

TRAITEMENT STATISTIQUE DE DONNEES PALYNOLOGIQUES DE L'Holocène DU LAC TANMA (SENEGAL)

J. GUIOT & J. MEDUS

Laboratoire de Botanique, Historique et Palynologie.
Faculté des Sciences St. Jérôme 13397-Marseille 13 F.

(Recibido el 7 de Junio de 1984)

RESUME. Le traitement statistique de données palynologiques de l'Holocène du Lac Tanma situé au Sénégal, au Nord de Dakar, au moyen de l'analyse canonique permet de préciser le schéma climatique exposé antérieurement. Il est notamment mis en évidence une pulsation positive de la pluviosité ca. 6000 BP.

RESUMEN. El procesamiento de datos palinológicos del Holoceno del lago Tanma situado al norte de Dakar en Senegal por medio de un análisis canónico ha permitido precisar la interpretación climática anteriormente propuesta, y particularmente, poner de manifiesto una pulsación positiva de la lluvia ca. 6000 BP.

INTRODUCTION

Le caractère édaphique prononcé des variations de fréquences dans les diagrammes palynologiques de l'Holocène de sites cotiers du Sénégal a constitué une difficulté pour leur interprétation climatique (MEDUS, 1984). L'influence notamment des variations de fréquences de *Rhizophora* dans ces milieux pose véritablement une question en préalable à toute tentative dans ce sens quand on sait la sensibilité de ce genre aux facteurs géomorphologiques. Toutefois, comme l'écarter complètement pour n'en faire qu'un indicateur de trasgression ne paraît pas non plus une solution satisfaisante dans la mesure où, manifestement, une part non négligeable des conditions géomorphologiques (i.e. la qualité du substrat) optimales nécessaires à la prospérité d'une mangrove à *Rhizophora* dépend de facteurs climatiques à travers la végétation, il a semblé intéressant d'effectuer une analyse plus poussée de l'information disponible dans les variations de fréquences de ce taxon sur le plan climatique.

Pour cela, on peut considérer que deux sites qui présentent des éléments de végétation en commun se trouvent nécessairement soumis à des influences communes. Si ces sites sont suffisamment distants l'un de l'autre pour que les contraintes édaphiques soient différentes, les

similitudes que ces deux sites ont en commun sur le plan phytogéographique proviennent alors d'une influence climatique identique (en l'absence de facteur anthropique). Il est donc nécessaire de disposer d'une méthode qui permettent de traiter simultanément deux groupes de variables pour extraire des données les facteurs qui influencent à la fois les deux groupes. L'analyse canonique, qui n'est pas utilisée ici dans l'esprit des fonctions de transfert classiques (WEBB & al., 1972) est dans ce cas toute indiquée puisqu'elle a pour but de rechercher des directions communes de variations des variables (axes canoniques). Avec cette méthode, si l'abondance d'un taxon est dans les sites considérés lié au climat il apparaîtra alors lié à un des principaux axes canoniques.

Trois des diagrammes présentés par ailleurs (MEDUS, sous presse) feront l'objet de cette étude, il s'agit des diagrammes S1, S3 et S4. Les diagrammes sont analysés par couple. Les variables seront les onze taxons assez constamment représentés et dont les variations de fréquences ne sont pas négligeables. Ce sont: *Adansonia* (ADAN), *Alchornea* (ALCH), *Ceiba* (CEIB), *Cyperaceae* (CYPE), *Elaeis* (ELAE), *Lygodium* (LYGO), *Phoenix* (PHOE), *Poaceae* (POAC), *Rhizophora* (RHIZ), *Syzygium* (SYZY), *Typhaceae* (TYPH).

TECHNIQUES

Notons X et Y (cf. Annexe) les deux matrices de dimension (n_x, t) et (n_y, t) composant le couple analysé: n_x et n_y sont les nombres de niveaux des deux carottes correspondantes et t , le nombre de taxons utilisés ($t=11$).

