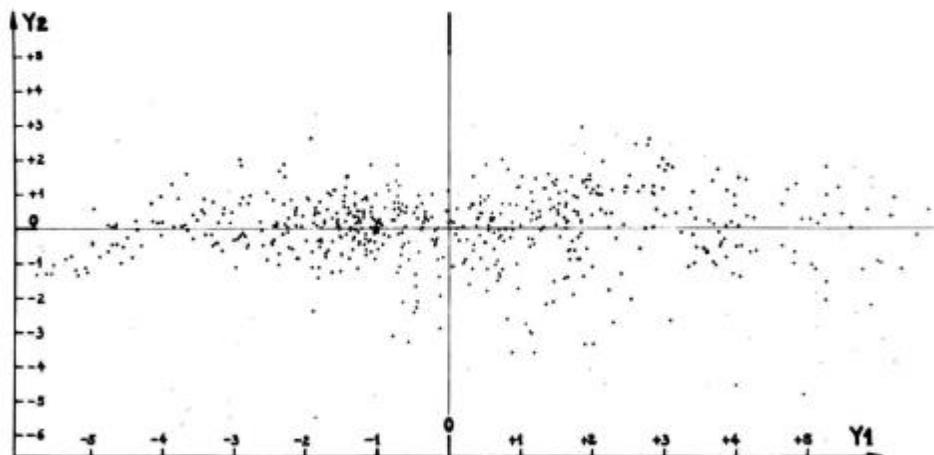


FIG. 4. — Le nuage des 565 points dans le plan des deux premières composantes principales

FIG. 4. — The cloud of 565 points in the plane of the two first principal components

Cette généralité de la valeur de l'analyse a plusieurs conséquences pratiques :

- dans les conclusions qu'elle apporte *sur les propriétés mesurées* (variables initiales, c'est-à-dire caractéristiques physiques et mécaniques), par exemple leurs dispersions comparées, leurs corrélations : on pourra peut-être modifier certains protocoles de laboratoire au vu de ces conclusions, parce que l'on sait maintenant qu'elles ne sont pas particulières à l'échantillonnage de bois utilisé. Ce genre d'étude pourrait même aboutir à la suppression de certaines mesures, mais la question n'est pas du tout abordée ici, où nous commençons par tirer les conclusions les plus évidentes sur les positions comparées de diverses essences,

— dans les conclusions qu'elle apporte *sur les positions comparées* de diverses essences. La stabilité des conclusions vis-à-vis de l'échantillonnage autorise à incorporer après coup des essais effectués postérieurement à la présente analyse; il suffit pour un arbre à partir des résultats d'essais de laboratoire, de calculer les valeurs de ses composantes principales (expressions numériques du tableau 1) et de représenter l'arbre par un point supplémentaire, bien qu'il n'ait pas participé à la détermination des axes principaux.

2.41. Signification de Y_1 (Contribution à l'inertie : 71,5 % — Cf. § 232)

Y_1 est corrélée positivement et fortement avec toutes les variables initiales et spécialement avec D, C, F, E, K, elles-mêmes très liées entre elles. Y_1 est donc typiquement une *composante « de taille »* : Y_1 fort indique un bois *fort en tout*, particulièrement en densité et en résistances mécaniques, la dureté et les rétractibilités étant moins étroitement associées à cette notion de taille que la densité et les résistances mécaniques.

Y_1 peut donc être baptisée *solidité*, évoquant à la fois le poids (densité) et la résistance aux efforts mécaniques.

2.42. Signification de Y_2 (Contribution à l'inertie : 12,4 % — Cf. § 232)

Y_2 fort qui met en opposition la dureté et les rétractibilités indique un bois à la fois *tendre et instable en dimensions*. On peut donc baptiser cette deuxième composante principale *sensibilité* (pour éviter le terme de nervosité, employé par ailleurs pour qualifier la variable initiale v , et celui d'instabilité, dont le sens de variation serait plus ambigu).

On sous-entend ainsi *sensibilité* à l'effet de pénétration d'un corps solide (faible dureté) et en même temps à l'effet de pénétration de l'eau (forte rétractibilité, forte nervosité).

On peut dire d'une autre façon :

les corrélations de la seconde composante principale avec la rétractibilité et la dureté (en opposition) signifient qu'à « taille » égale, le meilleur point de vue pour décrire la « forme » d'un bois est de considérer à quel point il réunit la conjugaison rétractibilité faible — dureté forte, tout en sachant que d'une façon encore plus globale les caractères initiaux, y compris ces deux-là, montrent une première tendance (première composante) à augmenter ou diminuer tous en même temps.

2.43. Directions obliques du plan Y_1Y_2

La figure 5 fournit l'interprétation des déplacements dans diverses directions obliques du plan Y_1Y_2 : certaines de celles-ci représentent assez bien la variation d'un des caractères initiaux; plus précisément, cette interprétation est d'autant plus juste que le vecteur unitaire correspondant est, dans l'espace, plus proche du plan Y_1Y_2 , c'est-à-dire que la longueur de sa projection sur ce plan est plus proche de l'unité. De ce point de vue, les caractères B et (T + R) sont les mieux représentés.

Autrement dit encore, entre deux points quelconques du plan Y_1Y_2 , un déplacement dans le sens du vecteur B de la figure 5 représente assez bien une augmentation de la rétractibilité volumétrique.

TABLEAU I

Caractères initiaux	Unité	Symbole	Moyenne (¹)	Ecart type (¹)	Expression numérique des deux premières C.P. (¹)	
					Y ₁	Y ₂
Dureté (Chalais-Meudon)	mm ⁻³	N	0,4945 E1	0,3562 E1	+ (0,8695 E-1) × N	— (0,1275 E0) × N
Densité	g/cm ³	D	0,7171 E0	0,1878 E0	+ (0,1954 E1) × D	— (0,1084 E1) × D
Rétractibilité volum. totale	%	B	0,1476 E2	0,3974 E1	+ (0,7551 E-1) × B	+ (0,1425 E0) × B
Coefficient de Retr. volum.	%	v	0,5215 E0	0,1439 E0	+ (0,2072 E1) × v	+ (0,2068 E1) × v
Somme des Retr. linéaires	%	(T+R)	0,1402 E2	0,3970 E1	+ (0,7564 E-1) × (T+R)	+ (0,1275 E0) × (T+R)
Résis. en compression de fil	Kg/cm ²	C	0,6109 E3	0,1838 E3	+ (0,1982 E-2) × C	— (0,1153 E-2) × C
Résis. en flexion statique	Kg/cm ²	F	0,1533 E4	0,5065 E3	+ (0,7356 E-3) × F	— (0,3214 E-3) × F
Module d'élast. apparent en flexion statique	Kg/cm ²	E	0,1261 E6	0,4097 E5	+ (0,8688 E-5) × E	— (0,2560 E-5) × E
Coefficient de résilience	100 Kg/cm ²	K	0,4123 E0	0,2188 E0	+ (0,1460 E1) × K — (0,9123 E1)	— (0,2920 E0) × K — (0,1922 E1)

(1) L'analyse a été faite sur les caractères initiaux centrés réduits mais on revient ici aux caractères initiaux en valeurs réelles, telles qu'elles sont mesurées en laboratoire.

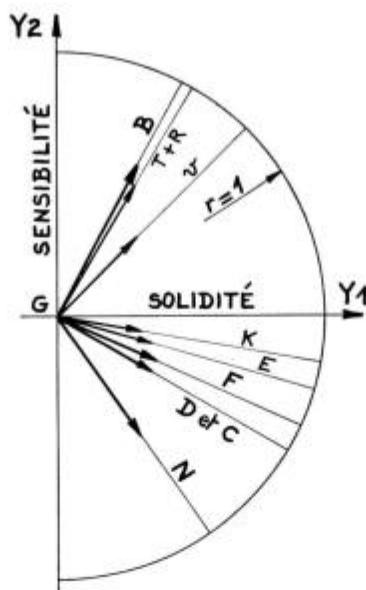


FIG. 5. — Projections dans le premier plan principal des vecteurs unitaires de l'ancien système de référence (caractères initiaux)

FIG. 5. — Projection in the first principal plane of the unit vectors of the initial axis

Les autres caractères initiaux ont leur représentation sur la fig. 5 et fournissent chacun l'interprétation plus ou moins rigoureuse selon les cas, d'une direction de déplacement dans le plan Y_1Y_2 .

Certains exemples particuliers du chapitre III ont employé ce type de raisonnement pour préciser les positions relatives de nuages d'espèces, non seulement en distance, mais en direction.

III. — DESCRIPTION DE QUELQUES BOIS

En visualisant certains individus extraits de la masse des résultats, par exemple les arbres de même identité botanique, on fait apparaître des sous-ensembles que l'on peut comparer entre eux d'un seul coup d'œil.

On peut ainsi concrétiser les nuages d'espèces dans le plan Y_1Y_2 , et observer :

— leur étendue (allongement dans diverses directions), qui représente la dispersion globale des propriétés technologiques dans cette espèce,

— leur composition interne (les origines géographiques peuvent apparaître clairement par une ségrégation à l'intérieur du nuage),

— leur proximité, d'une espèce à l'autre, leur imbrication partielle ou leur recouvrement complet, qui présumeront de la possibilité de grouper différents bois pour les mêmes emplois industriels,

— la position du nuage d'une espèce non encore commercialisée par rapport à des bois « jalons » déjà utilisés dans l'industrie, qui permettra d'orienter des vocations pour les espèces encore peu employées.

Les deux derniers points nécessitent d'ailleurs une réserve importante : l'éloignement de deux espèces dans cette représentation signifie sans ambiguïté qu'elles ont des propriétés

technologiques différentes et donc très vraisemblablement des emplois différents, mais par contre la proximité ou la superposition de deux bois n'est qu'une présomption, un premier pas, et n'implique pas forcément la possibilité de les employer similairement, et cela pour deux raisons :

— la proximité de deux bois en projection sur le premier plan principal n'implique leur proximité réelle dans l'espace à 9 dimensions que si l'on vérifie la bonne qualité de leur représentation; cette vérification peut être faite point par point en consultant le listing des distances des points au premier plan principal

— nombre de caractéristiques fort importantes, non prises en considération ici, peuvent rendre ces bois très différents (couleur, grain, durabilité, possibilités d'approvisionnement, dimension moyenne des grumes, etc...). Ces qualités doivent être considérées une fois l'analyse finie, pour en confirmer ou critiquer les premières présomptions.

Rappelons les raisons qui ont amené à entreprendre le présent travail : après plusieurs années passées à acquérir les données technologiques concernant les bois malgaches (récoltes, essais de laboratoire), il était indispensable de fournir à la profession du bois un catalogue des usages possibles de chaque essence.

Les conclusions qui y figurent (GUENEAU, 1971) proviennent de deux sources :

— l'examen détaillé pour chaque bois de sa représentation dans la présente analyse en composantes principales,

— la prise en compte d'informations complémentaires extérieures à l'analyse (d'ordre économique, esthétique, etc.).

Nota : Incertitude expérimentale sur les deux premières composantes principales — précision graphique.

Chaque caractéristique initiale est mesurée avec une certaine incertitude instrumentale au laboratoire; il est facile de calculer l'influence cumulée de ces incertitudes des 9 caractères initiaux sur la valeur numérique des deux premières composantes principales : l'erreur relative d'origine instrumentale commise sur Y_1 et Y_2 est inférieure à 2 %. A la dimension du graphique fourni par le traceur Benson (1,71 m \times 0,66 m), le rectangle d'erreur expérimentale maxima pour un point mesure 2,8 cm \times 1,3 cm; au-dessous de ces distances, il faut donc s'abstenir de tirer des conclusions sur la position relative de 2 points : tous les raisonnements ci-dessous se plient à cette règle et les distances considérées comme significatives entre polygones d'espèces sont toujours très supérieures à cette incertitude.

3.1. — Arofy — Okoumé — Ramy (fig. 6)

Nom vernaculaire	Échantillon utilisé — Identité	Provenance
Arofy (12 arbres)	<i>Commiphora barorum</i> <i>Commiphora pterocarpa</i> <i>Commiphora sp. p</i>	Andranolava } Sud-Ouest Sakaraha } Madagascar
Okoumé (11 arbres)	<i>Aucoumea klaineana</i>	Gabon Congo
Ramy (20 arbres)	<i>Canarium madagascariense</i> Diverses variétés locales	Antsohihy } Nord-Ouest Périnet } Est, moyenne altitude

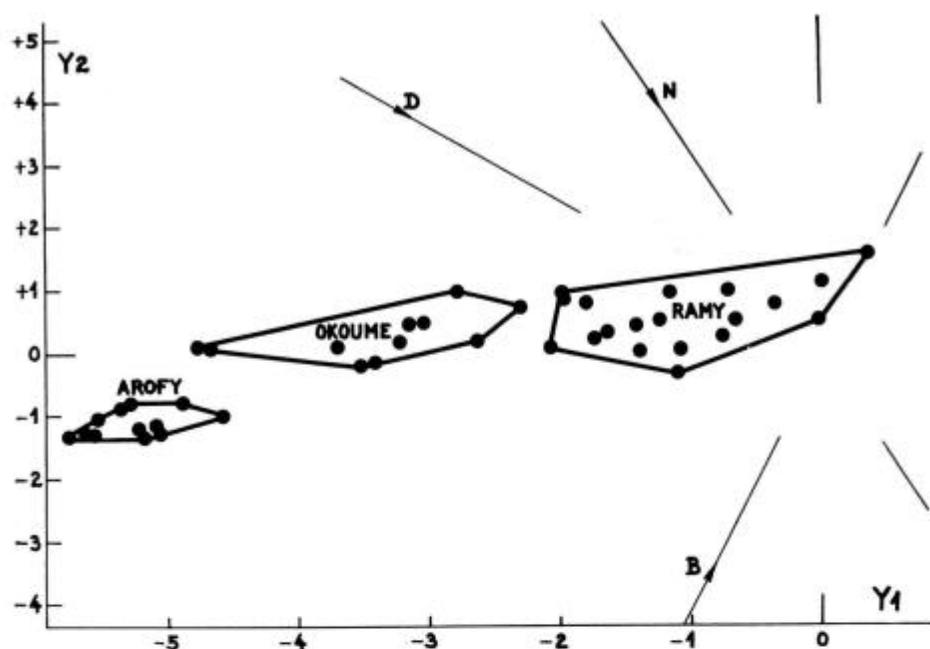


FIG. 6. — Arofy - Okoumé - Ramy

FIG. 6. — Arofy - Okoumé - Ramy

Certains projets industriels concernant les Ramy et les Arofy malgaches nous conduisent à les comparer à l'Okoumé.

