

# 34

## Relations statistiques pluie pollinique-végétation en forêt tropicale africaine

MAURICE ROUX, GASTON ACHOUNDONG, HILAIRE ELENGA,  
CHRISTIAN DE NAMUR, ISABELLE REYNAUD-FARRERA, ANNIE VINCENS

### Introduction

L'analyse palynologique des sédiments, dont l'objectif est la reconstitution des écosystèmes végétaux passés, implique que l'ensemble des pollens et des spores déposés en un même site représente le plus fidèlement possible la végétation ou, du moins, qu'il existe une relation étroite entre la pluie pollinique et la végétation. La connaissance de cette relation est la base même de toute interprétation des spectres polliniques anciens ; mais il est certain que celle-ci est complexe car elle est sous la dépendance de nombreux facteurs — écologiques, climatiques —, mais est également liée à la production pollinique des végétaux, au mode de dispersion des pollens. En Afrique, afin d'essayer de comprendre une telle relation, de nombreux travaux concernant l'étude du contenu pollinique d'échantillons actuels de sol ont été effectués. Cependant, et à une exception près (Gabon), les résultats obtenus n'ont jamais pu être comparés directement à des données botaniques précises et quantifiées au site d'échantillonnage. De ce fait, les interprétations fournies sur la relation pollen-végétation restent largement intuitives et empiriques.

Dans le cadre du programme ECOFIT, un atelier « calibration pollen-végétation » a été créé, qui a permis d'établir pour la première fois une collaboration étroite entre palynologues, botanistes et mathématiciens. Les travaux menés dans le cadre de cet atelier concernent deux régions forestières africaines, le Cameroun et le Sud-Congo, pour lesquelles, sur les mêmes sites (respectivement 53 et 64), des données botaniques et des données palynologiques quantifiées ont été recueillies. Ainsi, à chaque échantillon pollinique est associé un

relevé phytosociologique du couvert végétal de la parcelle (20 × 20 m) sur laquelle il a été prélevé.

Au Cameroun, mises à part les forêts montagnardes et les observations de lisières forêt-savane, les données de moyennes et basses altitudes se sont révélées très homogènes, tant du point de vue pollinique que du point de vue botanique. De plus, les gradients qui sous-tendent les deux types de données sont différents. Une analyse conjointe des deux tableaux n'a pas fourni de résultats cohérents, mais une approche analytique examinant l'ensemble des corrélations entre taxons botaniques et taxons polliniques a révélé l'existence de plusieurs groupes d'espèces associées. (Reynaud, 1995).

Au Congo, il semble que les types de forêts soient beaucoup plus marqués qu'au Cameroun (dans l'état actuel de l'échantillonnage). Nous en avons profité pour pousser plus avant les calculs statistiques, en allant jusqu'à tenter de prédire le type de végétation à partir des enregistrements polliniques. C'est sur ces travaux qu'est focalisé le présent exposé.

Au paragraphe suivant nous détaillons les méthodes statistiques utilisées. Ensuite nous présentons les résultats obtenus au Congo. Nous terminons par une conclusion résumant les acquis obtenus au cours de cette étude.

Pour l'échantillonnage et les problèmes liés à chaque type de données nous renvoyons à l'article de H. Elenga dans ce même volume.

## Méthodes statistiques

### Analyses factorielles

Les données de végétation sont constituées par des comptages d'arbres des différentes espèces rencontrées, tandis que les pollens se présentent sous forme de pourcentages de grains de chaque taxon identifiable, par rapport au nombre total de grains examinés. L'analyse factorielle de correspondances (AFC) s'applique bien aux données de végétation. En ce qui concerne les données polliniques, outre certaines méthodes spécifiques à ce type de tableau (Aitchison, 1986), on est tenté d'appliquer l'analyse en composantes principales qui pondère chaque taxon par l'écart-type de ses pourcentages. Nous avons toutefois utilisé de nouveau l'AFC pour les raisons suivantes : dans le cas des pourcentages, la double pondération par les poids des lignes et des colonnes se réduit à une simple pondération par les poids des colonnes (taxons) puisque les poids des lignes (relevés) sont tous égaux à cent. Par ailleurs, on sait que, en général, plus les pourcentages sont élevés, plus leur écart-type est grand, de sorte que pondérer par la somme des pourcentages donne des résultats très semblables. Nous avons préféré l'AFC pour le centrage simultané des deux nuages de points, taxons et prélèvements, qui facilite l'interprétation des résultats. L'AFC a également été utilisée pour l'étude directe des relations entre pollens et végétation (voir plus bas Premier essai de détermination de la forêt associée à un spectre pollinique).

### Tests de corrélation

On a utilisé le coefficient de corrélation usuel (Bravais-Pearson) pour tenter de déterminer les pollens les plus révélateurs de la végétation ambiante (taxons marqueurs). On sait qu'il est possible de tester la nullité de ce coefficient, mais le test employé habituellement repose sur l'hypothèse que les deux variables étudiées doivent être distribuées selon une loi de Gauss bidimensionnelle... ce qui est loin d'être notre cas (d'autre part, les tests non-paramétriques de Spearman ou de Kendall sont difficilement applicables à cause des nombreuses valeurs nulles représentant l'absence du taxon dans les parcelles correspondantes). C'est pourquoi nous avons préféré tester ce coefficient par permutations aléatoires (Edgington, 1987 ; Manly, 1991).