Avant de procéder à l'analyse canonique, il est nécessaire de synthétiser les deux matrices X et Y par une analyse en composantes principales dont les équations correspondantes sont (LEBART & al., 1977):

$$X'XA_x = A_x \lambda_x \quad \text{et} \quad Y'YA_y = A_y \lambda_y \quad (1)$$

avec $A'_x A_x = I$ $A'_y A_y = I$

CP	S1		S3		S4	
	VP	%	VP	%	VP	%
1	4.95	45.0	4.71	42.8	3.92	35.6
2	2.60	68.5	2.12	62.0	2.67	59.9
3	1.13	78.7	1.48	75.5	1.33	72.0
4	1.05	88.4	1.17	86.2	0.98	80.9
5	0.64	94.2	0.95	94.8	0.74	87.6
6	0.43	98.1	0.29	97.4	0.63	93.4
7	0.12	99.8	0.17	98.9	0.34	96.4
8	0.05	100.0	0.08	100.0	0.22	98.4
9	0.00	100.0	0.00	100.0	0.08	99.2
10	0.00	100.0	0.00	100.0	0.05	99.6
11	0.00	100.0	0.00	100.0	0.03	100.0

Tableau 1 : Valeurs propres (VP) et pourcentages cumulés de variance pour les trois diagrammes.

$X'X$ est une matrice de corrélation au facteur $1/nx$, dont la matrice des valeurs propres est Λ_x et celle des vecteurs propres est A_x , leur dimensions sont (t,t) et l'équivalent pour Y .

Le Tableau 1 présente les valeurs propres de ces deux analyses et les pourcentages cumulés de la variance pour les trois diagrammes.

Comme les deux premières composantes principales (CP) expliquent 60 à 70% de la variance totale (ce pourcentage est fonction inverse du nombre de niveaux inclus dans l'analyse), il suffit pour la suite de ne traiter que celles qui se trouvent sur le Tableau 2.

Nom	C.P.1.			C.P.2.		
	S1	S3	S4	1	3	4
(+) CEIB 17	ELAE 15	PHOE 16	ALCH 15	CEIB 5	ADAN 16	
ELAE à (22)	PHOE à (22)	TYPH 17(22)	PHOE 6	LYGO 12	CEIB 17	
CYPE 19	SVZY 20	CYPE 8	ALCH 2	SVZY 7	ELAE	
			7	SVZY		
(-) POAC 5	POAC 6	RHIZ 12b	RHIZ 20	THYP 17	POAC 7	
LYGO à (21)	LYGO à (21)	ADAN 15 (21)	SVZY 11	CYPE 21	LYGO à	
RHIZ 14	RHIZ 12	22a(23)	ADAN 25	ADAN		11
ADAN	ADAN					

Tableau 2 : Principaux taxons et niveaux influençant les deux premières C.P. des trois sites. Les taxons et les niveaux sont rangés par ordre d'importance décroissante. On distingue influence positive (+) ou négative (-).

Dans ce Tableau, alors que par rapport à la seconde composante principale, les sites ne montrent que peu de cohérence au niveau des taxons, celle-ci apparaît assez grande entre les sites 1 et 3 par rapport à la première composante principale; le site 4 étant assez distinct des deux autres de ce point de vue.

Le calcul des coordonnées des niveaux dans le système des composantes principales s'effectue à l'aide des équations suivantes:

$$F_x = XA_x \quad F_y = YA_y \quad (2)$$

F_x et F_y sont les matrices (nx,t) et (ny,t) qui permettent de projeter les niveaux des diagrammes x et y sur les composantes principales. Pour tenir compte de la variance, les matrices A_x et A_y sont multipliées par $\Lambda_x^{1/2}$. De cette façon, les plans des deux premières composantes principales peuvent être analysés (cf. Tab. 2) et ceci permet de montrer que la composante principale 1 semble partager les diagrammes en trois parties: avant le niveau 20 (Zone 3, Z3), entre les niveaux 20 et 15 (Zone 2, Z2) et après le niveau 15 (Zone 1, Z1) alors que la seconde composante principale ne suggère rien de précis.