Observations

- Aucun recouvrement entre ces trois bois. Pour utiliser Arofy et Ramy dans les emplois de l'Okoumé, on peut prévoir les difficultés suivantes :
 - aucun d'eux n'est exactement semblable à l'Okoumé,
 - l'Arofy, trop fragile, aux essais de déroulage a effectivement donné des feuilles pelucheuses se déchirant facilement,
 - le Ramy au contraire, doit s'usiner plus difficilement que l'Okoumé et exiger de plus fortes puissances.
- Aucune différence à faire dans le commerce entre *Commiphora barorum* et *Commiphora pterocarpa* et on peut raisonnablement accepter tous les *Commiphora* malgaches sous le nom pilote d'Arofy.

Ses emplois seraient ceux de bois très peu solides, moins que l'Ayous, moins que le Fromager, mais aussi nettement moins *rétractiles* que la majorité de tous les bois étudiés, ce qui est une qualité fort importante et peut éventuellement présenter un intérêt industriel particulier.

— Aucune différence pratique à faire dans le commerce entre les Ramy de l'est et de l'ouest, ni entre « Ramy Fotsy » et « Ramy Mena », ni entre « Ramy à grandes feuilles » et « Ramy à petites feuilles », distinctions qui sont cependant très ancrées dans l'esprit des utilisateurs. Deux études parallèles de systématique botanique et d'anatomie du bois sont d'ailleurs aussi arrivées à la même conclusion.

— La variabilité des propriétés d'un arbre à l'autre dans les Ramy n'est pas plus importante que dans les Okoumé.

3.2. — *Ayous (Samba)* — *Alampona* — *Fantsilotra* — *Voantsilana* (fig. 7)

Nom vernaculaire	Échantillon utilisé — Identité	Provenance	
Ayous (Samba) (10 arbres)	<i>Triplochiton scleroxylon</i>	Cameroun Côte d'Ivoire R.C.A.	
Alampona (2 arbres)	<i>Hibiscus lasiococcus</i>	Ranomafano (F.) Périnet	} Est moyenne altitude
Fantsilotra (8 arbres)	<i>Alluaudia procera</i>	Amboasary Ifotaka	} Sud
Voantsilana (7 arbres)	<i>Cussonia</i> sp. <i>Polyscias fraxinifolia</i> <i>Araliacées</i> sp. p	Périnet Antoetra	} Est moyenne altitude

Observations

— Seuls les contours des polygones sont représentés : les arbres tombant à l'intérieur ne sont pas figurés pour augmenter la clarté.

— Fantsilotra, Alampona, et Ayous interpénètrent le domaine de l'Okoumé mais le Fantsilotra est plus *sensible*, l'Ayous-Samba moins *sensible* mais légèrement moins *solide* que l'Okoumé.

L'Alampona semble très comparable à l'Ayous mais sa dispersion est mal connue par deux arbres seulement.

Fantsilotra et Alampona méritent des essais de déroulage : le premier étant un arbre abondant dans le Sud, mais de dimensions modestes, pourrait être essayé en caisserie légère.

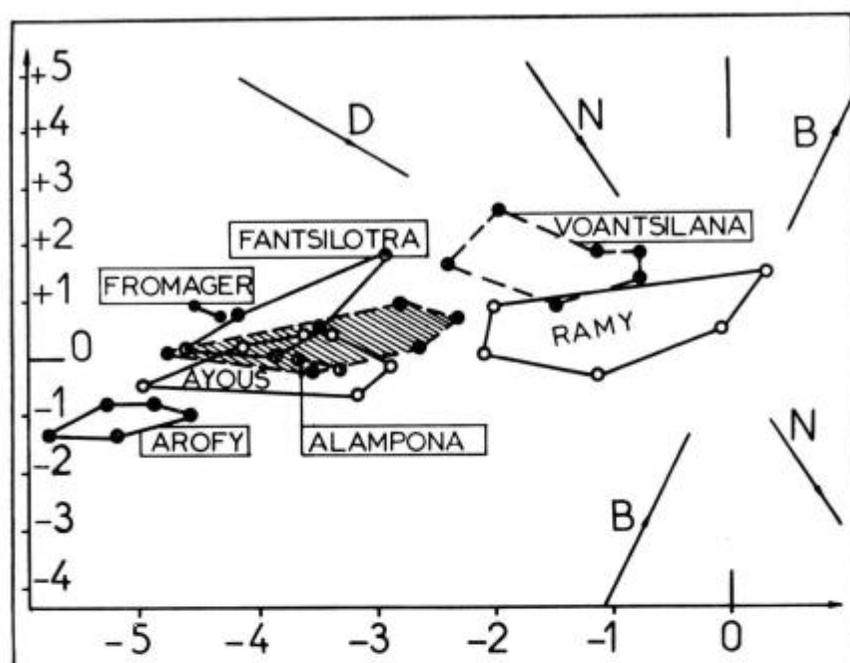


FIG. 7. — *Ayous (Samba) - Alampona - Fantsilotra - Voantsilana*

FIG. 7. — *Ayous (Samba) - Alampona - Fantsilotra - Voantsilana*

— Le groupe Voantsilana, bien que de composition botanique assez hétéroclite, est très acceptable sur le plan pratique sous ce nom pilote, étant technologiquement assez homogène.

Un peu moins *solide* que le Ramy mais nettement plus que l'Okoumé et l'Ayous, mais aussi plus *sensible*, le Voantsilana pourrait être essayé en tranchage ou déroulage pour caisserie légère, et les produits obtenus seraient nettement plus *solides* que ceux de l'Ayous par exemple. L'emploi en caisserie massive classique paraîtrait également judicieux.

— Fromager : non utilisé dans l'analyse, ses coordonnées Y_1 et Y_2 ont été calculées a posteriori (sur 2 arbres). Ils se placent en-dessous de l'Okoumé en *solidité*, légèrement inférieurs aussi à l'Ayous, mais surtout nettement plus *sensibles* que ces deux bois. Le Fantsilotra serait à tous égards intermédiaire entre Fromager et Okoumé.

— Structure interne du groupe Ayous-Samba : les quatre arbres récoltés sous le nom de Samba (origine Côte d'Ivoire) se trouvent groupés près de la frontière supérieure du polygone limitant l'ensemble du groupe : ils sont nettement à l'intérieur du polygone Okoumé, tandis que les Ayous (Côte d'Ivoire-Cameroun-R.C.A.) sont en-dessous, c'est-à-dire dans la direction des *sensibilités* plus faibles. Ces Ayous paraissent par ailleurs être de *solidité* plus variable que les Samba de Côte d'Ivoire. Ces observations s'appuient sur un faible nombre d'essais et pourraient être précisées sur un échantillonnage plus important.

3.3. — Hafotra — Molanga — Iroko (fig. 8) — Vory (fig. 9)

Nom vernaculaire	Échantillon utilisé — Identité	Provenance
Hafotra (11 arbres)	<i>Dombeya acerifolia</i> <i>Dombeya lucida</i> <i>Dombeya sp.</i> <i>Grewia bridellaefolia</i>	Beparasy Antoetra Fierenana Morafeno Périnet Ranomafana
Molanga (3 arbres)	<i>Croton mongue</i>	Anosibe Périnet
Iroko (8 arbres)	<i>Chlorophora excelsa</i> <i>Chlorophora sp.</i>	Gabon R. C. A. Côte d'Ivoire Cameroun Dahomey
Vory (fig. 9)	<i>Allaeanthus greveanus</i>	Maromiandra Ouest

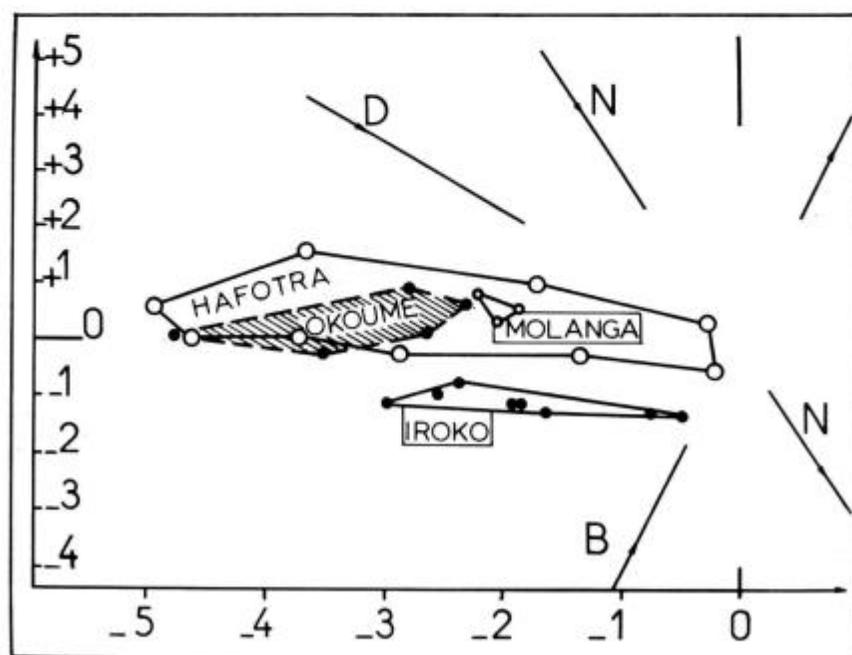


FIG. 8. — Hafotra - Molanga - Iroko

FIG. 8. — Hafotra - Molanga - Iroko

Observations

— Structure interne du groupe Hafotra :

La dispersion de ce groupe est grande. Il peut se scinder en *Dombeya lucida* + *Grewia brideliaefolia*, (moitié droite du polygone) et *Dombeya acerifolia* + *Dombeya sp.*, (moitié gauche).

Par ailleurs l'extrême droite du domaine est occupée par les *D. Lucida* provenant de Fierenana (7 arbres).

— Position du groupe Hafotra par rapport à l'Okoumé :

Le groupe Hafotra, pris au sens le plus large (*D. Acerifolia* + *D. Lucida* + *Grewia brideliaefolia*) recouvre et déborde l'Okoumé. Mais des deux sous-groupes composants, aucun ne s'ajuste à l'Okoumé : l'un est trop à gauche (moins *solide*), l'autre trop à droite (plus *solide*) et en son sein, les arbres provenant du massif de Fierenana sont les plus *solides*.

Parmi les Hafotra, certains arbres donnent donc un bois intéressant mais l'ensemble est trop dispersé pour avoir une vocation industrielle unique et malheureusement les moyens pratiques de sélection, par l'identité botanique, sont impuissants à faire un tri efficace.

On peut dire cependant que les *Dombeya lucida* d'origine Fierenana s'orienteraient vers l'utilisation en menuiserie, plus que le reste du groupe.

— Molanga : intermédiaire typique du point de vue *solidité* entre Okoumé et Ramy et équivalente aux deux en *sensibilité*, l'espèce paraît technologiquement très homogène. Son succès en déroulage pour la fabrication d'allumettes, prouvé par l'expérience, est donc très logique; c'est une vérification a posteriori.

— Iroko — Vory (fig. 8 et 9) :

Sont rapprochés ici parce que le Vory a été longtemps considéré comme un *Chlorophora* avant d'être nommé *Allaeanthus*. Les deux espèces se séparent bien, l'Iroko étant plus *solide*. Elles ont en commun une *sensibilité* remarquablement faible, surtout l'Iroko.

Le Vory est, en *solidité*, comparable à l'Okoumé, et sa *sensibilité* est plus faible; il serait donc dans les mêmes emplois, plus stable.

3.4. — Hazomalanga — Hazondrano — Volomborona — Voanana (fig. 9)

Nom vernaculaire	Échantillon utilisé — Identité	Provenance
Hazomalanga (2 arbres)	<i>Hazomalania voyroni</i>	Sud-Ouest
Hazondrano (5 arbres)	<i>Ilex mitis</i>	Morafenobe } Est Ranomafana } moyenne Périnet } altitude
Volomborona (3 arbres)	<i>Albizia gummifera</i>	Beparasy } Est moy. altitude
Voanana (12 arbres)	<i>Sloanea rhodantha</i>	Anosibe } Est Antoetra } moyenne Périnet } altitude Fierenana } Morafeno } Ranomafana }

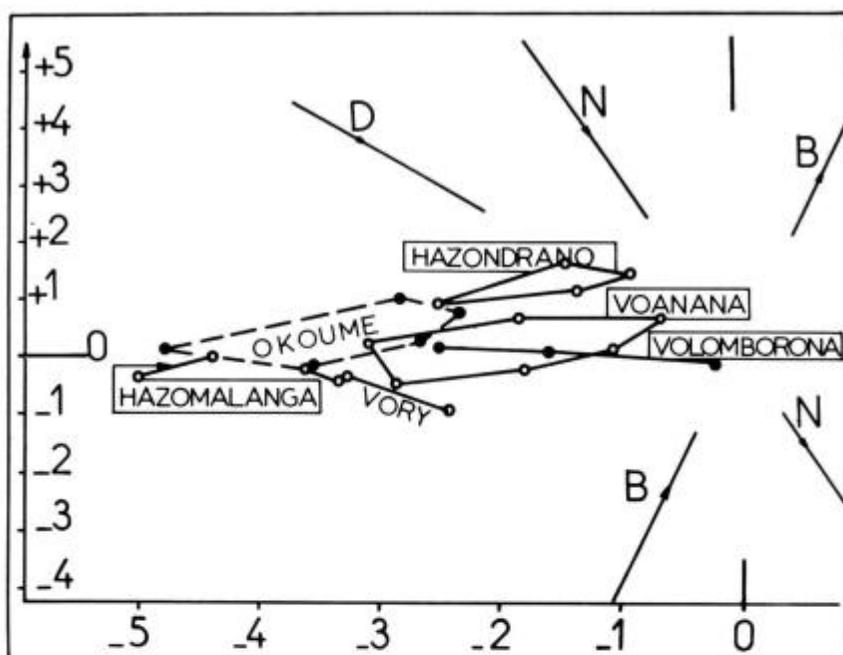


FIG. 9. — Hazomalanga - Hazondrano - Volomborona - Voanana - Vory

FIG. 9. — Hazomalanga - Hazondrano - Volomborona - Voanana - Vory

Observations

— Hazomalanga : intermédiaire à tous égards entre Arofy et Okoumé, assimilable aux Ayous les moins *solides*, de sensibilité faible, l'Hazomalanga, en placage ou en lattes massives, peut vraisemblablement être incorporé à l'intérieur de divers types de panneaux. Il semble trop peu *solide* pour les emplois en menuiserie où on le trouve parfois localement.