Le principe de cette méthode consiste à permuter arbitrairement les valeurs de l'une des variables par rapport à l'autre variable étudiée ; ici un taxon pollinique par rapport à une espèce arborée. De sorte que le pourcentage de pollens attribués à un site n'a plus rien à voir avec l'abondance des arbres correspondants sur ce même site. On simule ainsi une situation de corrélation nulle, ou de non-corrélation, mais en conservant la distribution des valeurs de chaque variable. On compare alors le coefficient réellement observé avec les valeurs de ce coefficient sur quelques centaines (deux cents dans le cas présent) de permutations aléatoires. Si le coefficient observé est du même ordre de grandeur que les valeurs obtenues par simulation on retient l'hypothèse de nullité de la corrélation. En revanche, si moins de 5 % des valeurs simulées sont supérieures ou inférieures à la

vraie valeur, on considère celle-ci comme significative (au risque 0,05) d'une liaison entre les deux variables considérées.

Une variante de cette méthode, le test de Mantel, a également été utilisée pour tester globalement la liaison entre deux tableaux portant sur les mêmes observations ; ici le tableau des données polliniques et le tableau contenant les effectifs des arbres. Ce test consiste à calculer deux matrices de distances entre les observations, l'une à partir des données de végétation, l'autre à partir des données polliniques. En conformité avec le choix de l'AFC, la formule utilisée pour ces deux matrices de distance est la métrique du Chi<sup>2</sup>, qui utilise les mêmes pondérations que l'AFC. On permute alors aléatoirement les observations de l'une des matrices par rapport à l'autre, ce qui nécessite un déplacement simultané des lignes et des colonnes. Puis l'on compare le coefficient de corrélation entre ces deux matrices non permutees, avec les coefficients de corrélations obtenus après permutations. La conclusion du test se réalise comme précédemment.

### Détermination des espèces caractéristiques d'un site : rapport de corrélation

Il est possible de déterminer facilement les espèces caractéristiques d'un, ou de quelques-uns des sites étudiés grâce au « rapport de corrélation », calculé pour chaque espèce. Si l'on appelle

- ▷  $x_{ik}$  l'abondance d'une espèce  $x$  dans le relevé  $i$  du site  $k$ ,
- ▷  $\bar{x}_k$  l'abondance moyenne de cette espèce dans le site  $k$ ,
- ▷  $\bar{x}$  l'abondance moyenne de cette espèce sur l'ensemble des sites,
- ▷  $n_k$  le nombre de relevés du site  $k$ ,

alors on peut écrire l'équation classique d'Analyse de la variance :

$$\sum_k \sum_i (x_{ik} - \bar{x})^2 = \sum_k n_k (x_k - \bar{x})^2 + \sum_k \sum_i (x_{ik} - x_k)^2$$

En utilisant l'abréviation SCE pour « somme des carrés des écarts à la moyenne », cette équation s'écrit encore :

$$SCE_{totale} = SCE_{inter} + \text{Somme des } SCE_{intra}$$

(Les mots inter et intra sont mis pour inter-site et intra-site).

Dans le cas où l'on prend en compte plusieurs espèces à la fois, l'équation reste valable en remplaçant l'expression SCE par le mot « inertie ». Le rapport de corrélation est le rapport :

$$r_x^2 = SCE_{inter} / SCE_{totale}$$

Par sa définition même, il est toujours compris entre 0 et 1. Il est proche de 1 lorsque l'espèce est abondante dans certains sites et rare ou absente dans d'autres.

Il est voisin de zéro lorsque ses fluctuations d'abondance n'ont rien à voir avec les sites. Ici aussi, il est possible d'appliquer une procédure de permutations pour obtenir un seuil de significativité du rapport de corrélation, mais le programme de calcul n'est pas encore au point. On s'est contenté d'une approche simplement exploratoire dans laquelle on calcule la contribution de chaque site au rapport de corrélation, ce qui permet de trouver rapidement quelles sont les espèces typiques de chaque site.

## Analyse factorielle discriminante

En vue de déduire le type de végétation à partir de la composition pollinique d'un échantillon de sol, nous avons utilisé l'analyse factorielle discriminante. Les données à fournir pour cette méthode sont, d'une part les compositions polliniques d'un ensemble d'observations, et d'autre part le type de végétation ou groupe associé à chaque observation. Le principe de la méthode se décompose en quatre étapes : on calcule d'abord la composition pollinique moyenne correspondant à chaque type de végétation ; on effectue ensuite une analyse factorielle particulière sur ces compositions moyennes ou centres de gravité des groupes ; on projette alors les points moyens et l'ensemble des observations sur les axes factoriels obtenus ; la dernière étape consiste à calculer les distances entre chaque observation et l'ensemble des centres de gravité, et d'adopter pour règle de déduction que le type de végétation associé à une composition pollinique donnée est vraisemblablement celui dont le centre de gravité est le plus proche de cette composition.