A ce stade de l'analyse, l'on dispose de deux matrices A_x et A_y . Les dernières composantes principales expliquant peu de la variance et contenant surtout du bruit sont éliminées pour la suite du calcul par l'utilisation du critère de PVP (GUIOT, 1981). Les six premières composantes principales des trois diagrammes ($t'=6$) sont alors retenues. Il est nécessaire maintenant de trouver une dimension commune à ces

diagrammes. Mais dans la mesure où l'on ne saurait affirmer que tous les niveaux sont réellement synchrones, la seule dimension commune disponible est constituée par les taxons. Les six premières colonnes de A_x et A_y vont être mises en corrélation par une analyse canonique; les trois sites étant analysés par couple et le couple (1/3) devant être rejeté par manque de niveaux communs, seuls les couples (1/4) et (3/4) seront considérés.

La méthode de l'analyse canonique développée par HOTTELING (1936), ultérieurement décrite par différents auteurs (LEBART & al., 1977); CLARK, 1975), permet de découvrir les facteurs qui influencent simultanément deux groupes de variables.

En partant des composantes principales on peut définir une combinaison linéaire pour x:

$$u_i(x) = \sum_{j=1}^{t'} a_{ij}(x) c_j(x) \quad i=1, \dots, t' \quad (3)$$

Et il en serait de même pour y.

Le vecteur $u(x) = (u_1(x) \dots u_{t'}(x))$ donne la valeur de cette combinaison pour tous les niveaux. Le vecteur $c(x) = (c_1(x), \dots, c_{t'}(x))$ donne le poids des t' composantes principales dans cette combinaison linéaire. La corrélation entre $u(x)$ et $u(y)$ peut être calculée et elle dépend bien sûr des coefficients $c_j(x)$ et $c_j(y)$ ($j=1, \dots, t'$). La valeur de ces coefficients qui rend maximale la corrélation sera le premier axe canonique. On peut également chercher une autre combinaison linéaire indépendante de la première et qui rend également maximale la corrélation (bien sûr inférieure à la première) et ainsi de suite. Comme les t' variables sont indépendantes, il y aura t' axes canoniques. Ils sont classés par ordre de corrélation décroissante. Ces corrélations sont appelées corrélations canoniques. t' peut être différent pour x et y, le nombre d'axes sera alors le minimum des dimensions des matrices A_x et A_y . La solution du problème passe, pour finir, par le calcul d'une matrice C_x et d'une matrice C_y dont chaque colonne définit les poids des variables de départ sur un axe canonique.

On démontre finalement que C_y est la matrice des vecteurs propres de:

$$(A_y' A_y)^{-1} A_y' A_x (A_x' A_x)^{-1} A_x' A_y \quad (4)$$

La matrice diagonale des valeurs propres Δ^2 fournit le carré des corrélations canoniques. On peut déduire C_x directement de C_y et de Δ :

$$C_x = (A_x' A_x)^{-1} A_x' A_y C_y \Delta^{-1} \quad (5)$$

Les matrices U_x et U_y (généralisant à t' dimension les vecteurs $u(x)$ et $u(y)$) donnent la projection des taxons sur les axes canoniques. On a:

$$U_x = A_x C_x \quad U_y = A_y C_y \quad (6)$$

avec $U_x' U_x = I$ (matrice unité dim. t'), $U_y' U_y = J$ et $U_x' U_y = \Delta$

On peut trouver le niveau de signification de ces corrélations (et par là, le nombre maximum de celles-ci à analyser) par les formules suivantes:

$$\text{lam} = \prod_{i=1}^c (1 - \lambda_i) \quad \text{lambda de Wilks (c=nb axes pris)}$$

$$X^2 = -(nx-1-(2t'+1)/2) \ln(\text{lam}) \quad \text{avec } (t'-c)(t'-c) \text{ d.l.} \quad (7)$$

Le Tableau 3 montre que les trois premières au maximum sont intéressantes.