— Volomborona : intermédiaire entre Ramy et Okoumé, semble pouvoir être mélangé au premier dans certaines fabrications, menuiseries intérieures en particulier. Si son homologue de l'ouest, le Sambalahy (qui n'a pas fait l'objet d'essais) lui ressemble, ce qui paraît établi par l'usage, on pourra classer quatre essences souvent citées ensemble car relativement abondantes dans tout l'ouest : par ordre de *solidité* croissante, Arofy — Hazomalanga — Vory — Sambalahy, le seul relativement assimilable à l'Okoumé restant le Vory.

Le Volomborona est vraisemblablement utilisable en caisserie légère ou classique.

— Voanana : se classe aussi entre Ramy et Okoumé. Des essais de déroulage peuvent être recommandés, en raisons des dimensions atteintes habituellement par ces arbres, et aussi de leur couleur et de leur veinage. L'emploi en caisserie est également prévisible.

— Hazondrano : dans la même zone de *solidité* que le Voanana, le Hazondrano se signale par une forte *sensibilité* qui semble limiter son utilisation à la caisserie.

3.5. — *Arina* — *Baby* — *Tafanala* — *Varongy* (fig. 10)

Nom vernaculaire	Échantillon utilisé — Identité	Provenance
Arina (5 arbres)	<i>Bridelia tulasneana</i>	Ranomafana } Est Périnet } moyenne altitude
Baby ou Mandravoky (2 arbres)	<i>Anthostema madagascariensis</i>	Manantenina } Sud-Est
Tafanala (5 arbres)	<i>Terminalia tetrandra</i>	Périnet } Est moyenne altitude
Varongy (17 arbres)	<i>Ocotea auriculata</i> <i>Ocotea cymosa</i> <i>Ocotea faucherei</i> <i>Ocotea platydisca</i> <i>Ocotea laevis</i> <i>Ocotea trichophlebia</i> <i>Ocotea sp.</i>	Antoetra } Périnet } Est Morafeno } moyenne Moramanga } altitude Ranomafana } Fierenana } Fandradava } Beparasy } Anosibe }
Varongy (5 arbres)	<i>Ocotea macrocarpa</i>	Beparasy } Est Moramanga } moyenne altitude

Observations

— Problème des Varongy.

Le commerce mélange actuellement tous les *Ocotea* (espèces échantillonnées et de nombreuses autres) sous le nom de Varongy. Les utilisateurs les plus consciencieux ont quelques scrupules sur la légitimité de ce mélange mais jusque là aucune réponse précise n'a pu être donnée. La présente analyse tranche clairement la question : les espèces *auriculata*, *cymosa*, *faucherei*, *laevis*, *platydisca*, *trichophlebia* peuvent être mélangées sous le nom pilote de Varongy, mais l'espèce *O. macrocarpa* en est radicalement différente.

L'écart entre les deux polygones Varongy et *Ocotea macrocarpa* est spectaculaire. Il montre que des bois botaniquement très différents des *Ocotea*, tels Arina, Baby, Tafanala auraient des vocations voisines (menuiserie), alors que l'*Ocotea macrocarpa* est totalement rejeté de cette catégorie d'emplois pour rejoindre d'autres groupes étudiés plus loin, et composés de bois beaucoup plus solides.

— Arina — Baby — Tafanala

Leur proximité du Varongy nécessite d'être nuancée. Dans cet ordre ils sont de *solidité* croissante et de *sensibilité* décroissante. On aborde ici une zone du plan où les vocations sont moins claires qu'ailleurs parce qu'elle renferme un très grand nombre d'espèces variées et que toutes leurs propriétés se rapprochant d'une certaine « moyenne », les vocations de ces bois peuvent raisonnablement être multiples, comme celles de certaines essences utilisées couramment par plusieurs industries : le Sipo par exemple, qui est justement dans cette zone, et que nous représentons plus loin.

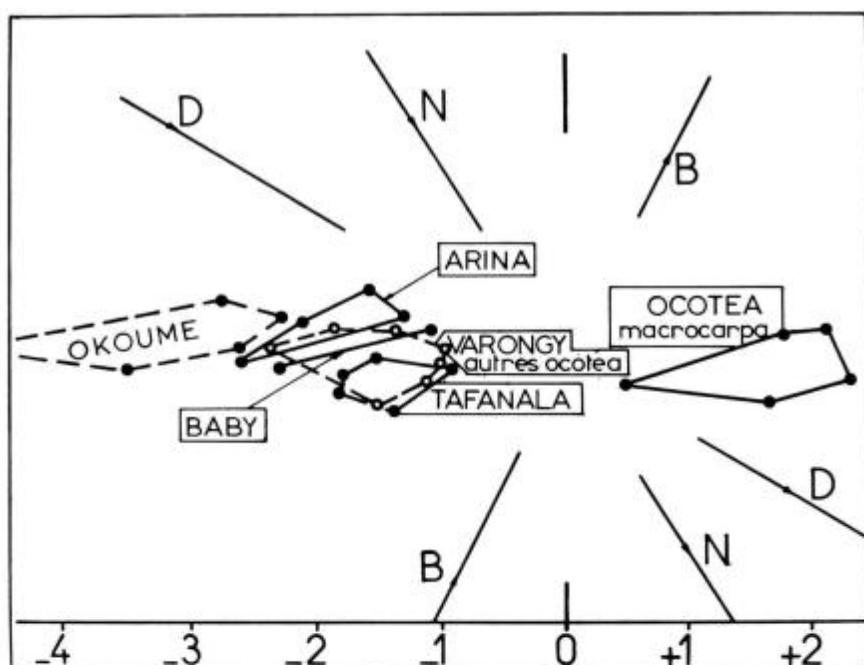
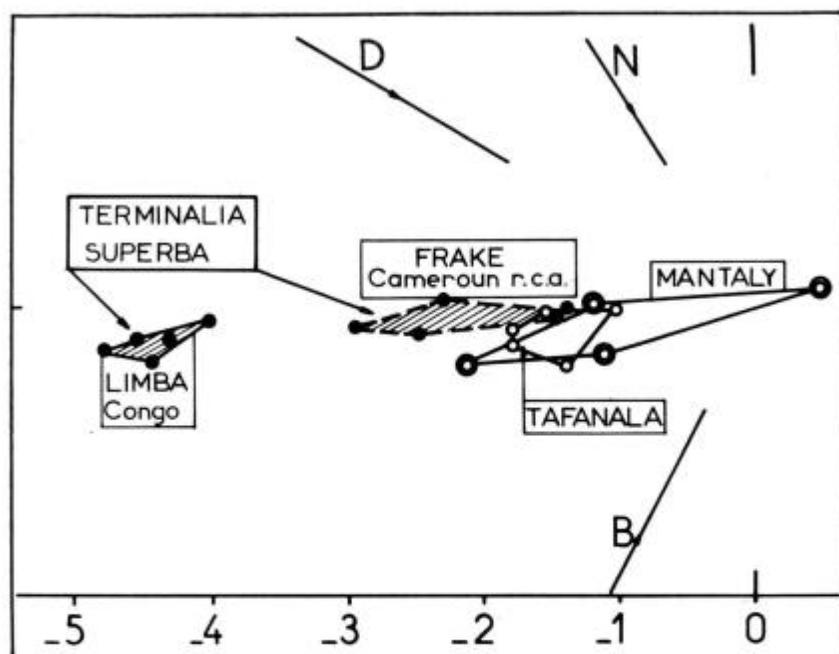


FIG. 10. — Arina - Baby - Tafanala - Mantaly

FIG. 10. — Arina - Baby - Tafanala - Mantaly

3.6. — Limba — Frake — Akom — Tafanala — Mantaly (fig. 11)

Nom vernaculaire	Échantillon utilisé — Identité	Provenance
Limba (6 arbres)	<i>Terminalia superba</i>	Congo (5) R.C.A. (1)
Frake (4 arbres)	<i>Terminalia superba</i>	Cameroun
Tafanala (5 arbres)	<i>Terminalia tetrandra</i>	Madagascar
Mantaly (6 arbres)	<i>Terminalia mantaly</i>	Madagascar

FIG. 11. — *Limba - Frake - Akom - Tafanala - Mantaly*FIG. 11. — *Limba - Frake - Akom - Tafanala - Mantaly*

Observations

L'analyse devrait être reprise sur un échantillonnage plus abondant mais sur les dix *T. superba* utilisés ici il apparaît clairement que le Limba du Congo se sépare du reste, et différence entre les deux sous-groupes ainsi constitués est uniquement dans la *solidité* (inférieure à celle de l'Ayous pour le Limba du Congo).

L'un et l'autre sous-groupes sont moins sensibles que l'Okoumé et l'encadrent en *solidité*.

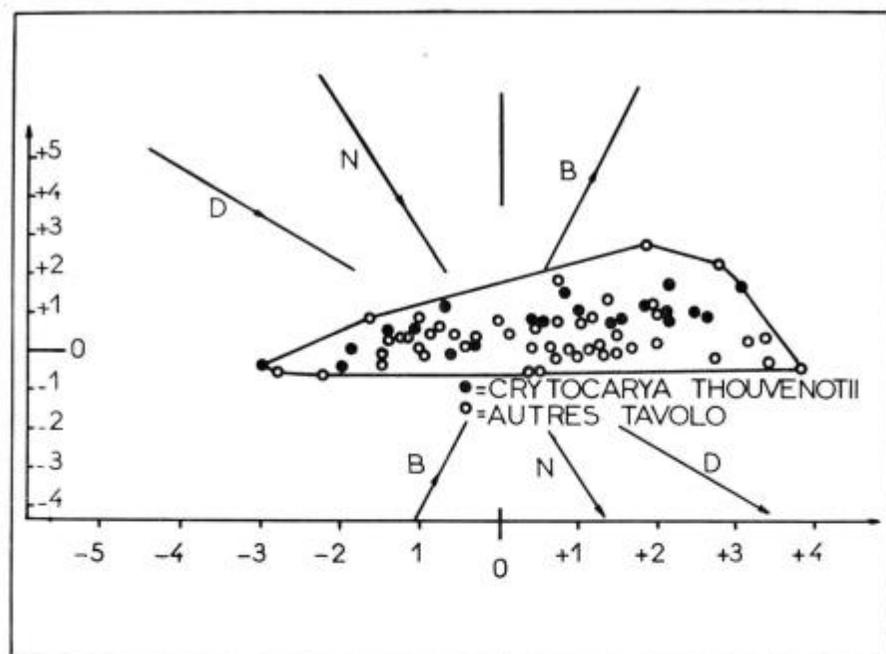


FIG. 12. — Tavolo - Situation générale

FIG. 12. — Tavolo - General outlook

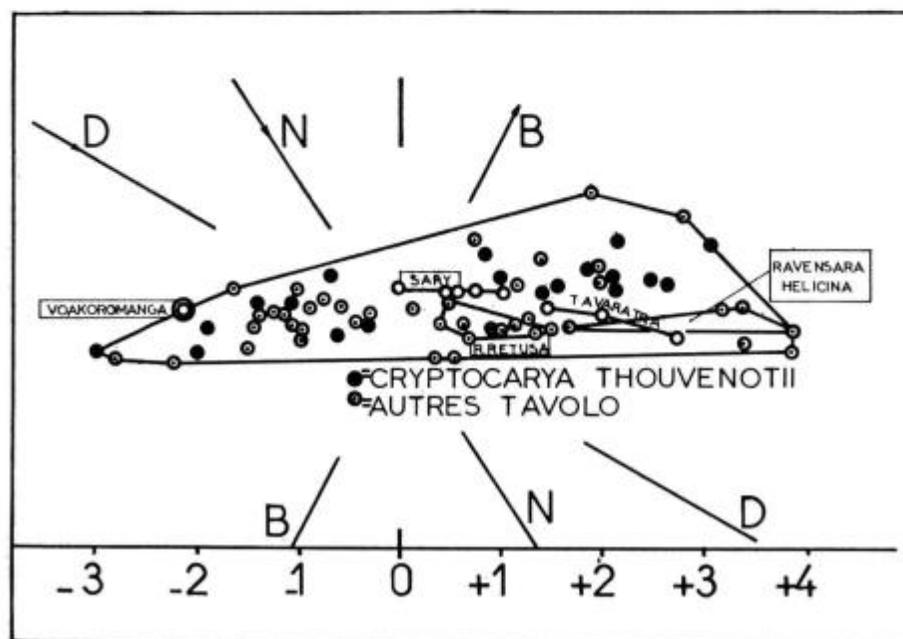


FIG. 13. — Tavolo - Sous-groupes

FIG. 13. — Tavolo - sub-groups

Observations

Structure du groupe du point de vue de l'identité

— Sous le nom de Tavolo, on a identifié à peu près autant de *Cryptocarya* que de *Ravensara*, et les variantes de nomenclature vernaculaire (Tavolo Manitra, etc.) ne présentent pas de relation avec l'identité réelle. Les genres *Beilschmiedia* et *Potameia* sont parfois reconnus par les bûcherons comme étrangers au groupe. Le botaniste lui, les sépare bien. Mais il rapproche étroitement les diverses espèces de *Cryptocarya* entre elles, de même que les espèces de *Ravensara* entre elles et parfois ne peut pas se prononcer sur les herbiers stériles entre genre *Ravensara* et genre *Cryptocarya*.

— Aucune station n'a fourni plus spécialement un genre ou l'autre, *Cryptocarya* et *Ravensara* étant très généralement représentés ensemble.

— L'ensemble du groupe occupe une place importante dans le volume de bois sur pied dans la forêt orientale de moyenne altitude.

Structure du groupe dans la représentation graphique

— Dispersion extraordinaire en *solidité* (— 3 à + 4)
très élevée aussi en *sensibilité* (— 0,5 à + 3)

— on peut extraire assez facilement :

Beilschmiedia velutina, bois plus faible que le reste du groupe en *solidité* (confirmé par des essais postérieurs à l'analyse), que l'on peut individualiser sous le nom pilote de Voakoromanga, utilisable en menuiserie. Signalons un service que rend l'analyse : elle attire l'attention sur les espèces intéressantes et insuffisamment échantillonnées et oriente le planning des récoltes ultérieures.

Beilschmiedia oppositifolia : *solidité* comprise entre 0 et + 1 donc moyenne pour le groupe, et *sensibilité* voisine de 0,75. Peut être individualisé sous le nom pilote de Sary, utilisable en charpente (sous réserve de traitement de préservation).