L'analyse factorielle utilisée est particulière car, non seulement elle donne une dispersion maximum au nuage des centres de gravité, mais elle fait en sorte que les nuages de points associés aux différents groupes d'observations se chevauchent aussi peu que possible, d'où le nom d'analyse discriminante.

Il faut remarquer que la règle de déduction du type de végétation s'applique aussi bien aux observations actuelles ayant servi à établir les axes factoriels, qu'à des observations fossiles pour lesquelles le type de végétation est inconnu. L'application de la règle aux données actuelles révèle cependant un certain nombre d'erreurs de détermination ; le pourcentage d'erreurs par rapport au nombre d'observations actuelles est un indice de la qualité de la discrimination. On préfère toutefois exprimer cette qualité par « le pourcentage de bien classées » qui varie en sens contraire du pourcentage d'erreurs.

Ce pourcentage de bien classées (PBC) doit être utilisé avec circonspection. En effet, si l'on modifie l'échantillon des sols actuels il est probable que les axes factoriels discriminants vont être modifiés ainsi que la position des centres de gravité sur ces axes. La règle de détermination des types de végétation sera donc changée... sans grand changement, en général, dans le PBC. Cela suggère que le pourcentage d'erreurs sur les données fossiles doit être plus élevé que la valeur indiquée par le PBC.

Pour se rapprocher des conditions réelles d'application de la règle de détermination on recommande de n'utiliser celle-ci que sur des observations actuelles n'ayant pas servi à déterminer les axes factoriels. Ceci implique de mettre de côté un certain nombre d'observations, qu'on peut appeler « observations d'épreuve », avant d'appliquer l'analyse factorielle discriminante sur les autres observations, appelées « observations de base ». Lorsque l'échantillon des données actuelles est suffisamment grand on peut en prendre une partie (le quart ou le cinquième) comme « observations d'épreuve » et calculer le PBC sur ces observations qui doivent être tirées au hasard (Romeder, 1973). Dans le cas contraire on peut utiliser la méthode de « l'observation écartée » (méthode

« *leave-one-out* », ter Braak). Cette procédure consiste à enlever une observation et à utiliser, la règle de détermination obtenue à l'aide des autres observations, qui constituent temporairement, les « observations de base ». On réitère donc les calculs d'analyse discriminante pour chacune des observations, mise tour à tour de côté, puis réaffectée à un type de végétation. C'est sur ces réaffectations qu'est alors calculé le PBC, qui donne une estimation réaliste du taux de succès de l'analyse

## Résultats obtenus au Congo

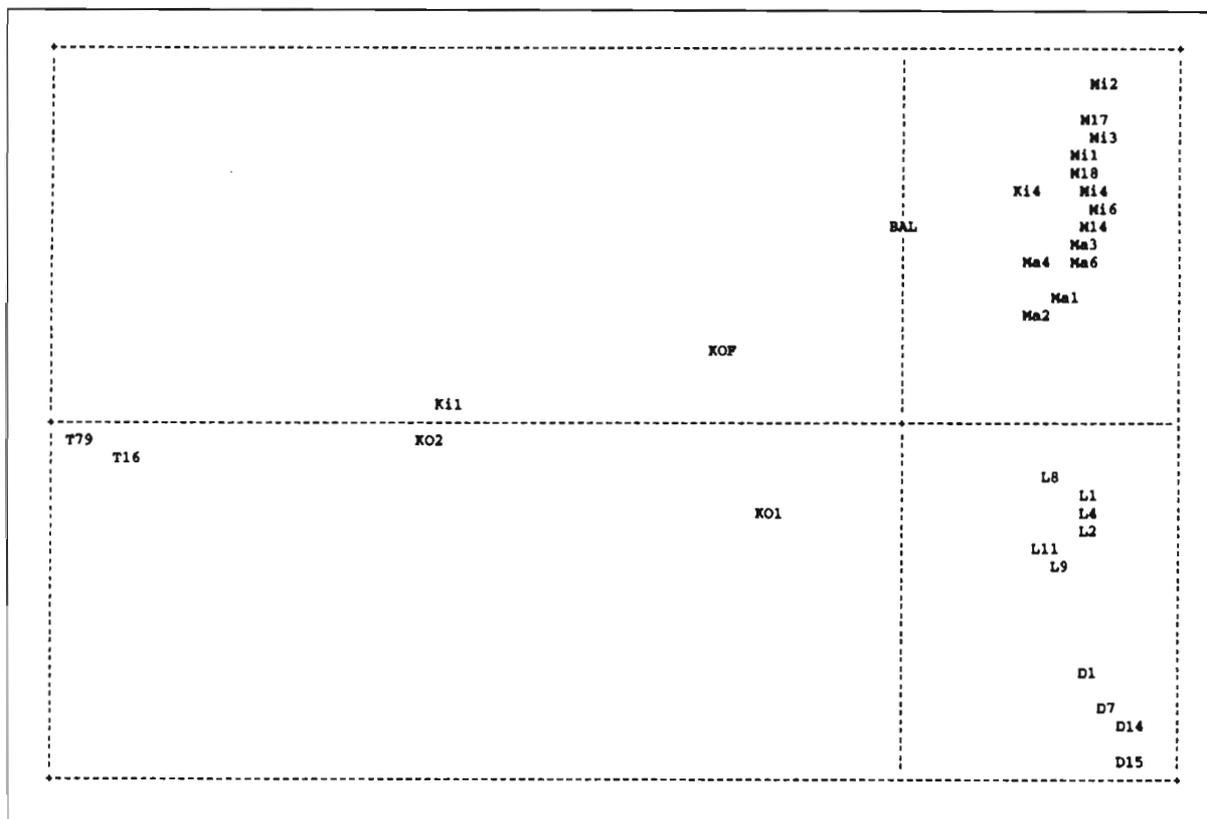
### Relation globales entre pollens et végétation