C	Couple 1/4			Couple 3/4		
	corr. can.	chi2	ndf	corr. can.	chi2	ndf
1	1.000	51.8***	36	1.000	163.7***	36
2	1.000	51.8***	25	1.000	51.8***	25
3	0.939	27.8 *	16	0.996	33.6 **	16
4	0.796	8.0	9	0.895	11.5	9
5	0.732	3.5	4	0.768	4.3	4
6	0.098	0.0	1	0.235	0.3	1

Tableau 3 : Corrélations canoniques, chi-carrés et nombres de degrés de liberté. ***, ** et * indiquent un seuil de signification de 0.001, 0,01 et, 0.05).

Le Tableau 3 montre également que le couple 3/4 est plus homogène que le couple 1/4 puisque ses corrélations canoniques sont plus élevées. Le Tableau 4 présente les taxons qui influencent le plus les trois axes canoniques significatifs.

Signes	1/4			3/4		
	1	2	3	1	2	3
(+)	CEIB ADAN	LYGO	RHIZ	CYPE ALCH TYPH	ADAN	CEIB ALCH
(-)	POAC	TYPH RHIZ	CYPE ELAE PHOE	RHIZ	ELAE PHOE SYZY	TYPH CYPE

Tableau 4 : Taxons influençant le plus les 3 premiers axes canoniques.

Le Tableau 4 montre qu'il n'y a pas de cohérence entre le couple 3/4 et le couple 1/4 excepté entre l'axe 1 de 3/4 et l'axe 3 de 1/4 (en signe opposé).

Si on multiplie les matrices C_x et C_y par F_x et F_y respectivement, on a les chronologies correspondantes:

$$F_x C_x \quad \text{et} \quad F_y C_y$$

dessinées aux Figs. 1 et 2 pour le couple 1/4 et les Figs. 3 et 4 pour le couple 3/4. On distingue bien une corrélation négative entre l'axe 1 (3/4) et 3 (1/4).

DISCUSSION

Dans l'analyse précédente, tout concorde pour qu'il soit donné un sens climatique à ce couple d'axes dominé par *Rhizophora*. Pour l'interpréter, il faut remarquer au préalable que:

1°. Les taxons autres que *Rhizophora* impliqués par ce couple d'axes se rapportent à des plantes (Cyperaceae, *Alchornea* (*A. chordifolia*), *Elaeis* (*E. guineensis*), *Typha*, *Phoenix*) qui se signalent par un besoin en eau douce temporaire, ou continu, non négligeable.

2°. La prospérité d'une mangrove est fonction (pour une part, l'autre ayant trait au niveau de la mer) d'un apport continu et suffisant de particules fines pour renouveler le substrat; or un tel apport n'est réalisé que par un écoulement hydrique suffisant à travers un filtre (le tapis végétal) efficace.

Il s'ensuit que ce couple d'axes exprime selon toute vraisemblance la pluviosité. Et, les deux événements majeurs qui apparaissent dans la succession (Fig. 2, axe 3 et Fig. 4, axe 1) indiquent deux moments durant lesquels, la pluviosité a cru suffisamment pour influencer la végétation. Le premier de ces moments se centre vers 9000 - 10000 BP et le second ca 6000 BP.

Lors de la tentative de reconstitution de la végétation et du climat holocènes de la côte septentrionale du Sénégal, les variations des fréquences observées dans la zone datée 9000 - 10000 BP, avaient été clairement attribuées à une augmentation de la pluviosité (MEDUS, 1984). En revanche, l'importance des variations des fréquences qui se situaient ca 6000 BP avaient été grandement sous-estimées; car, effectivement interprétées comme la manifestation d'un marécage, il avait été supposé que cette phase de marécage provenait de la rétraction d'un lac alors qu'elle correspond, selon l'analyse ci-dessus, à un regain de la pluviosité. Or, cette pulsation positive de la pluviosité correspond bien à la seconde "phase humide" ou "période lacustre" holocène déjà discutée (BEAUDET & al., 1976; ELOUARD, 1973; NICHOLSON, 1978) et située au Nouakchottien sur la base de critères pédologiques et géomorphologiques.