Potameia antevaratra : *solidité* élevée (entre + 1,5 et + 2,8) *sensibilité* moyenne (0 à + 0,6). Peut être séparé sous le nom pilote de Tavaratra et être utilisé en charpente forte (sous réserve de traitement de préservation).

Ces trois sous-groupes, moyennant une attention particulière d'identification, sont individualisables dans la pratique.

Par contre, les suivants, pourtant aussi nets sur échantillons identifiés par herbier complet, sont déjà beaucoup plus délicats à faire passer dans l'usage car les confusions d'identité à la réception (sur examen visuel du bois) sont très probables en l'absence d'herbier. Sont dans ce cas les sous-groupes suivants :

Ravensara retusa : bien groupés en *solidité* entre + 0,5 et + 1,5 et en *sensibilité* entre 0 et 0,75, conviendraient à la charpente (sous réserve de traitement de préservation), et pourraient sans inconvénient être employés comme *Beilschmiedia oppositifolia* (Sary).

Ravensara helicina : bien groupés vers les *solidités* élevées, conviendraient (sous réserve de traitement de préservation) à la charpente forte et pourraient être employés comme *Potameia antevaratra* (Tavaratra).

— le total de ces « extractions », dont l'application pratique est déjà discutable à cause des difficultés d'identification, ne concerne que 16 arbres sur 72, cinq espèces sur une vingtaine composant le groupe, et malheureusement pas des plus fréquentes en forêt.

— *les espèces restantes*, dont la dispersion n'est aucunement diminuée par les extractions, *semblent désespérément inutilisables industriellement* : à elle seule l'espèce unique *Cryptocarya thouvenotii* fournit, suivant l'arbre abattu (dans une même station), un bois voisin du Frake, du Hafotra ou bien un bois plus *solide* et plus *sensible* que l'*Eucalyptus robusta* ou le Vintanina, avec tous les intermédiaires possibles.

Dès lors il est inutile d'observer que certains Tavolo avoisinent une espèce « sage » comme le Niangon puisque dans un même lot commercial on trouvera en même temps des pièces plus dures et plus lourdes que certains Nato (bois incomparablement plus lourd, plus dur, plus *solide* et plus *sensible* que le Niangon).

Devant l'ampleur des conséquences de ces graves constatations l'analyse a été poussée plus loin; on pourrait penser que les artifices du calcul mathématique soient responsables de la dispersion étonnante constatée, puisque l'analyse factorielle étale au maximum l'échantillonnage total pour optimiser l'observation des individus qui le composent. Malheureusement il n'en est rien puisque :

— certains groupes pourtant eux aussi botaniquement hétérogènes, se trouvent graphiquement bien concentrés (ex. : Voantsilana = *Cussonia* + *Polyscias* + d'autres Araliacées);

— la dispersion des Tavolo dans la représentation graphique en composantes principales est le reflet d'une dispersion intrinsèque, très profonde, des propriétés physiques et mécaniques de base (caractères initiaux retenus pour l'analyse). Les histogrammes de distribution de ces propriétés de base sont formels : l'étendue de chaque propriété étant divisée en 20 classes pour l'ensemble de tous les bois tropicaux étudiés, les Tavolo « s'éparpillent » :

en dureté : de la classe 1 à la classe 6,

en densité : de la classe 6 à la classe 14,

en rétractibilité volumétrique : de la classe 5 à la classe 14,

en résistance en flexion : de la classe 4 à la classe 13,

en module d'élasticité : de la classe 5 à la classe 16.

Remarque :

Si l'on regarde la dispersion à l'intérieur de chaque espèce et non plus dans l'ensemble du groupe, hormis les trois *Cryptocarya*, celle de chaque espèce de *Ravensara* peut être considérée comme normale mais c'est là une précision de peu d'intérêt réel puisque les espèces sont indiscernables dans un lot de bois façonnés.

Suggestion pratique :

Dans les permis d'exploitation, rendre obligatoire l'abattage des Tavolo éventuellement sans redevance, pour qu'ils laissent la place, dans la génération suivante, à des essences de lumière, occupatrices de places vides, dont certaines peuvent être intéressantes (ex. Molanga).

3.8. — Niangon — Ogoûe (fig. 14)

Nom vernaculaire	Échantillon utilisé — Identité	Provenance
Niangon (10 arbres)	<i>Tarrietia utilis</i>	Côte d'Ivoire
Ogoûe (10 arbres)	<i>Tarrietia densiflora</i>	Gabon

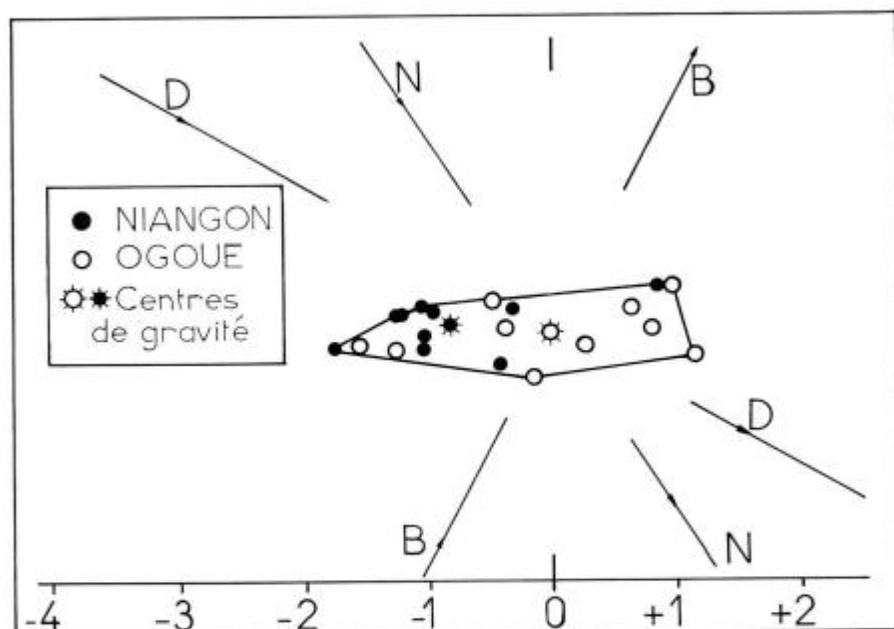


FIG. 14. — Niangon - Ogoûe

FIG. 14. — Niangon - Ogoûe

Observations

Niangon et Ogoûe sont considérés comme des bois « voisins » par certains utilisateurs. Que peut-on en dire ?

— Seul est représenté le contour extérieur global des deux espèces, dont les domaines propres sont très imbriqués.

— Ce contour global est à peine plus étendu que les domaines propres à chacune des deux espèces regardées séparément.

— On peut donc valablement réunir les deux espèces dans un domaine unique.

— On pressent pourtant graphiquement que l'imbrication se fasse plutôt par les extrémités et que le Niangon se groupe plutôt dans la partie gauche du domaine commun, alors que l'Ogoûe paraît plutôt occuper la partie droite.

— Pour préciser cela on peut observer la position des centres de gravité (G) de chacun des deux nuages; celui de l'Ogoué (couleur claire sur le graphique), à droite, implique un bois en moyenne plus *solide* que le Niangon (à gauche).

Un simple test de conformité des moyennes effectué sur la variable *densité*, par exemple, a suffi à déceler une différence significative au niveau 5 % entre les deux espèces.

On a ici l'exemple typique de deux bois très voisins, pour lesquels les individus s'entremêlent, mais dont l'analyse montre cependant qu'ils sont « centrés » différemment.

— L'Ogoué se trouve représenter presque exactement « le » bois moyen de toute l'analyse : son centre de gravité est pratiquement confondu avec le centre de gravité de l'ensemble de tous les bois malgaches et africains analysés. C'est pourquoi nous faisons figurer le genre *Tarrietia* ici, bien qu'il ne soit pas représenté dans la flore malgache, pour son rôle de jalon et de bois moyen de référence, bien connu dans les industries européennes.

3.9. — *Kijy* (fig. 15)

Nom botanique	Provenance	
<i>Symphonia louveli</i> (12 arbres)	Anosibe Beparasy Fierenana Moramanga Périnet Ranomafana (F)	} Est moyenne altitude
<i>Symphonia tanalensis</i> (4 arbres)	Périnet	
<i>Symphonia microphylla</i> (2 arbres)	Antoetra	
<i>Symphonia fasciculata</i> (2 arbres)	Périnet	
<i>Symphonia sp.</i> (2 arbres)	Manantenina	
<i>Garcinia</i> (4 arbres)	Morafeno Périnet	} Est moyenne altitude

Observations

Le groupe des *Kijy* (= *Hazinina* en zone côtière) donne des bois jaunes d'aspects très voisins. Les diverses espèces se distinguent très difficilement sur l'herbier et pas du tout par simple examen visuel du bois. Les *Garcinia* se distinguent par contre des *Symphonia* à la fois sur bois et sur herbier.

— Analyse des *Symphonia*

Deux sous-groupes apparaissent nettement, de *sensibilité* équivalente, mais de *solidité* différente :

a) *Symphonia tanalensis*, auquel s'apparente *S. microphylla* : sous-groupe de *solidité* plus faible.

b) *Symphonia louveli*, auquel s'apparentent les deux échantillons *Symphonia sp.* (du côté des fortes *solidités* et fortes *sensibilités*) et aussi les deux *S. fasciculata* (du côté des plus faibles *solidités*).

— Genre *Garcinia* :

Intermédiaire entre les deux sous-groupes (a) et (b) de *Symphonia*, le genre botanique *Garcinia* peut donc très raisonnablement être inclus dans les Kijy.

— Vocations

Tous ces bois ont une *sensibilité* élevée, et sont fort étalés en *solidité*. Ils peuvent donc fournir de la charpente (ordinaire pour le sous-groupe (a), lourde pour le sous-groupe (b)) et à la rigueur de la menuiserie forte d'intérieur malgré leur *sensibilité* qui les exclut cependant de la parqueterie.

— Note sur la rétractibilité des Kijy

L'analyse factorielle attire l'attention sur la forte *sensibilité* du groupe Kijy. Comme cette *sensibilité* est une combinaison des caractéristiques physiques où prédominent les variables de rétractibilité, il est indiqué de vérifier que ces variables (B, T + R, et v) ont effectivement des valeurs élevées pour les Kijy. Cela se révèle bien exact : dans l'histogramme de rétractibilité volumétrique totale (B), on trouve les Kijy dans les classes 9 à 17, comme on trouve les *Eucalyptus robusta* (espèce bien connue pour sa forte rétractibilité) dans les classes 9 à 16.

Les Kijy se comportent aussi mal que les Eucalyptus Robusta du point de vue du « jeu ».

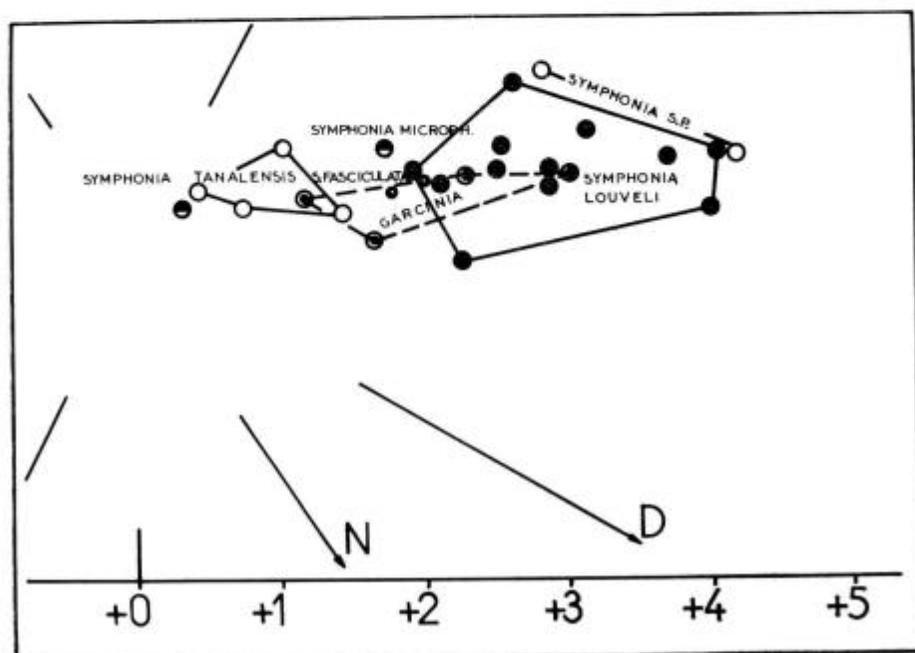


FIG. 15. — Kijy
FIG. 15. — Kijy

3.10. — Doussié — Lingué — Hintsy (fig. 16)

Nom vernaculaire	Échantillon étudié — Identité	Provenance
Doussié (8 arbres)	<i>Azelia pachyloba</i> <i>Azelia bipendensis</i>	Cameroun Cameroun
Lingué (5 arbres)	<i>Azelia africana</i>	Côte d'Ivoire R.C.A.
Hintsy (3 arbres)	<i>Intsia bijuga</i>	Est basse altitude

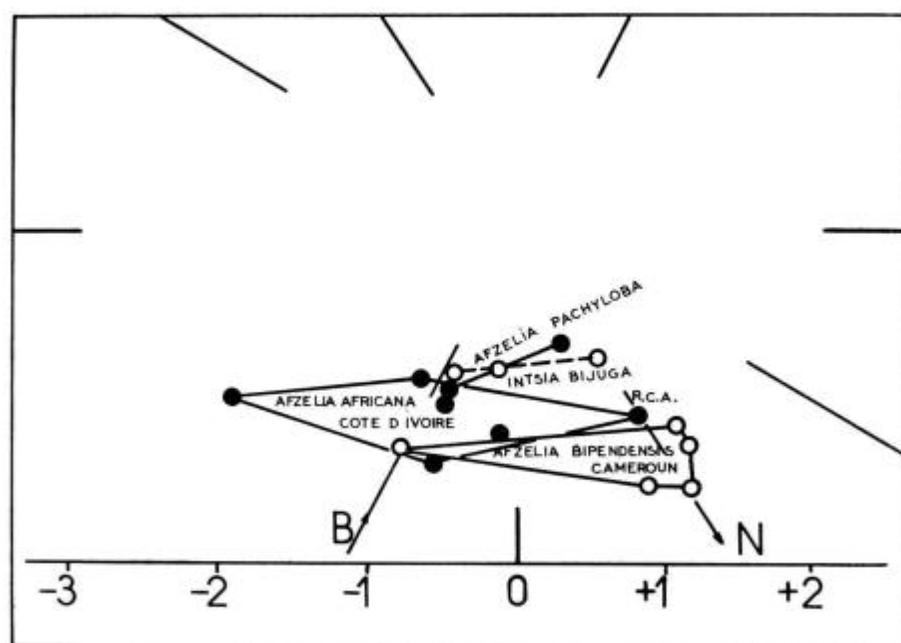


FIG. 16. — Doussié - Lingué - Hintsy
 FIG. 16. — Doussié - Lingué - Hintsy

Observations

Tous ces bois présentent une *sensibilité* exceptionnellement faible (forte *dureté*, *rétractibilité* très basse) ce qui est une situation très favorable pour les emplois en parquets, menuiserie forte d'extérieur ou d'intérieur, bardeaux, constructions navales.