Dans ce paragraphe nous examinons les analyses factorielles des correspondances des données botaniques et des données polliniques séparées. Puis nous examinons les liaisons entre ces deux types de données à l'aide du test global de Mantel, et aussi par l'étude des corrélations entre les axes factoriels d'une analyse et les axes factoriels de l'autre analyse.

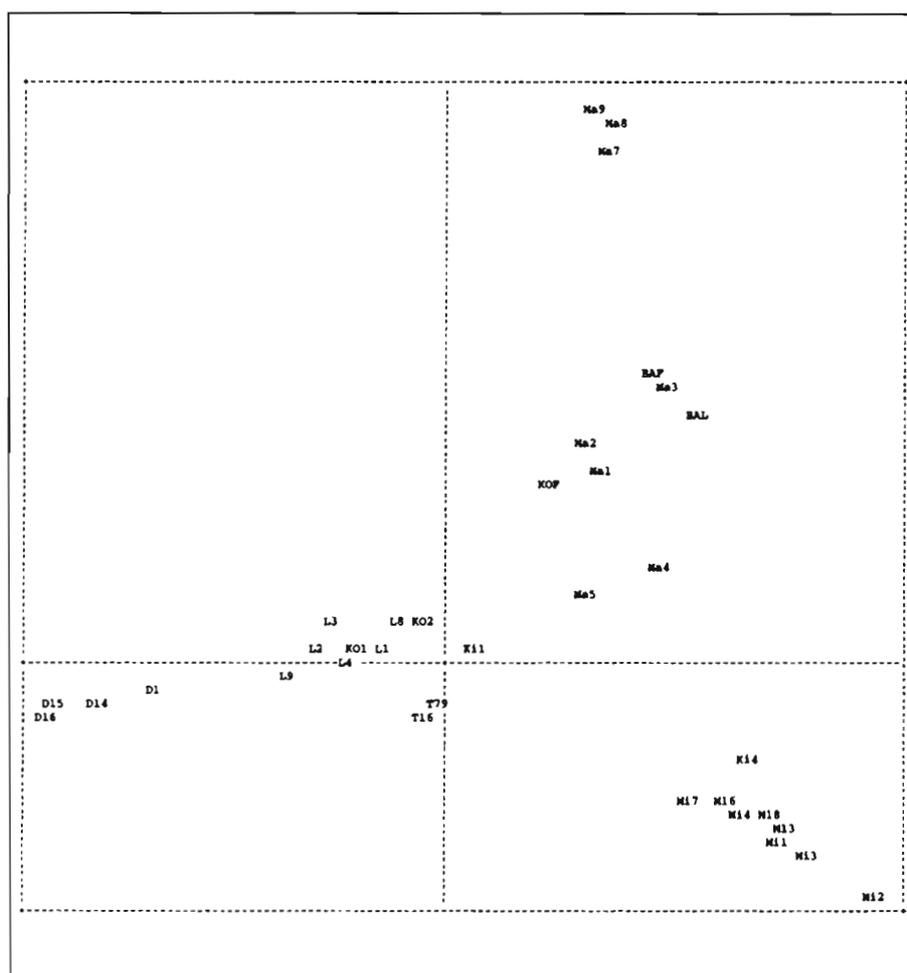
Au cours de l'étude le nombre de parcelles traitées a varié selon la progression du dépouillement des données de terrain. De plus on a essayé des AFC avec toutes les espèces ou bien seulement avec les espèces les plus fréquentes, etc. Dans tous les cas, les relevés botaniques du Cameroun, comme les relevés polliniques, se sont rassemblés de façon très nette selon leur station de provenance, comme le montrent les Figures 1 et 2, pour la botanique, et 3 pour les pollens.