L'analyse statistique de données palynologiques précise donc le schéma proposé précédemment. En outre, cette deuxième pulsation positive de la pluviosité permet maintenant de mieux comprendre comment la niaye a pu résister apparemment au déficit de l'approvisionnement hydrique manifeste vers 7700 BP. Par ailleurs, les fréquences relativement conséquentes de *Typha*, *Lycopodium* (cf. *L. cernuum*), *Alchornea* et *Cyperaceae* qui apparaissent dans les deux niveaux supérieurs du diagramme de Keur Marsal (MEDUS & MONTEILLET, 1981) trouveraient peut-être elles aussi une explication pertinente dans le cadre de ce schéma plus précis.

REFERENCES

- BEAUDET, G., P. MICHEL, D. NAHON, P. OLIVA, S. RISER & A. RUELLAN (1976). Formes, Formations superficielles et variations climatiques récentes du Sahara occidental. *Rev. Géographie physique et Géol. Dyn* (2). 18:157 - 174.
- CLARK, D. (1975). *Understanding canonical correlation analysis*. University of East Anglia, cat. no. 3, Norwich.
- ELOUARD, P. (1973). Oscillations climatiques de l'Holocène à nos jours en Mauritanie Atlantique

et dans la vallée du Sénégal. Colloque de Nouakchott sur les problèmes de la desertification, prétrirage.

- GUIOT, J. (1981). Analyse mathématique de données géophysiques. Applications à la dendroclimatologie. Vol. 1. Thèse Univ. Catho. Louvain. Fac. Sciences.
- HOTTELING, H. (1936). Relation between two sets of variables. *Biometrika* 28:129 - 149.
- LEBART, L., A. MORINEAU & N. TABARD (1977). Techniques de la description statistique. Dunod. Paris.
- MEDUS, J. (1984). Essai de reconstitution de la végétation et du climat holocènes de la côte septentrionale du Sénégal. *Rev. Palaeobot. Palynol.* 41:31 - 38.
- _____ Analyse pollinique des sédiments holocènes du Lac Tanna, Sénégal. *Palaeoecology of Africa*. 16. Sous presse.
- _____ & J. MONTEILLET (1981). Données préliminaires sur la palynologie de l'Holocène du delta du fleuve Sénégal. *Géobios*. 14:801 - 805.
- NICHOLSON Sh., E. (1978). Comparison of Historical and recent african rainfall anomalies with Late Pleistocene and early holocene. In: ZINDEREN-BAKKER (ed.). *Palaeoecology of Africa*, ed. 10. 99 - 123.
- WEBB III, T., R. A. BRYSON (1972). Late and postglacial climatic change in the northern midwest U.S.A.: quantitative estimates derived from pollen spectra by multivariate analysis. *Quaternary Research*. 2:70 - 115.

LEGENDE DES FIGURES

- FIGURE 1. Représentation des niveaux dans l'espace des six premiers axes canoniques ($F_x C_x$) pour le site 1 mis en corrélation avec le site 4.
- FIGURE 2. Représentation des niveaux dans l'espace des six premiers axes canoniques ($F_y C_y$) pour le site 4 mis en corrélation avec le site 1.
- FIGURE 3. Représentation des niveaux dans l'espace des six premiers axes canoniques ($F_x C_x$) pour le site 3 mis en corrélation avec le site 4.
- FIGURE 4. Représentation des niveaux dans l'espace des six premiers axes canoniques ($F_y C_y$) pour le site 4 mis en corrélation avec le site 3.