Structure interne du groupe.

On voit se séparer :

— *Azelia africana* de Côte d'Ivoire (Lingué), le moins *solide*.

— *Afzelia bipendensis* du Cameroun (dont se rapproche beaucoup *Afzelia africana* de R.C.A.), sous-groupe le plus *solide* et le moins *sensible*.

— *Afzelia pachyloba* du Cameroun et *Intsia bijuga* de Madagascar, très imbriqués, de *solidité* intermédiaire et de *sensibilité* légèrement plus forte que les *A. africana* et *A. bipendensis*.

Doussié du Cameroun et Hintsy de Madagascar peuvent être mélangés dans les mêmes fabrications industrielles sans inconvénient.

IV. — ZONES D'UTILISATION INDUSTRIELLE DES BOIS DANS LE PREMIER PLAN PRINCIPAL

Les quelques exemples précédents montrent des utilisations possibles de la représentation plane :

— la *position* d'un polygone d'espèce ou de groupe dans le plan permet de synthétiser les propriétés de l'espèce ou du groupe et d'en décrire les propriétés globales,

— les *dimensions* d'un polygone rendent compte de la dispersion des propriétés,

— la *forme* indique éventuellement si cette dispersion est plus accentuée pour l'une des composantes principales, ou dans une direction oblique interprétable,

— l'*identification* graphique des différentes espèces constituant un groupe et de leurs provenances géographiques explique parfois l'allure de la dispersion observée et suggère des discriminations à faire dans le domaine de l'utilisation, ou confirme des discriminations déjà faites par les utilisateurs, ou encore permet la suppression de certaines de ces discriminations empiriques,

— la *comparaison* de plusieurs polygones sur la même représentation introduit la notion de proximité, de *ressemblance de plusieurs bois*.

C'est à partir de cette dernière façon de raisonner que nous donnerons ci-dessous quelques exemples de rapprochements multiples et simultanés, sur le même graphique, d'un nombre important de bois différents. La représentation complète de chaque bois par tous les arbres qui composent son échantillonnage rend la superposition graphique de plusieurs espèces assez rapidement illisible; même en représentant seulement le contour du polygone qui ceinture chaque bois, l'enchevêtrement des espèces est complexe.

C'est pourquoi l'on utilise dans ce qui suit une représentation simplifiée de chaque bois, par le centre de gravité de son échantillon, qui représente pour les propriétés, la moyenne d'espèce ou de groupe. La perte d'information ainsi consentie (la dispersion interne aux groupes n'est plus représentée) est compensée par un gain en clarté. Cependant cette simplification n'a été adoptée que pour les groupes ou espèces dont l'homogénéité raisonnable a été vérifiée préalablement sur la représentation complète.

* * *

On peut ainsi matérialiser dans divers cas un nuage de points (moyennes d'espèces ou de groupes) correspondant tous à un même emploi industriel et trouver pour certaines zones du plan une étiquette assez précise d'emploi.

Il est évident que chaque emploi présente des tolérances vis-à-vis des propriétés qu'il réclame : la zone graphique correspondante a donc une certaine surface. On peut en présumer les frontières approximatives en utilisant des bois-jalons connus de l'industrie.

Par exemple l'Ayous, qui n'a pas été reçu immédiatement ni sans quelque difficulté dans le commerce des bois de déroulage, peut être pris comme jalon de *solidité minimum* requise pour cet emploi. De même il est possible de localiser la frontière des bois trop *solides* et celle des bois trop *sensibles* pour le déroulage.

Rappelons que pour un emploi où des critères spéciaux interviennent, critères non incorporés dans l'analyse factorielle souvent parce que de caractère qualitatif (couleur, grain du bois, forme des arbres, aire géographique de l'espèce), on introduit soi-même la sélection nécessaire : c'est le cas par exemple de la séparation des bois aptes à fournir des placages déroulés décoratifs au sein de la population des bois simplement aptes technologiquement au déroulage.

On peut également noter que des emplois différents correspondent souvent à des surfaces qui se recouvrent partiellement dans le premier plan principal (et à des volumes imbriqués dans l'espace d'origine à 9 dimensions); ce fait traduit simplement la multiplicité des vocations pratiques convenant à un même bois.

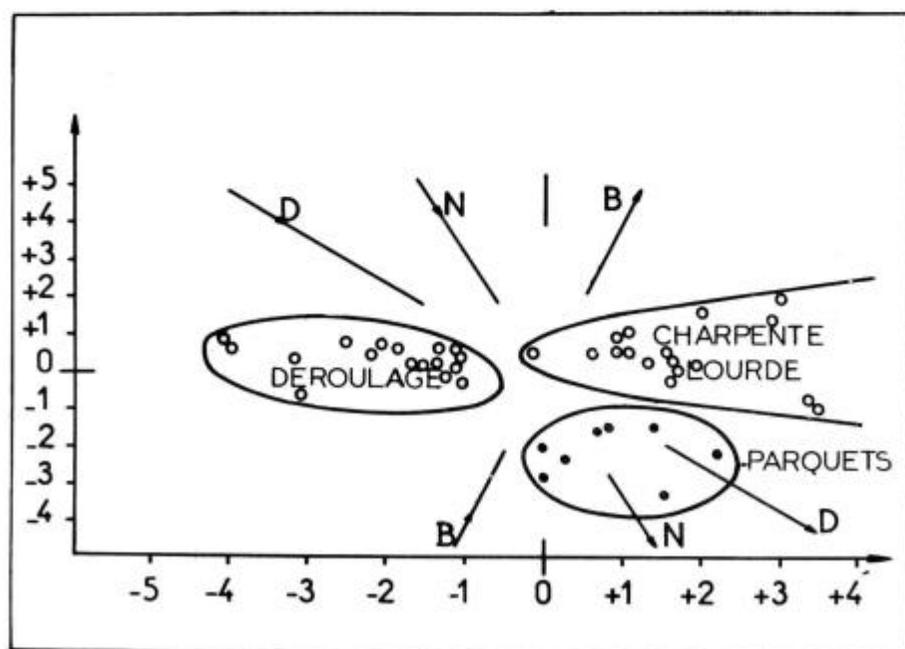


FIG. 17. — Zones d'utilisation industrielle des bois dans le premier plan principal

FIG. 17. — Zones of industrial uses of timbers in the first principal plane

Il est également logique d'admettre que des aires d'emploi ne soient limitées que dans certaines directions; par exemple, il n'y a pas de limite inférieure de *sensibilité* à la zone des bois à parquets (alors que la limite supérieure est primordiale), puisque la qualité d'avoir une faible *sensibilité* peut être poussée à l'extrême et avec le plus grand bénéfice, pour cet emploi. De même la limite supérieure de *solidité*, pour les bois de charpente forte, peut être choisie très soupagement car l'excès de *solidité* ne devient défaut que lorsqu'il s'assortit d'une impossibilité matérielle d'usinage.

Pour ces raisons on observe généralement que les aires d'emploi situées dans la partie droite du plan (fortes *solidités*) sont plus étirées et plus vastes que celles de la zone médiane du plan, où les bornes supérieure et inférieure en *solidité* sont plus précises et plus serrées.

Nous allons examiner trois emplois à titre d'exemple : placage déroulé, parquets, charpente lourde (fig. 17).

4.1. — Bois malgaches présumés aptes à fournir des feuilles de placages déroulés (fig. 18)

1 - <i>Isolona</i> sp.	Ambavy
2 - <i>Alluaudia procera</i>	Fantsilotra
3 - <i>Hibiscus lasiococcus</i>	Alampona
4 - <i>Allaeanthus greveanus</i>	Vory
5 - <i>Pinus khasya</i>	Pin d'Indochine
6 - <i>Pinus patula</i>	Pin Argente
7 - <i>Beilschmiedia velutina</i>	Voakoromanga
8 - <i>Croton mongue</i>	Molanga
9 - <i>Bridelia tulasneana</i>	Arina
10 - <i>Sloanea rhodantha</i>	Voanana
11 - <i>Anthostema madagascar.</i>	Mandravoky
12 - <i>Dombeya lucida</i>	Hafobalo
13 - <i>Terminalia tetrandra</i>	Tafanala
14 - <i>Albizia gummifera</i>	Volomborona
15 - <i>Ocotea</i> sp. p.	Varongy
(sauf <i>Ocotea macrocarpa</i>)	
16 - <i>Populus deltoides</i>	Peuplier
17 - <i>Canarium madag.</i>	Ramy
18 - <i>Zanthoxylum</i> sp.	Fahavalonkazo
19 - <i>Weinmannia minutiflora</i>	Lalona
20 - <i>Astrotrichilia</i> sp.	Hetakoaka

Observations

Les bois africains suivants, déjà utilisés dans l'industrie du déroulage, ont été utilisés comme « jalons » : *Ayous*, *Okoumé*, *Iomba*, *Tiama*, *Sipo*, *Bete*. Cet ordre les classe d'ailleurs par *solidité* croissante. Ces jalons sont notés en toutes lettres sur la figure 18, alors que les bois malgaches sont représentés, pour laisser au graphique une certaine clarté, par des numéros se rapportant à la nomenclature ci-dessus, dans laquelle l'ordre n'a pas de signification particulière.

C'est la première occasion que nous avons de nous faire une idée de la grandeur de la zone correspondant à un emploi particulier : on pensait a priori trouver, pour un usage aussi précis que le déroulage, une zone plus restreinte.

Le graphique laisse prévoir que des bois tels que *Terminalia tetrandra* (13) ou *Albizia gummifera* (14) puissent avoir un comportement voisin de celui du *Sipo* des points de vue physique et mécanique (les critères d'aspect n'étant pas incorporés dans le calcul ne sont évidemment pas traduits dans la position graphique). Des mêmes points de vue *Hibiscus lasiococcus* (3) se rapproche de l'*Okoumé*; dans ce cas particulier d'ailleurs la couleur, le grain et l'aspect du bois sont comparables à ceux de l'*Okoumé*.

Il est évident que le contour qui cerne ce groupe des bois de déroulage n'a pas en réalité de position, de forme, ni de dimensions aussi précises que l'ellipse figurant sur le graphique : celle-ci a seulement été tracée à titre indicatif, pour localiser visuellement la zone intéressée.

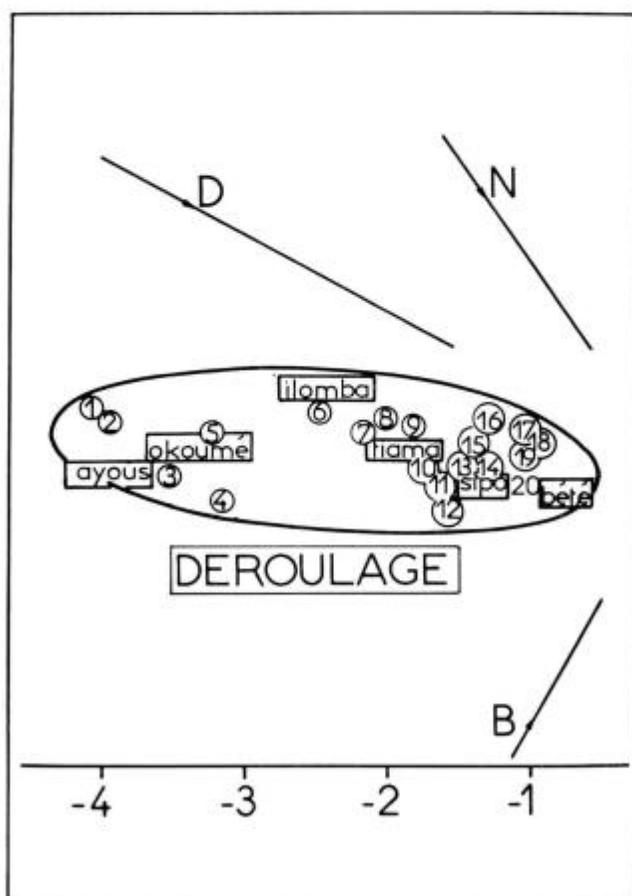


FIG. 18. — Bois de déroulage

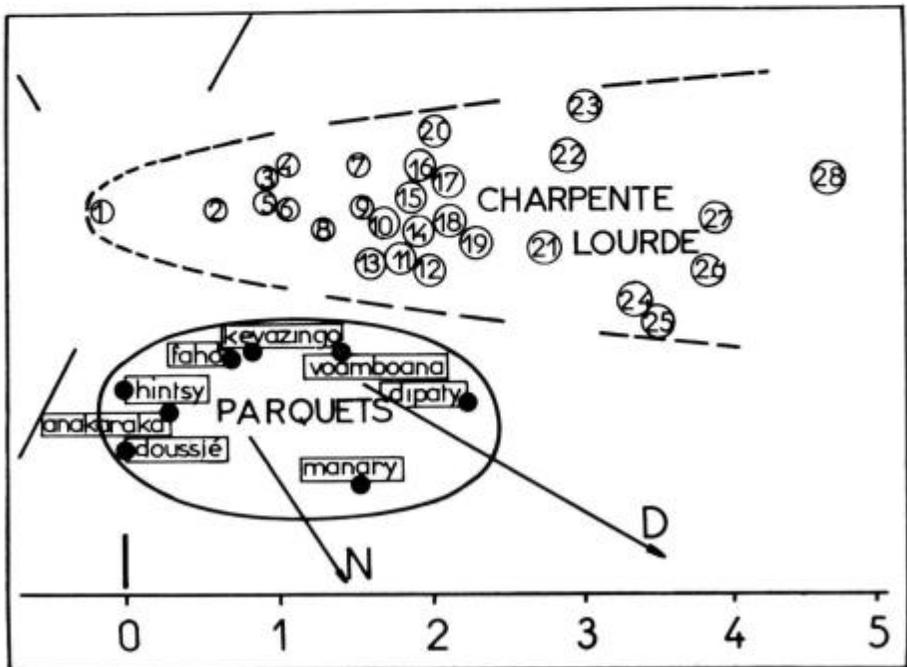
FIG. 18. — Peeling wood

4.2. — *Bois malgaches présumés aptes à fournir des parquets* (fig. 19)

<i>Intsia bijuga</i>	Hintsy
<i>Cordyla madagascariensis</i>	Anakaraka
<i>Chloroxylon faho</i>	Faho
<i>Dalbergia baroni</i>	Voamboana
<i>Dalbergia sp. s.</i>	Manary
<i>Pachytrophe dimepate</i>	Dipaty

Observations

Le Doussié et le Kevazingo ont été utilisés comme jalons et figurent sur le même graphique que les bois malgaches cités ci-dessus. Certains de ces bois, tel *Dalbergia baroni* (palisandre de la région orientale) ont effectivement fait leurs preuves auprès de nombreux architectes et entrepreneurs, et l'analyse ne fait que confirmer les connaissances empiriques; mais par contre le *Cordyla madagascariensis* se voit ici promu pour un emploi dans lequel il n'est pour l'instant pas utilisé.