La principale caractéristique de la Figure 1 est de montrer l'originalité des forêts de Tchissanga, Koubotchi et Kitina, situées dans la partie négative de l'axe 1 et s'opposant à toutes les autres (sauf pour le relevé Ki 4 qui se rapproche des relevés de Mindu). Il s'agit en effet d'une végétation littorale très particulière liée aux terrains marécageux. Nous désignerons ces forêts par le sigle FLM (forêts littorales marécageuses). Sur le plan de ses axes 2 et 3 (Figure 2) ce même nuage a la forme d'une étoile à 3 branches. Si l'on fait abstraction du groupe FLM, isolé par l'axe 1, chaque pointe de l'étoile est occupée par un groupe de relevés provenant d'une seule station : Mindu (Axe 2  $\geq 0$ , Axe 3  $\leq 0$ ), Di Monika (Axe 2  $\leq 0$ , Axe 3  $\geq 0$ ), Mandzi (Axe 2  $\geq 0$ , Axe 3  $\geq 0$ ). Au centre de l'étoile se trouvent les relevés de La Tour. Cette disposition admet cependant une exception : les deux relevés provenant de la station des Bandas qui se mêlent à ceux de Mandzi. De plus les relevés 7, 8 et 9 de Mandzi, correspondant à un deuxième site au sein de la station de Mandzi, se séparent nettement de ceux du premier site (relevés 1 à 6) à l'extrémité positive de l'axe 2.

Les relevés polliniques (Figure 3) présentent également une disposition en étoile à trois branches, mais ce ne sont plus les mêmes stations qui occupent les extrémités et le centre de l'étoile : Mindu est toujours à une extrémité (Axe 1  $\geq 0$ , Axe 2  $\leq 0$ ) mais La Tour en occupe une autre (Axe 1  $\leq 0$ , Axe 2  $\leq 0$ ), tandis que la troisième pointe (Axe 2  $\geq 0$ ) est formée par le groupe FLM. Les stations de Di Monika et Mandzi se confondent vers le centre du graphique.



**Figure 1** Analyse factorielle des correspondances des relevés botaniques. Plan des axes 1 (horizontalement) et 2 (verticalement) : D = Di Monika ; L = La Tour ; Ma = Mandzi ; Mi = Mindu ; BAF = les Bandas-Forêt ; BAL = les Bandas-Lisière ; T = Tchissanga ; Ki = Kitina ; KO = Koubotchi ; KOF = Koubotchi (Friche). Les relevés provenant des forêts littorales marécageuses se détachent du côté négatif de l'axe 1.



**Figure 2** Analyse factorielle des correspondances des relevés botaniques. Plan des axes 2 (horizontalement) et 3 (verticalement) : D = Di Monika ; L = La Tour ; Ma = Mandzi ; Mi = Mindu ; BAF = les Bandas-Forêt ; BAL = les Bandas-Lisière ; T = Tchissanga ; Ki = Kitina ; KO = Koubotchi ; KOF = Koubotchi (Friche). Les relevés provenant d'une même station sont bien groupés.