Sondage S1 Comportant 10 niveaux

RHIZ	39	36	42	60	32	48	45	12	45	39
POAC	30	32	27	20	31	26	12	5	13	12
CYPE	5	4	6	0	3	2	5	13	6	8
TYPH	10	1	1	1	5	1	5	7	5	6
LYGO	1	3	2	3	4	1	1	1	1	1
ELAE	1	0	0	0	0	0	1	9	2	0
CEIB	1	0	0	1	1	1	2	3	2	2
ADAN	0	0	0	0	0	1	2	1	1	2
ALCH	3	4	3	2	3	2	1	4	2	4
SYZY	0	0	1	0	0	1	1	1	1	0
PHOE	1	1	1	0	1	1	1	1	0	1

Sondage S3 Comportant 13 niveaux

RHIZ	35	57	44	42	38	47	87	16	4	21	6	43	26
POAC	36	23	25	21	29	18	3	9	5	15	4	10	8
CYPE	1	1	1	6	2	8	3	45	14	26	24	25	31
TYPH	1	0	0	2	5	8	2	6	2	1	2	3	5
LYGO	6	3	3	2	3	1	0	0	0	1	1	0	0
ELAE	0	0	0	0	0	0	0	0	15	6	17	2	10
CEIB	1	1	1	1	1	1	1	1	1	1	1	1	0
ADAN	1	0	0	1	1	1	1	1	0	0	0	0	0
ALCH	3	3	3	2	2	1	2	6	4	4	4	2	1
SYZY	0	0	0	1	1	1	0	1	19	6	8	1	1
PHOE	0	0	0	1	0	1	1	0	3	1	3	1	1

Sondage S4 Comportant 18 niveaux

RHIZ	40	40	38	36	50	95	85	94	32	16	20	83	56	52	62	66	35	67
POAC	33	36	25	31	25	0	2	1	7	7	17	4	4	4	7	10	17	10
CYPE	1	1	2	6	4	1	3	1	7	16	3	4	5	0	4	2	4	3
TYPH	5	3	5	5	3	1	3	1	5	10	6	1	2	0	4	2	5	6
LYGO	1	5	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
ELAE	0	0	0	0	0	0	0	0	12	4	1	1	1	0	1	2	1	2
CEIB	1	0	1	1	0	0	1	1	3	10	1	1	1	14	1	2	1	1
ADAN	0	0	0	1	0	0	1	1	0	1	0	1	1	1	1	1	0	1
ALCH	1	1	1	4	3	0	2	1	4	2	2	1	2	0	1	0	0	2
SYZY	0	0	0	1	0	0	0	1	2	1	1	1	1	0	1	0	0	0
PHOE	1	1	1	2	1	0	0	0	3	1	1	0	1	0	0	0	1	0

ANNEXE: Tableau des données (Exprimées en pourcents).