FIG. 19. — *Bois à parquets - Bois de charpente lourde*FIG. 19. — *Flooring timber - Heavy carpentry timber*

4.3. — Bois présumés aptes à fournir de la charpente lourde (fig. 19)

1 - <i>Gambeya boiviniana</i>	Famelona
2 - <i>Protorhus ditimena</i>	Ditimena
3 - <i>Symphonia tanalensis</i>	Kijy
4 - <i>Symphonia microphylla</i>	Kijy
5 - <i>Uapaca</i> sp. p.	Voapaka
6 - <i>Syzygium</i> sp. p.	Rotra
7 - <i>Eucalyptus robusta</i>	Kininy
8 - <i>Cryptocarya scintillans</i>	Longotra
9 - <i>Neotina isoneura</i>	Ramaindafy
10 - <i>Noronhia</i> sp.	Vazanaomby
11 - <i>Weinmannia rutenbergiana</i>	Herehitsika
12 - <i>Ocotea macrocarpa</i>	Varongy
13 - <i>Androstachys</i> sp.	Merana
14 - <i>Calophyllum drouhardi</i>	Vintanina
- <i>Calophyllum parviflorum</i>	Vintanina
15 - <i>Breonia decidua</i>	Molompangady
- <i>Breonia keliravina</i>	Molompangady
- <i>Breonia perrieri</i>	Molompangady
- <i>Neonauclea foveolata</i>	Molompangady
16 - <i>Symphonia fasciculata</i>	Kijy
17 - <i>Garcinia</i> sp.	Kijy
18 - <i>Potameia antevaratra</i>	Tavaratra
19 - <i>Scolopia</i> sp.	Menavahatra
20 - <i>Mauloutchia</i> sp.	Mahafotra
21 - <i>Mammea bongo</i>	Vongo
22 - <i>Symphonia louveli</i>	Kijy
23 - <i>Elaeodendron</i> sp.	Hazomalany
24 - <i>Faucherea ambrensis</i>	Nato
- <i>Faucherea parvifolia</i>	Nato
25 - <i>Asteropeia rhopaloides</i>	Manoka
26 - <i>Rhodolaena bakeriana</i>	Fotona
27 - <i>Canthium</i> sp.	Vandrika
- <i>Craspidospermum verticillatum</i>	Vandrika
- <i>Pyrostria</i> sp.	Vandrika
28 - <i>Dilobeia thouarsii</i>	Vivaona

Observations

Le contour cernant cet emploi est laissé ouvert à droite, vers les fortes *solidités*, qui ne peuvent devenir un défaut, par leur excès, qu'au delà d'une valeur impossible à préciser. Les bois se trouvant à l'extrême droite deviennent d'ailleurs fort rares et forment des cas très particuliers. L'éventail devient également de plus en plus large en *sensibilité* : cette forme du

contour est à rapprocher d'une observation faite sur la forme semblable des *polygones d'espèces* rencontrés dans la zone des fortes *solidités*.

Tout se passe comme si ces valeurs exceptionnelles de la *solidité* étaient presque systématiquement accompagnées d'une plus grande dispersion en *sensibilité*. Il est difficile pour l'instant de savoir si cette représentation graphique provient réellement des propriétés intrinsèques du bois, qui auraient cette distribution, ou bien seulement d'une dispersion accrue des mesures de ces propriétés, dispersion qui serait alors liée aux modes opératoires et non au matériau.

V. — POSSIBILITÉS D'UTILISATION DES RÉSULTATS COMPLÉMENTAIRES FOURNIS PAR LE PROGRAMME

5.1. — Analyse de la distribution séparée de chaque variable initiale

Le programme utilisé fournit, préalablement au calcul et à la représentation graphique des composantes principales, des renseignements précieux sur les variables initiales, en particulier sur leur distribution.

Le test de normalité pratiqué sur chacune d'elles, et surtout l'histogramme basé sur 20 classes (cf. § 31) permettront de discuter par rapport à la distribution réelle, les limites des catégories actuellement utilisées : certains de ces rapprochements suggèrent de reconsidérer le découpage de ces catégories, parfois par un simple décalage des limites, parfois en changeant le nombre.

Ces révisions pourront concerner certaines des 9 variables initiales qui ont été retenues ici, mais le cas des variables mesurées en laboratoire et non utilisées dans la présente analyse (cf. chapitre I) peut être repris également; dans un souci d'allègement de ce premier calcul exploratoire on a en effet écarté en particulier deux catégories de variables (1) : celles qui semblaient avoir une dispersion réellement anarchique due à un mode opératoire discutable (exemple : résistance au fendage) et celles qui au contraire semblaient varier peu d'un bois à l'autre et ne fournir que des valeurs de correction (exemple : hygroscopicité à l'air pour corriger la densité en fonction de l'humidité). Dans ces deux cas la confection des histogrammes précisera si ce point de vue initial était justifié.

5.2. — Analyse simultanée de couples de variables initiales

La représentation ponctuelle des arbres (et polygonale des espèces) a été utilisée dans les plans de divers couples de variables. Cette représentation débouche sur deux sortes d'informations :

— dans les plans formés par deux variables initiales fortement corrélées (fig. 20)

(1) En plus des caractéristiques n'apportant pas d'information par elles-mêmes parce que non mesurées directement mais calculées à partir de variables mesurées (exemple : S calculé par $\frac{B}{V}$).

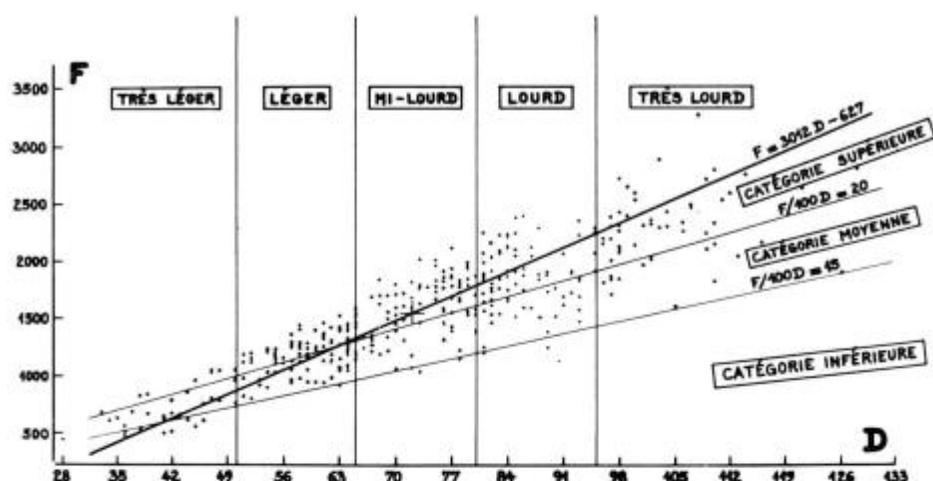


FIG. 20. — Le nuage dans le plan D - F

FIG. 20. — The cloud in D - F plane

Par exemple avec D et F pour lesquelles $r = 0,90$, on a une matérialisation de la liaison étroite entre les variables. L'idée a en effet été émise depuis longtemps de qualifier un bois non pas par les valeurs absolues de ses caractéristiques N, F, C, ... mais par leur rapport à la densité ou au carré de celle-ci : ces rapports N/D , N/D^2 , F/D , F/D^2 ... etc. sont appelés « cotes ». Si elles conservent un intérêt pratique pour ramener à poids égal des pièces dont on veut comparer la résistance, on voit par contre qu'elles ne doivent pas être interprétées comme étant de très bons modèles mathématiques des liaisons Résistances-Densité.

La figure 20 suggère les observations suivantes :

- les « catégories » de densité se trouvent en bonne correspondance avec la distribution globale, en particulier le point moyen de la population est bien voisin du centre de la catégorie médiane (bois mi-lourds),

- la droite de régression orthogonale de l'ensemble de la population, obtenue en minimisant la somme des carrés des distances des points perpendiculairement à la direction recherchée a pour équation $F = 3012 D - 627$.

Elle ne passe pas par l'origine, ce qui incrimine immédiatement toutes les tentatives de classification par la cote de flexion $F/100 D$.

- les droites limitant les catégories utilisées actuellement, basées sur les valeurs de la cote $F/100 D$ sont représentées : la figure 20 matérialise leur décalage dans la distribution réelle. En effet, le centre de gravité de la population et avec lui la majorité des points se trouvent dans la *catégorie supérieure*, alors que la *catégorie inférieure* n'accueille que très peu de points.

L'utilisation de cette classification amène à surestimer systématiquement les bois puisqu'elle dénomme très généralement « supérieurs » des bois qui n'ont qu'une valeur *moyenne* par rapport à l'ensemble.

Ce type de graphique donne également des indications immédiates pour la sélection de bois présentant des propriétés particulières ou plus exactement une association particulière de deux propriétés (exemple : bois à forte résistance au choc pour la plus faible densité possible). Ces associations précises sont évidemment plus faciles à détecter dans le plan des deux caractères en cause que dans le premier plan principal où elles sont masquées et compensées par le jeu des autres variables.

Une utilisation prometteuse de graphique de ce genre avait été proposée par A. AUBREVILLE (1948) qui prévoyait toute la puissance de cet outil d'interprétation, puissance qui ne semble pas avoir été pleinement mise à profit depuis cette première proposition.

— dans les plans formés par deux variables initiales peu corrélées

(exemple fig. 21 : D et B pour lesquelles $r = 0,59$), on a une *ébauche* intéressante de la représentation complète en composantes principales. En effet D peut être considérée comme la variable « chef de file » de celles qui entrent dans la première composante principale et B comme la variable « chef de file » de celles qui entrent dans la seconde.

Cette ébauche permet évidemment les mêmes utilisations directes que les graphiques précédents (exemple : sélection de bois à forte densité et faible rétractibilité), mais de plus elle donne aux polygones d'espèces des formes beaucoup moins étirées que celles qu'ils prennent par exemple en coordonnées D - F. Ces formes permettent déjà des raisonnements sur la proximité des espèces, et leur groupement par emplois, mais la puissance de description obtenue avec deux variables reste forcément très inférieure à celle qui est permise dans le premier plan principal.

Le progrès réalisé en passant de l'analyse des caractères un par un ou deux par deux à l'analyse en composantes principales est montré par la comparaison des figures 21 et 22.

La question était de savoir si toutes les espèces botaniques du genre *Ocotea* peuvent être légitimement amalgamées pour former une seule essence commerciale (c'est ce qui est fait sur le marché local et cette essence porte le nom unique de Varongy).

— L'analyse monovariante faite en considérant séparément les histogrammes montre que sur les 9 caractères mesurés, tous sauf deux (C et F) donnent une imbrication ou un contact de deux sous-populations : d'un côté l'espèce *Ocotea macrocarpa*, de l'autre côté l'ensemble des autres espèces d'*Ocotea* (*auriculata*, *cymosa*,... etc.). Cette constatation ne permet encore aucune séparation sûre.

— L'analyse bidimensionnelle sur tous les couples de caractères pris deux à deux nécessite, avec 9 caractères, de construire 36 graphiques dont la figure 21 représente l'un d'eux, celui du couple B - D. Ce graphique n'interdit toujours pas d'admettre le mélange de tous les *Ocotea* en une essence commerciale unique.

— L'analyse en composantes principales (fig. 22) donne, quant à elle, une réponse évidente et globale à la question initiale : il est indispensable de séparer les *Ocotea macrocarpa* des autres *Ocotea* et une fois la séparation faite, l'ensemble des autres *Ocotea* peut être considéré comme une essence commerciale sous le nom pilote de Varongy.

Cette conclusion est apportée par la lecture d'un document synthétique (le plan principal) alors que les méthodes analytiques (mono- ou bivariantes) sont d'un emploi beaucoup plus laborieux et moins sûr : en effet, dans le cas présent, l'examen de la densité, souvent considérée comme très instructif à lui seul, n'est d'aucun secours.