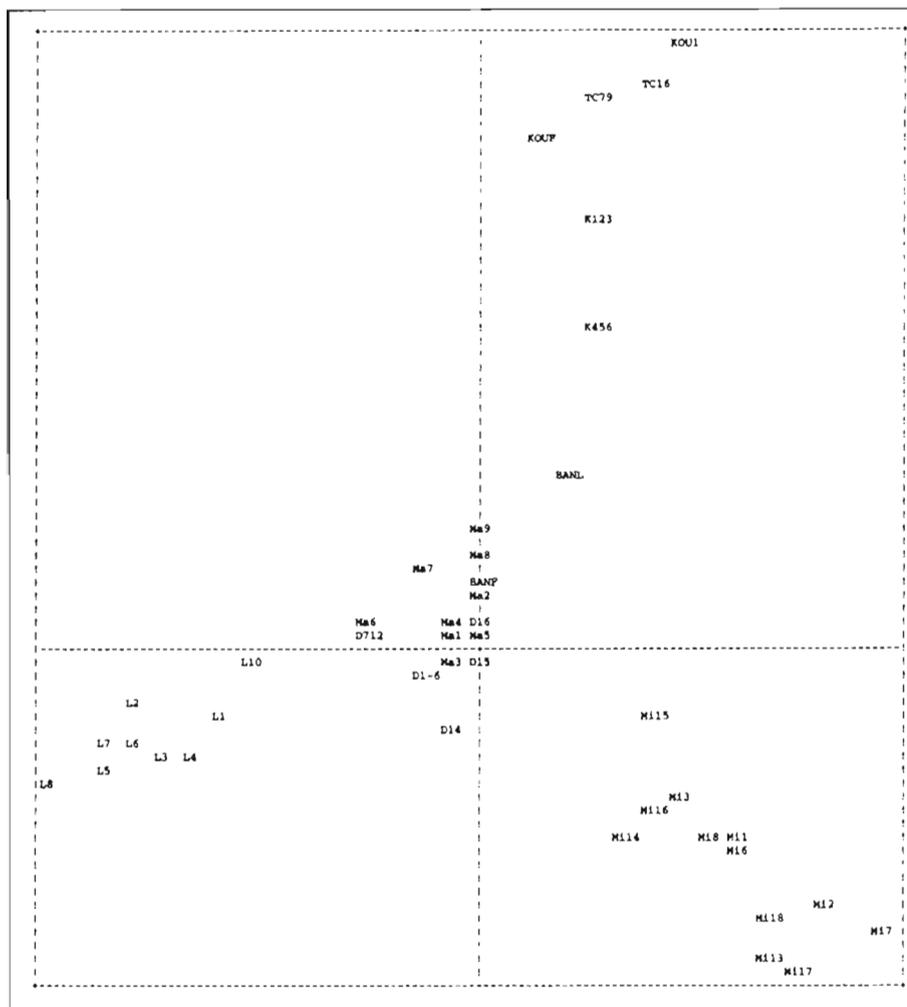


Figure 3 Analyse factorielle des correspondances des relevés polliniques, en pourcentages. Plan des axes 1 (horizontalement) et 2 (verticalement) : D = Di Monika ; L = La Tour ; Ma = Mandzi ; Mi = Mindu ; BANF = les Bandas-Forêt ; BANL = les Bandas-Lisière ; TC = Tchissanga ; K = Kirima ; KOU = Koubotchi ; KOUF = Koubotchi (Friche). Les relevés provenant d'une même station sont bien groupés.

Ainsi chaque station est bien individualisée, tant du point de vue pollinique que du point de vue des arbres ; malheureusement les axes factoriels qui les sous-tendent ne peuvent pas être identifiés les uns aux autres.

Seule la corrélation entre l'axe 2 des pollens et l'axe 1 de la botanique est assez élevée (0,8 Tableau 1), mais cela n'est guère satisfaisant car cette corrélation est imputable à la position excentrée des relevés FLM.

On a, par ailleurs, calculé les distances du Chi<sup>2</sup> entre les relevés botaniques d'une part, et entre les relevés polliniques d'autre part. Et l'on a appliqué le test de Mantel sur ces deux tableaux de distances : la corrélation entre les deux tableaux est  $r = 0,1757$  avec une probabilité  $p = 0,09$  d'être dépassée, ce qui n'est pas significatif (au risque 5 %). Curieusement, la même procédure basée sur les seules présences ou absences des espèces botaniques ou des taxons polliniques a donné un coefficient  $r = 0,4174$  avec une probabilité  $p = 0,002$  révélatrice d'une corrélation très significative (risque  $\leq 1\%$ ).

### Recherche des taxons-marqueurs

À l'inverse de l'approche précédente nous avons essayé une démarche analytique, étudiant les données taxon par taxon, en cherchant à déterminer les taxons caractéristiques de chaque type de forêt ; ces taxons caractéristiques sont appelés taxons-marqueurs. Dans cette démarche nous avons privilégié les taxons polliniques, ceux-ci devant être à la base des reconstitutions paléo-environnementales.

L'analyse des correspondances du tableau pollinique permet de faire une première sélection des taxons les plus typiques, en retenant ceux qui ont de fortes contributions sur les premiers axes factoriels (Tableau 2). On dégage ainsi une liste de vingt-deux taxons susceptibles de révéler l'appartenance des relevés à telle ou telle station.

Une autre approche du problème a été de calculer les corrélations entre taxons polliniques et taxons botaniques, en ne retenant que les paires de taxons hétérogènes, c'est à dire un taxon pollinique différent du taxon botanique.

	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	P <sub>5</sub>	P <sub>6</sub>
B <sub>1</sub>	-0,1709	-0,8034	0,2996	0,0033	0,1082	-0,1476
B <sub>2</sub>	0,6294	-0,2501	0,2139	-0,1372	0,0439	-0,0980
B <sub>3</sub>	-0,3488	0,3479	0,6054	0,2963	0,0844	-0,1311
B <sub>4</sub>	0,0053	-0,1682	0,3179	0,2192	0,1941	0,3460
B <sub>5</sub>	0,0417	-0,0440	-0,2703	0,1511	-0,3137	-0,3484
B <sub>6</sub>	0,2435	0,0256	-0,1877	-0,2441	-0,0731	-0,1385

Tableau 1 Corrélations usuelles entre axes factoriels botaniques (B<sub>1</sub> à B<sub>6</sub>) et axes factoriels polliniques (P<sub>1</sub> à P<sub>6</sub>).

Sigles	Taxons	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>
AREL	<i>Elaeis guinaeensis</i>	–	46	–
ARPO	<i>Podococcus</i>	–	50	–
ARSC	<i>Sclerosperma</i>	–	28	–
BABW	<i>Balanites wilsoniana</i>	55	70	–
BURS	<i>Burseraceae</i>	25	–	–
CADI	<i>Dialium</i> sp.	33	–	–
DITE	<i>Tetracera</i> sp.	–	82	–
EUAA	<i>Anthostema aubryanum</i>	52	22	–
EUAL	<i>Alchornea</i> sp.	–	65	–
EUMA	<i>Macaranga</i>	–	–	27
EUUA	<i>Uapaca</i> sp.	–	–	–
GUSG	<i>Symphonia globulifera</i>	–	225	59
MINE	<i>Newtonia duparquetiana</i>	–	–	23
MOTR	<i>Treculia</i> sp.	432	44	120
MYP A	<i>Pycnantus angolensis</i>	30	–	–
MYSY	<i>Sysygium</i> sp.	184	79	197
RUAM	<i>Aidia micrantha</i>	–	–	51
RUTF	<i>Fagara</i> sp.	–	–	177
SAGG	<i>Ganophyllum giganteum</i>	24	–	–
SAPF	<i>Pancovia floribunda</i>	–	–	39
SPTR	<i>Tridemostemon</i>	–	–	26
TIGR	<i>Grewia</i> sp.	–	–	30