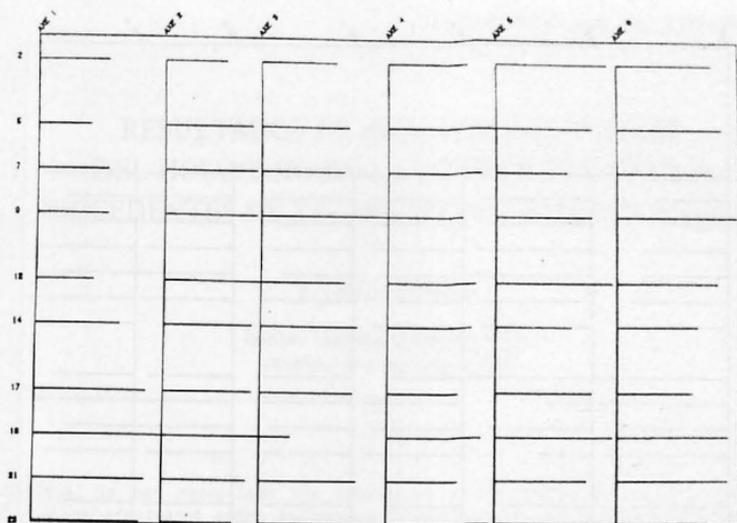


Figure 1

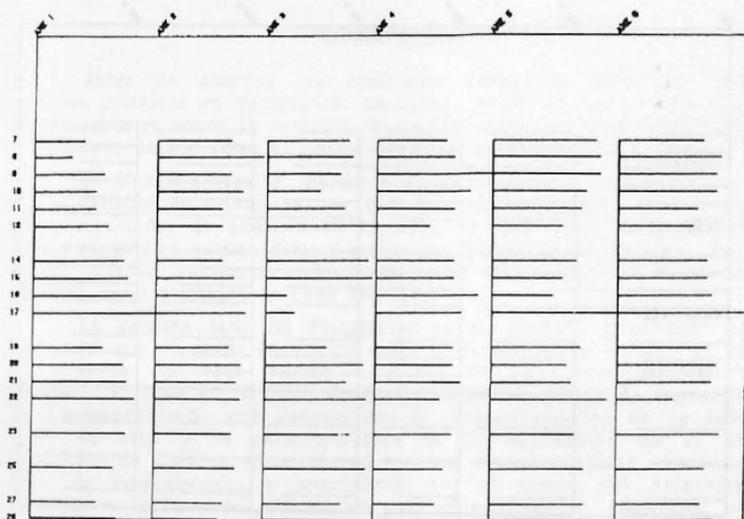


Figure 2

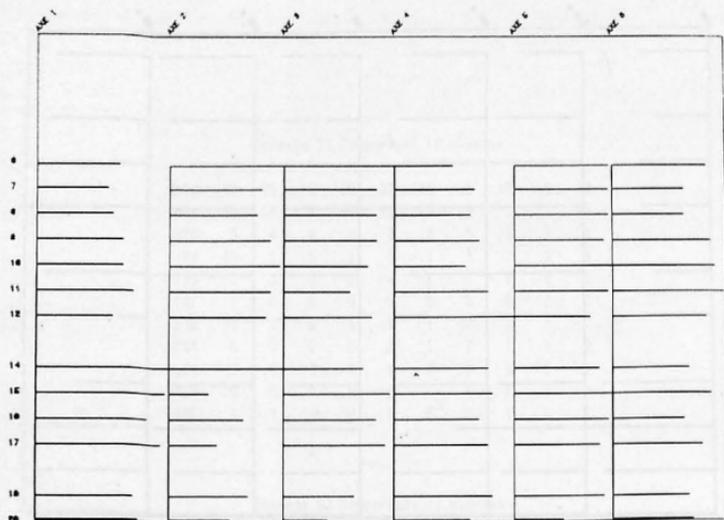


Figure 3

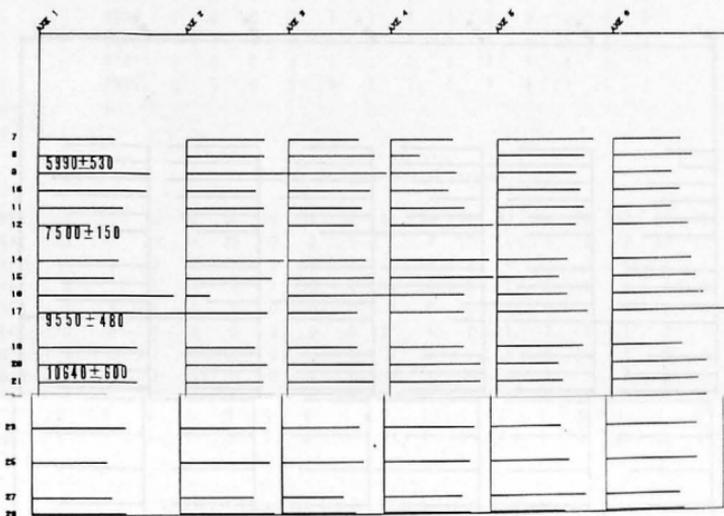


Figure 4