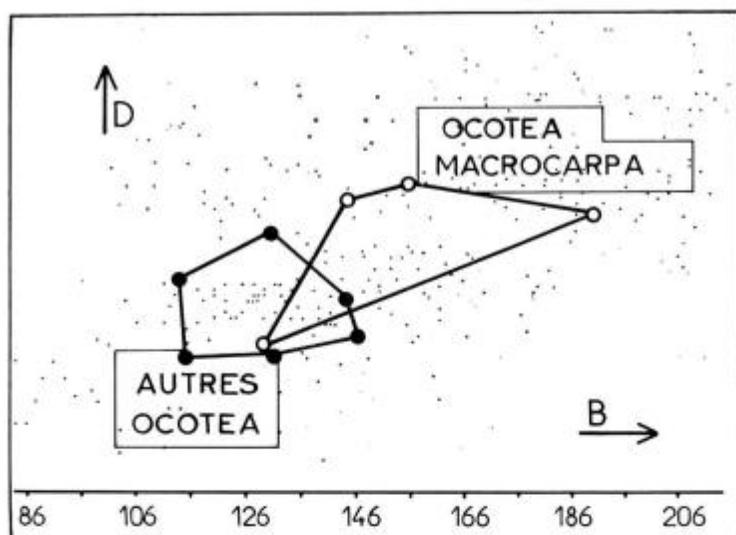


FIG. 21. — *Comparaison Ocotea macrocarpa - Autres Ocotea dans le plan B - D*

FIG. 21. — *Comparaison Ocotea macrocarpa - Other Ocotea in B - D plane*

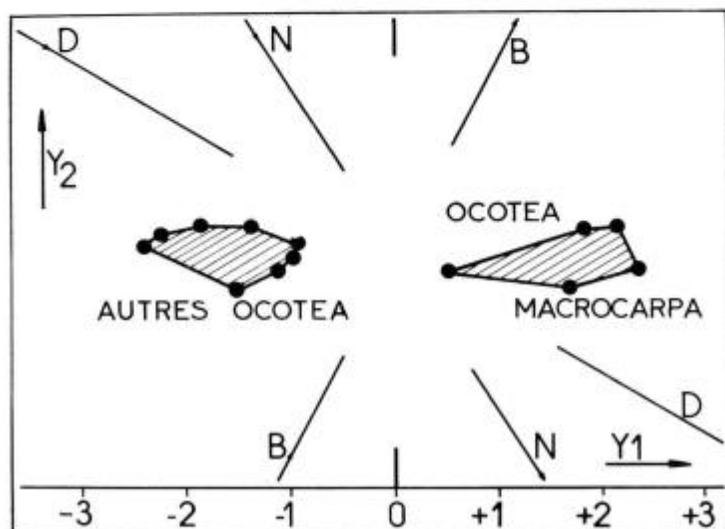


FIG. 22. — *Comparaison Ocotea macrocarpa - Autres Ocotea dans le plan Y₁ Y₂*

FIG. 22. — *Comparaison Ocotea macrocarpa - Other Ocotea in Y₁ Y₂ plane*

VI. — CONCLUSION

Rappelons les deux objectifs initiaux :

- trouver dans l'arsenal des méthodes d'analyse de données un outil approprié à la description technologique des bois,
- utiliser cet outil pour établir un catalogue des vocations industrielles des bois malgaches.

La méthode d'analyse en composantes principales s'est révélée un outil bien adapté au problème posé et nous préconisons la généralisation de son emploi à d'autres fichiers de résultats d'essais sur les bois.

L'expérience d'utilisation de cet outil sur l'exemple des bois malgaches a permis l'établissement d'un catalogue de leurs vocations industrielles.

Pour que ce catalogue soit le plus complet et le plus exact possible, on a tenu compte dans sa rédaction d'informations extérieures à celles que fournit l'analyse, dans le souci de concentrer en un seul document tous les renseignements pouvant être utiles aux utilisateurs. Ces informations extérieures sont de deux sortes :

- notions qualitatives : couleur du bois, facteurs économiques, etc.
- notions quantitatives : résultats d'essais d'usinage, de durabilité naturelle, d'imprégnabilité, etc.

On peut envisager l'utilisation ultérieure de la même méthode pour dépouiller et exploiter les essais sur les bois, caractériser des espèces nouvelles étudiées et prévoir leurs emplois : l'introduction des renseignements qualitatifs nécessitera toujours l'intervention et la réflexion du chercheur responsable de l'étude, mais il serait très intéressant d'incorporer aux calculs la seconde catégorie d'informations supplémentaires.

En effet, l'entrée de ces nouvelles variables élargirait grandement le champ des investigations possibles, en donnant au sens propre de l'expression, des « dimensions nouvelles » au profil dressé pour chaque bois : dès la première phase des calculs (corrélation et formation des premières composantes principales) on observerait facilement l'effet de l'entrée de ces nouvelles variables dont l'importance pratique est incontestablement aussi déterminante dans l'industrie que celle de certaines des variables utilisées ici.

En résumé, la méthode permet, outre une saisie globale des données :

- de diriger d'une façon très éclairée l'exécution d'essais industriels en vraie grandeur,
- de compléter l'échantillonnage des espèces qu'elle révèle prometteuses,
- d'orienter l'élaboration de protocoles expérimentaux futurs, destinés à l'acquisition de données plus riches que celles dont nous disposions.

Reçu pour publication en février 1972.

REMERCIEMENTS

Nous tenons à remercier :

- M. MILLIER C. du C.N.R.F. qui nous a aidés à choisir la méthode, grâce à quelques études préliminaires sur un fichier partiel.
- MM. PAGES J.-P. et BOUVET C. du C.E.A. qui ont supervisé l'exécution du calcul sur l'I.B.M. 360-91 de Saclay.

SUMMARY

PRINCIPAL COMPONENT ANALYSIS OF
TECHNOLOGICAL PROPERTIES OF MALGASIAN TIMBERS

Because of insularity, malgasian timbers are practically unknown by modern industries, and only used on domestic market.

No other results are available than those coming from physical and mechanical tests on small clear specimens, in laboratory.

The purpose of the present work is double :

— to find the most suitable method of analysis able to give a technological description of numerous different timbers, taking altogether in account their properties,

— to use the method to deliver a review of probable industrial abilities of malgasian timbers.

A principal component analysis has been performed on 9 physical and mechanical properties, measured on specimens coming from :

— 414 malgasian trees, belonging to 73 genera and 125 botanical species, sampled in 22 ecological areas spread throughout the Malgasian Island,

— 151 african trees, used as references, belonging to 16 genera and 35 botanical species, sampled in 7 countries of equatorial and western Africa.

The first two principal components (linear combinations of initial properties) explain 83,9 % of the total variation of this population.

They have a practical meaning : the first one can be called *solidity*, or *strength*, the second one can be called *sensibility*. These two new synthetic properties are not dependant on the sampling, as seen when joining african trees to malgasian trees; they can then describe » the tropical timber ».

Using the first principal plan, lots of comparisons are studied between woods from many species, provenances and the natural clustering of different woods in this plan suggests similarity of industrial uses; some well known tropical woods as Okoume give references for these types of comparisons.

The method has allowed in fact to give in an other report the prediction of possible uses of various malgasian timbers, adding to the results of the present analysis other no-quantified informations such as grain, colour, economic disponibility,... etc.

The method is proposed to be applied to similar problems in matter of selecting secondary species of tropical forest, including in further analysis additional variables such as anatomic factors, natural durability, impregnability for example.

ZUSAMMENFASSUNG

PRINZIPALKOMPONENTENANALYSE DER TECHNOLOGISCHEN
EIGENSCHAFTEN MADEGASSISCHER HÖLZER

Wegen der Insellage sind die madegassischen Hölzer der modernen Industrie unbekannt und werden nur auf dem inneren Markt verkauft.

Die einzigen sie betreffenden Wertzahlen sind die Ergebnisse der im Laboratorium nach den französischen Normen ausgeführten physikalischen und mechanischen Versuche.

Der Zweck der vorliegenden Arbeit ist doppelt :

— eine geeignete Untersuchungsmethode auszufinden um eine synthetische technologische Beschreibung einer grossen Anzahl Hölzer zu geben, die allen gemessenen Eigenschaften Rechnung trägt,

— diese Methode anzuwenden, um ein Verzeichnis der mutmasslichen industriellen Verwendungsmöglichkeiten aufzustellen.

Es wurde eine Prinzipalkomponentenanalyse ausgeführt mit 9 physikalischen und mechanischen Eigenschaften von Holzproben aus :

— 414 madegassischen, von 73 Gattungen und 125 botanischen Arten stammenden Bäumen aus 22 verschiedenen in Madagascar verteilten Standorten,

— 151 afrikanischen von 16 Gattungen und 35 botanischen Arten stammenden, Referenzbäumen aus 7 Ländern Aquatorial- und Westafrikas.

Die zwei ersten Prinzipalkomponenten erklären 83,9 % der gesamten Variationen dieser Bevölkerung.

Sie haben eine praktische Auslegung : die erste kann man als Festigkeit, die zweite als Empfindlichkeit bezeichnen. Diese zwei neuen synthetische Eigenschaften hängen nicht von der Probenauswahl ab, wie es daraus hervorgeht wenn man die afrikanischen Bäume mit den madegassischen zusammenzählt. Diese zwei Eigenschaften beschreiben also *alle* Tropenhölzer sehr gut.

Im ersten Prinzipalplan konnten zahlreiche Vergleiche zwischen Hölzern von verschiedenen Arten und Herkunft gemacht werden und die Gruppierung verschiedener Hölzer in gewissen Gebieten dieses Planes führen auf ähnliche industrielle Verwändbarkeit. Gewisse gut bekannte Tropenhölzer wie Okoumé geben Andeutungen für diese Vergleiche.

Das Verfahren erlaubte ein Verzeichnis (siehe besondere Auflage) der Brauchbarkeit auf industriellem Gebiet von verschiedenen madegassischen Hölzer aufzustellen in dem, ausser den Ergebnissen der vorliegenden Analyse, andere, nicht quantitative Auskünfte wie Struktur, Farbe, wirtschaftliche Verfügbarkeit gegeben werden.

Es wird vorgeschlagen dieses Verfahren anderen ähnlichen Problemen anzuwenden besonders um Holzarten zweiten Ranges des Tropenwaldes auszuwählen, mit Erwägung zusätzlicher Eigenschaften : anatomische Faktoren, natürliche Dauerhaftigkeit, Imprägnierbarkeit usw.

RÉFÉRENCES BIBLIOGRAPHIQUES

- AUBREVILLE A., 1948. Les essais mécaniques des bois coloniaux africains. Bois et Forêts des Tropiques n° 6 p. 133-144.
- CAILLIEZ F., MAILLES J. P., NAKACHE J. P., PAGES J. P. 1971. Analyse des données multidimensionnelles — Tome II. Centre d'Études Économiques d'Entreprises — Paris, — p. 113-186.
- GUENEAU P., 1971. Bois de Madagascar, possibilités d'emplois. Centre Technique Forestier Tropical — Paris — 75 p., 1 vol.
- SALLENAVE P., 1955. Propriétés physiques et mécaniques des bois tropicaux de l'Union Française. Centre Technique Forestier Tropical — Paris — p. 7-25.

ANNEXE

RAPPEL DE LA THÉORIE DE L'ANALYSE EN COMPOSANTES PRINCIPALES

I. — Notations

n individus ont été observés; sur chacun d'eux, on a mesuré p caractères.

X est le tableau de description (p lignes, n colonnes)

x_i , i^{me} colonne de X , est l'individu « i » : $x_i \in E = \mathbb{R}^p$

x^j , j^{me} ligne de X , est le caractère « j » : $x^j \in F = \mathbb{R}^n$

E et F sont « l'espace des individus » et « l'espace des caractères »; ce sont des espaces vectoriels à p et n dimensions respectivement

E est muni de la base canonique $\{ e_1, \dots, e_p \}$ (*)

F est muni de la base canonique $\{ f_1, \dots, f_n \}$

(*) e_j a toutes ses coordonnées nulles sauf la j^{me} qui vaut 1,
 f_i a toutes ses coordonnées nulles sauf la i^{me} qui vaut 1.

Chaque individu est muni d'un poids p_i (nombre positif) avec

$$\sum_{i=1}^n p_i = 1. \text{ Ce poids mesure « l'importance » donnée à l'individu « } i \text{ ».}$$

Quand tous les individus ont la même « importance », $p_i = \frac{1}{n}$ quel que soit $i = 1, \dots, n$.

C'est le cas dans la présente étude.

On suppose les caractères *centrés*, c'est-à-dire que l'origine de E est placée au centre de gravité du nuage des individus :

$$\sum_{i=1}^n p_i \underline{x}_i = \underline{0}$$

E est muni de la métrique M : la distance de deux individus \underline{x}_i et \underline{x}_j est

$$d(\underline{x}_i, \underline{x}_j) = \|\underline{x}_i - \underline{x}_j\|_M \quad (*)$$

matriciellement : $d(\underline{x}_i, \underline{x}_j) = \sqrt{(\underline{x}_i - \underline{x}_j)' M (\underline{x}_i - \underline{x}_j)}$

où $(\underline{x}_i - \underline{x}_j)'$ désigne le vecteur (ligne) transposé du vecteur $\underline{x}_i - \underline{x}_j$

F est muni de la « métrique des poids » D_p , de matrice associée

$$D_p = \begin{pmatrix} p_1 & & & 0 \\ & \ddots & & \\ 0 & & & p_n \end{pmatrix}$$

la distance de deux caractères \underline{x}^i et \underline{x}^j s'écrit matriciellement :

$$d(\underline{x}^i, \underline{x}^j) = \|\underline{x}^i - \underline{x}^j\|_{D_p} = \sqrt{(\underline{x}^i - \underline{x}^j)' D_p (\underline{x}^i - \underline{x}^j)}$$

2. — Principe de l'analyse

L'œil ne sait pas regarder un nuage dans l'espace à p dimensions. L'analyse en composantes principales qui se propose de visualiser les données, consiste à chercher une représentation approximative, mais simple et fidèle, du nuage de points.

— « simple » signifie « dans un sous-espace vectoriel de E, de dimension k faible », l'idéal étant $k = 1$ ou 2 puisque l'œil éprouve déjà des difficultés à regarder dans \mathbb{R}^3 .

— « fidèle » signifie « qui ressemble le plus possible » au nuage de départ. Comment décider que deux nuages se « ressemblent »? Divers critères sont possibles : l'analyse en composantes principales se base sur celui d'inertie :

(*) $\|\underline{x}_i - \underline{x}_j\|_M$ se lit "norme du vecteur $\underline{x}_i - \underline{x}_j$ au sens de la métrique M".

\underline{c}_i^k désignant la projection M-orthogonale (c'est-à-dire orthogonale au sens de \underline{L} la métrique M) du point \underline{x}_i sur un sous-espace vectoriel W_k de dimension $k < p$,

$$I_{W_k} = \sum_{i=1}^n p_i \left\| \underline{x}_i - \underline{c}_i^k \right\|_M^2$$

est l'inertie du nuage des \underline{x}_i par rapport à W_k .

De tous les espaces vectoriels de E à k dimensions, on retient W_k si I_{W_k} est minimum.

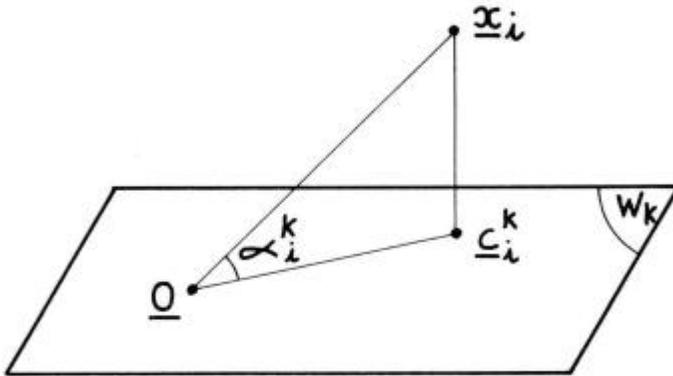


FIG. 23. — Projection d'un individu dans un sous-espace vectoriel

FIG. 23. — Projection of an element in a linear subspace

2.1. — La tactique.