**Tableau 2** Liste des taxons à fortes contributions sur les trois premiers axes de l'AFC des pollens. Les axes sont numérotés F<sub>1</sub>, F<sub>2</sub> et F<sub>3</sub>. Les valeurs inscrites en regard des taxons sont leurs contributions lorsque celles-ci dépassent 2% de l'inertie de l'axe.

Sigle	Taxons	Corrélation	Fréquence du taxon
AATA	<i>Trichoscypha acuminata</i>	0,3250	21
ACTB	<i>Thomandersia</i> sp.	0,9960	1
APPM	<i>Pleiocarpa mutica</i>	0,5938	6
GUAF	<i>Allanblackia.floribunda</i>	0,4591	14
GUMA	<i>Mammea africana</i>	0,6514	1
GUSG	<i>Symphonia globulifera</i>	0,8152	11
HYHU	<i>Hymenocardia ulmoides</i>	0,7713	10
LEPM	<i>Petersianthus macrocarpus</i>	0,2549	2
MIPE	<i>Pentaclethra eetveldeana</i>	0,7555	8
MIPM	<i>Pentaclethra macrophylla</i>	0,5225	25
MYSY	<i>Sysygium</i> sp.	0,8629	29
RHAM	<i>Anisophyllea</i> sp.	0,3695	3
RUAM	<i>Aidia micrantha</i>	0,4389	31
SAAL	<i>Allophyllus</i>	0,6654	6

**Tableau 3** Liste des taxons pour lesquelles la corrélation entre l'arbre et le pollen est significative (à 5% de risque).

Si l'on se limite aux paires homogènes on obtient une liste très restreinte de quatorze taxons dont certains ont des fréquences très faibles dans les relevés polliniques (Tableau 3).

Dans cette liste seuls trois taxons (GUSG, MYSY et RUAM) ont déjà été signalés comme ayant de fortes contributions aux axes factoriels de l'AFC (Tableau 2).

### Premier essai de détermination de la forêt associée à un spectre pollinique

Les résultats du paragraphe précédent nous ont paru trop parcellaires pour être utilisés de façon fiable. D'autre part il est clair que la production de pollens est très variable suivant les taxons, et que la dissémination de ces pollens est soumise à de nombreux aléas. C'est pourquoi il nous a paru plus raisonnable de ne prendre en compte que la présence ou l'absence des taxons dans les prélèvements de pollens, comme pour la végétation. Des AFC basées sur cette information restreinte nous ont confirmé que les stations restent distinguables les unes des autres, aussi bien sur le plan botanique que d'après les données polliniques.

Et pour mettre en relief les relations pollen-végétation nous avons construit un tableau de contingence « croisant » les taxons polliniques avec les espèces botaniques : à l'intersection de la ligne *i* et de la colonne *j* de ce tableau nous avons placé le nombre de parcelles contenant à la fois l'espèce botanique *i* et le taxon pollinique *j*. Dans ce contexte il n'est plus nécessaire d'harmoniser espèces botaniques et taxons polliniques. Le tableau analysé comporte, dans notre cas, 352 lignes (espèces botaniques) et 112 colonnes (taxons polliniques). L'AFC de ce tableau permet de dégager les principaux groupes de taxons les plus révélateurs des principaux types de forêts. Cependant elle ne permet pas d'apprécier l'homogénéité des parcelles provenant d'une même station, ni sur le plan botanique ni sur le plan pollinique. Fort heureusement il existe une solution à ce problème qui consiste à accoler au tableau de contingence ci-dessus les deux tableaux, botanique et pollinique, qui en sont à l'origine comme indiqué sur la Figure 4. Ces deux tableaux, mis en éléments supplémentaires (passifs) n'interfèrent donc pas avec l'AFC du tableau de contingence mais fournissent un moyen de contrôler la validité d'une telle analyse.



L'objectif que nous poursuivons maintenant est de trouver pour un spectre pollinique quelconque le type de forêt dont il peut provenir. Pour cela nous appliquons l'analyse factorielle discriminante en prenant pour variables explicatives les trois premiers facteurs de l'analyse précédente (paragraphe précédent). C'est-à-dire que l'on entre dans ce programme, d'une part, les types de forêts (numéros arbitraires entre 1 et 6) et, d'autre part, les coordonnées des spectres polliniques correspondants sur les trois premiers axes factoriels. Ces coordonnées sont en effet disponibles puisque les relevés polliniques ont été mis en éléments supplémentaires dans l'AFC.