On cherche W_1 , puis W_2 si la représentation dans W_1 n'est pas assez fidèle, puis W_3 si la représentation dans W_2 n'est pas encore assez fidèle, et ainsi de suite jusque $W_p = E$.

2.2. — Les résultats.

On montre :

— que $g = 0 \subset W_1 \subset W_2 \dots \subset W_p$, c'est-à-dire que les sous-espaces vectoriels de dimension croissante à inertie minimum sont inclus les uns dans les autres.

— que $W_1 = \Delta_{u_1}$, droite engendrée par le vecteur propre normé \underline{u}_1 de $V M$ associé à la plus grande valeur propre λ_1 .

$$V M \underline{u}_1 = \lambda_1 \underline{u}_1 \quad \text{avec} \quad V = X D_p X'$$

(du fait que les caractères sont centrés, V est la matrice des variances-covariances des p caractères \underline{x}^j)

Δ_{u_1} est le premier axe principal.

— que W_2 est le plan engendré par Δ_{u_1} et la droite Δ_{u_2} (deuxième axe principal) définie par le vecteur propre \underline{u}_2 de $V M$ associé à la deuxième plus grande valeur propre λ_2 et ainsi de suite...

Le processus, une fois mené jusqu'à son terme, revient donc à faire un changement de base dans E , la nouvelle base étant formée des vecteurs propres normés \underline{u}_1 de $V M$ associés aux valeurs propres λ_1 rangées par ordre décroissant.

Deux vecteurs propres de $V M$ associés à des valeurs propres différentes étant M -orthogonaux, la base $\{\underline{u}_1, \dots, \underline{u}_p\}$ de E est M -orthonormée.

Les composantes principales.

A chaque axe principal sont associées les projections M -orthogonales des points-individus sur lui.

c_i^j , coordonnée de la projection de l'individu i sur le j^{me} axe principal, est la valeur que prend la j^{me} composante principale chez l'individu i .

Matriciellement :

$$c_i^j = x_i' M \underline{u}_j \implies \underline{c}^j = \begin{pmatrix} c_1^j \\ \vdots \\ c_n^j \end{pmatrix} = X' M \underline{u}_j \in F$$

\underline{c}^j est la j^{me} composante principale.

Comme tout point de F , \underline{c}^j peut être regardée comme un caractère.

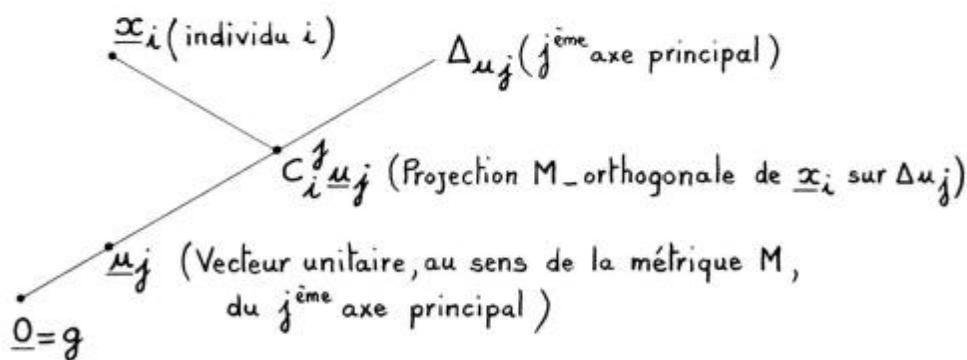


FIG. 24. — Projection d'un individu sur un axe principal

FIG. 24. — Projection of an element on a principal axis

3. — Description des individus

On observe la projection du nuage des individus dans le plan principal (plan formé par les deux premiers axes principaux) et, si cette projection est peu fidèle, dans les plans formés par les axes principaux suivants.

On « oriente » le plan principal en y faisant figurer les projections M-orthogonales des vecteurs unitaires $\underline{e}_1, \dots, \underline{e}_p$ des axes associés aux caractères initiaux, (cf. fig. 5).

3.1. — *Un critère pour juger de la « fidélité » de la représentation du nuage par sa projection M-orthogonale dans W_k .*

Appliquons le théorème de Pythagore à chaque triangle $\{ \underline{0}, \underline{x}_i, \underline{c}_i^k \}$ où \underline{c}_i^k désigne la projection M-orthogonale de \underline{x}_i sur W_k

$$\| \underline{x}_i \|^2 = \| \underline{c}_i^k \|^2 + \| \underline{x}_i - \underline{c}_i^k \|^2$$

Après multiplication par le poids p_i et sommation sur les n individus, on obtient (I_0 désigne l'inertie du nuage des individus par rapport à l'origine) :

$$I_0 = \sum_{i=1}^n p_i \| \underline{x}_i \|^2_M = \sum_{i=1}^n p_i \| \underline{c}_i^k \|^2_M + I_{W_k}$$

La représentation du nuage des n points \underline{x}_i par le nuage des n points \underline{c}_i^k est d'autant plus fidèle que :

$$\frac{I_{W_k}}{I_0} \text{ est faible} \iff \frac{\sum_{i=1}^n p_i \| \underline{c}_i^k \|^2_M}{I_0} \text{ est fort.}$$

C'est le deuxième rapport, appelé *part d'inertie expliquée par le sous-espace principal* W_k , qui est couramment utilisé pour mesurer la qualité de la représentation du nuage par sa projection M-orthogonale dans W_k .

Cette représentation est d'autant meilleure que ce rapport est proche de 1.

$$0 \leq \frac{\sum_{i=1}^n p_i \| \underline{c}_i^k \|^2_M}{I_0} = \frac{\lambda_1 + \dots + \lambda_k}{\lambda_1 + \dots + \lambda_p} \leq 1$$

3.2. — *Comment mesurer la qualité de la représentation d'un individu dans W_k .*

Le pourcentage d'inertie expliqué par W_k ne donne qu'une idée *globale* de la qualité de la représentation du nuage par sa projection dans W_k ; des nuages très différents peuvent avoir la même part d'inertie expliquée par leur sous-espace principal à k dimensions; il est donc nécessaire, quand on émet des jugements sur la proximité de deux individus \underline{x}_i et \underline{x}_j au vu des proximités de leurs projections \underline{c}_i^k et \underline{c}_j^k dans W_k , de s'assurer que la qualité de la représentation de \underline{x}_i par \underline{c}_i^k et celle de \underline{x}_j par \underline{c}_j^k sont semblables.

On peut prendre comme *mesure de la qualité de la représentation* de \underline{x}_i par \underline{c}_i^k , le nombre :

$$\frac{\|\underline{c}_i^k\|}{\|\underline{x}_i\|} \frac{M}{M} = \cos \alpha_i^k$$

L'individu « i » est d'autant mieux représenté dans W_k que son indice est proche de 1.

Le programme que nous avons employé utilise cet indice; il fournit en fait le carré du cosinus qui mesure la qualité de la représentation de chaque individu sur chaque axe principal, ce qui permet de calculer facilement la qualité de la représentation sur tout plan défini par deux axes principaux; en effet si $\alpha_i^1, \alpha_i^2, \alpha_i^{1,2}$ désignent respectivement les angles formés par \underline{x}_i et ses projections M-orthogonales sur le premier axe principal, le second axe principal et le premier plan principal,

$$\cos^2 \alpha_i^{1,2} = \cos^2 \alpha_i^1 + \cos^2 \alpha_i^2$$

L'indice $\cos \alpha_i^k$ n'est intéressant que si \underline{x}_i est éloigné de W_k^\perp , sous-espace vectoriel supplémentaire M-orthogonal de W_k dans E (si \underline{x}_i est dans W_k^\perp , il se projette sur W_k en θ); un meilleur indice pourrait être :

$$\frac{p_i \|\underline{x}_i - \underline{c}_i^k\|}{I_{W_k}} = \frac{p_i \|\underline{x}_i - \underline{c}_i^k\|}{\sum_{i=1}^n p_i \|\underline{x}_i - \underline{c}_i^k\|} \frac{M}{M}$$

qui présente aussi l'intérêt d'être compris entre 0 et 1.

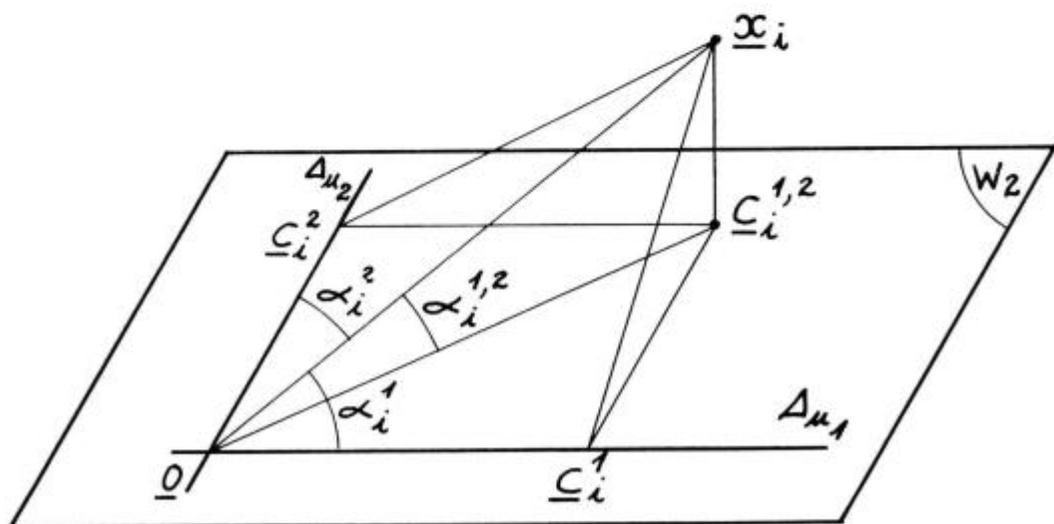


FIG. 25. — Qualité de la représentation sur deux axes principaux et dans le plan qu'ils forment
 FIG. 25. — Quality of representation on two principal axis and on the plane which they span

4. — Description des caractères

Dans F , les p caractères \underline{x}^j déterminent un sous-espace vectoriel à p dimensions dont les \underline{c}^j constituent une base D_p -orthogonale. De plus, la norme (au sens de D_p) de la composante principale \underline{c}^j est $\sqrt{\lambda_j}$.

Montrons en effet que deux composantes principales \underline{c}^j et \underline{c}^k (associées respectivement aux valeurs propres λ_j et λ_k de VM) sont D_p -orthogonales et que leurs normes (au sens de D_p) sont respectivement $\sqrt{\lambda_j}$ et $\sqrt{\lambda_k}$:

$$\left. \begin{array}{l} \underline{c}^j = X'M \underline{u}_j \\ \underline{c}^k = X'M \underline{u}_k \\ VM \underline{u}_j = \lambda_j \underline{u}_j \\ VM \underline{u}_k = \lambda_k \underline{u}_k \\ (\underline{u}_j)' M \underline{u}_j = 1 \\ (\underline{u}_j)' M \underline{u}_k = 0 \end{array} \right\} \implies \begin{array}{l} (\underline{c}^j)' D_p \underline{c}^k = \underline{u}_j' M X D_p X' M \underline{u}_k = \underline{u}_j' M V M \underline{u}_k \\ \quad \quad \quad = \underline{u}_j' M (\lambda_k \underline{u}_k) = 0 \\ (\underline{c}^j)' D_p \underline{c}^j = \underline{u}_j' M V M \underline{u}_j = \underline{u}_j' M (\lambda_j \underline{u}_j) = \lambda_j \end{array}$$

($\|\underline{c}^j\|_{D_p}^2 = \lambda_j$ signifie que la j^{me} composante principale, du fait qu'elle est centrée, a pour variance λ_j).

Tout caractère \underline{x}^j peut donc être repéré par ses coordonnées dans la base des \underline{c}_k

$$\frac{(\underline{x}^j)' D_p \underline{c}^k}{\sqrt{\lambda_k}} = \frac{\text{cov}(\underline{x}^j, \underline{c}^k)}{\sqrt{\lambda_k}} \quad k = 1, \dots, p$$

A la représentation des points-individus dans le plan principal est donc associée la représentation des points-caractères dans le plan $(\underline{c}^1, \underline{c}^2)$ des deux premières composantes principales.

Le caractère \underline{x}^j intervient d'autant plus dans la description du nuage des individus par sa projection dans le plan principal qu'il est « proche » du plan $(\underline{c}^1, \underline{c}^2)$, c'est-à-dire que

$$\|\underline{x}^j\|_{D_p}^2 \text{ est voisin de } \frac{1}{\lambda_1} \left(\text{cov}(\underline{x}^j, \underline{c}^1) \right)^2 + \frac{1}{\lambda_2} \left(\text{cov}(\underline{x}^j, \underline{c}^2) \right)^2$$

L'interprétation des composantes principales \underline{c}^1 et \underline{c}^2 se fait au vu des proximités entre les projections des caractères \underline{x}^j dans le plan $(\underline{c}^1, \underline{c}^2)$ et les composantes principales \underline{c}^1 et \underline{c}^2 .

5. — Cas de l'analyse en composantes principales « sur matrice de corrélation »
(cas de la présente étude)

Pour s'affranchir du fait que les résultats d'une analyse en composantes principales dépendent des unités de mesure des caractères, il est courant de faire l'analyse sur les données réduites (division de chaque caractère par son écart-type); la métrique M a alors pour matrice la matrice diagonale des inverses des variances des caractères; on la note $M = D_1/\sigma^2$.

Alors :

- la matrice dont on cherche les vecteurs propres est la matrice de corrélation
- les caractères \underline{x}^j sont de norme 1 :

$$\| \underline{x}^j \|_{D_p} = 1 \quad j = 1, \dots, p$$

et les coordonnées des caractères réduits \underline{x}^j dans la base des composantes principales sont les corrélations de ces caractères avec les composantes principales :

$$\frac{(\underline{x}^j)' D_p \underline{e}^k}{\sqrt{\lambda_k}} = \text{cor}(\underline{x}^j, \underline{e}^k)$$

Pour voir si un point-caractère est bien représenté par sa projection dans le plan $(\underline{e}^1, \underline{e}^2)$, il suffit donc de voir si cette projection est proche du cercle de centre $\underline{0}$ et de rayon 1.