Les résultats obtenus sont très satisfaisants puisque, en appliquant la règle d'affectation des spectres au centre de gravité du groupe le plus proche (dans l'espace factoriel), on obtient 45 spectres « bien classés » (c'est-à-dire classés dans le groupe auquel ils appartiennent) sur 46 spectres au total, soit 97,8 % de bien classés.

Naturellement un tel résultat ne reflète pas forcément le taux de succès qu'on pourrait avoir avec des spectres fossiles puisque les spectres actuels ont eux-mêmes contribué à élaborer la règle d'affectation. C'est pourquoi nous avons utilisé la méthode de l'observation écartée : la règle d'affectation d'un spectre  $s$  particulier est basée sur l'ensemble des spectres sauf  $s$ . Avec cette méthode et toujours à partir des trois premiers facteurs de l'AFC le taux de succès reste inchangé : 97,8 % de bien classés. Cet excellent résultat est probablement dû à la grande stabilité des axes de l'AFC et à la forte individualité des groupes de relevés.

### Conclusion

Les principaux enseignements de la présente étude peuvent se résumer comme suit. L'analyse des seules données botaniques, ou l'analyse des seules données polliniques, montre que la forêt du Mayombe au Congo est très diversifiée, et que chaque station fait preuve d'une forte spécificité. Cependant l'étude des corrélations entre taxons polliniques et espèces botaniques, même si elle fait apparaître des groupes de taxons et d'espèces caractéristiques des stations échantillonnées, ne semble pas permettre à elle seule d'asseoir sur des bases solides la reconstitution du faciès forestier à partir des pourcentages de pollens.

En revanche la prise en compte simultanée de l'ensemble de tous les taxons polliniques et de toutes les espèces botaniques dans un tableau de co-occurrences, ne faisant pas intervenir les abondances, s'avère être une méthode prometteuse. L'apparente perte d'information qu'on pourrait craindre en ignorant les abondances est compensée par l'élimination des fluctuations aléatoires liées aux productions et aux dispersions très hétérogènes des pollens.

Par ailleurs une astuce technique dans le traitement statistique, la mise en éléments supplémentaires passifs des prélèvements polliniques, permet d'obtenir des coordonnées factorielles stables pour ces prélèvements, coordonnées à partir desquelles on peut alors facilement déterminer le faciès forestier correspondant.

Ces excellents résultats ne sont qu'une première étape dans la reconstitution des paléo-végétations car il

reste quelques problèmes en suspens. En effet, ces résultats sont liés à la forte individualité des stations retenues initialement pour faire les relevés botaniques. Que deviendraient-ils avec un choix différent de stations, en maintenant constant le nombre de relevés ? ou bien que se passerait-il avec un effort d'échantillonnage plus important, en doublant ou triplant le nombre de relevés, par exemple ?

Par ailleurs les prélèvements polliniques ont été dans notre étude des échantillons de sols ; lorsque qu'on voudra reconstituer les forêts passées, il nous faudra utiliser des tranches de sédiments déposés dans des lacs ou dans des tourbières pour lesquels l'accumulation des pollens est sans doute différente de celle des sols.

Quoi qu'il en soit de ces problèmes, la qualité des résultats obtenus, par des méthodes relativement frustes, nous permet d'être optimistes pour la reconstitution des paléo-végétations tropicales humides.

### Références

- Aitchison, 1986. *The analysis of compositionnal data*. Chapman & Hall, London.
- Edgington, E. S. 1987. *Randomization tests*, 2nd ed., Marcel Dekker, New-York.
- Manly, B. 1991. *Randomization and Monte-Carlo methods in Biology*. Chapman & Hall, London. 281 p.
- Mantel, L. 1967. The detection of disease clustering and a general regression approach. *Cancer Research*, 27, 209-220.
- Reynaud, I. 1995. Histoire des paléoclimats forestiers du Sud-Cameroun à partir d'analyses palynologiques et statistiques de dépôts holocènes et actuels. Thèse. Université Montpellier. 98 p.
- Romeder, J.-M. 1973. *Méthodes et programmes d'analyse discriminante*. Dunod, Paris.

# Dynamique à long terme des écosystèmes forestiers intertropicaux

MICHEL SERVANT, SIMONE SERVANT-VILDARY,  
ÉDITEURS SCIENTIFIQUES



IRD

UNESCO

MAB

CRS



Les responsables d'édition adressent leurs sincères remerciements à  
Christian Levêque, Samy Mankoto, Bernard Riéra et Léo Rona-Beaulieu.

Ouvrage publié avec le soutien de :

Centre national de la recherche scientifique, Programme Environnement,  
vie et sociétés, 3, rue Michel-Ange, F-75016 Paris

UNESCO, 7 place de Fontenoy, F-75007 Paris  
Programme sur l'Homme et la Biosphère (MAB)  
Projet PNUD ZAI/97/001-ERAIFT

Ministère des affaires étrangères  
Comité MAB France

IRD (Institut de recherche pour le développement),  
313, rue Lafayette, F-75010 Paris

ISBN 92-3-203753-X  
Mise en page : Valérie Herman  
Impression : Imprimerie Jouve  
Photo de couverture : Lac Tabéré, Adamaoua, Cameroun

© UNESCO 2000