

# Détermination de poissons du genre *Clarias* par une méthode généralisable aux poissons non écaillés<sup>(1)</sup>

par C. DEBOUCHE, J.-P. MARQUET et H. TEUGELS

## SOMMAIRE

Résumé & Summary.....	844
Introduction.....	845
1. Méthodologie statistique.....	846
2. Utilisation de l'analyse discriminante.....	850
3. Test visuel.....	857
4. Généralisation et conclusions.....	860
Bibliographie.....	861

## RÉSUMÉ

Dans le but de différencier plusieurs espèces de poissons sans écailles à l'intérieur d'un même genre (le genre *Clarias*), les auteurs ont travaillé sur la mise au point d'une méthode biométrique simple.

Cette méthode s'appuie sur différentes mesures prises sur deux espèces de *Clarias* (*Clarias lazera* et *Clarias anguillaris*) traitées par analyse discriminante. Les auteurs examinent ensuite les conditions d'application de cette méthode en vue de sa généralisation.

(1) Cette étude reprend des données d'une thèse d'Ingénieur agronome (Élevage) soutenue le 3 novembre 1978 par J.-P. MARQUET à la Faculté des Sciences agronomiques de B-5800 Gembloux, Belgique. Cette thèse a été réalisée sous la direction du Professeur Dr. C. REIZER (Centre interuniversitaire des Sciences de l'Environnement, Fondation universitaire luxembourgeoise, B-6700 Arlon) avec la collaboration, pour la partie systématique, du Professeur Dr. D. THYS VAN DEN AUDENAERDE (Musée royal de l'Afrique centrale, B-1980 Tervuren).

## SUMMARY

In order to differentiate several species of scaleless fish inside a same genus (the *Clarias* genus), the authors have worked on setting up a simple biometric method.

This method is based on various measurements taken on two species of *Clarias* (*Clarias lazera* and *Clarias anguillaris*) treated by discriminant analysis. Further, the authors examine the conditions of application of this method with a view to its generalization.

## INTRODUCTION.

Afin d'exploiter rationnellement son environnement, l'homme a de tout temps cherché à classer les éléments qui le constituent. Un groupe d'individus est défini par un ensemble de caractères héréditaires, subdivisés le plus souvent en caractères morphologiques et en caractères physiologiques.

Dans la super-classe des Poissons, certains caractères morphologiques, qualifiés de méristiques, se prêtent à une numération précise (VIBERT & LAGLER, 1961). Ainsi, les comptages d'écaillés le long d'un axe de référence rendent souvent de précieux services au taxonomiste.

Si on envisage la détermination spécifique de poissons non écaillés, le problème se pose avec plus d'acuité, et ce d'autant que les autres caractères méristiques sont dépendants de l'environnement. Il est alors nécessaire d'envisager des méthodes de prospection biométriques plus sophistiquées afin d'éliminer cette dépendance. Cette démarche nécessite cependant quelques restrictions car si les poissons ont une croissance continue, certaines de leurs proportions peuvent varier avec l'âge.

Ainsi, dans le problème qui nous occupe, la population des poissons du genre *Clarias* du bassin du Sénégal est divisée en deux espèces, à savoir *Clarias lazera* et *Clarias anguillaris*. Leur détermination est réalisée à partir du rapport de leur nombre de branchiospines sur leur longueur standard. L'utilisation de ce critère, seul connu à ce jour (DAGET, 1954 et DE HUISSER, 1974), nécessite le sacrifice de l'individu pour prélever son arc branchial.

Notre étude a pour but de définir un certain nombre de mesures à prendre sur l'animal pour parvenir à le déterminer sans devoir

le sacrifier. Nous apporterions alors une preuve supplémentaire de l'opportunité de ces deux espèces parfois groupées sous le nom de « complexe *anguillaris-lazera* ».

### 1. MÉTHODOLOGIE STATISTIQUE.

Le matériel disponible pour établir cette règle de classement est un lot de *Clarias* identifiés comme *lazera* pour certains d'entre eux ( $n_1$ ) et comme *anguillaris* pour les autres ( $n_2$ ).

La répartition du lot de *Clarias* étant ainsi réalisée, on effectue sur tous les individus un certain nombre de mensurations sur l'extérieur de l'animal.

Les individus étant d'âges inconnus, les mensurations réalisées peuvent être plus influencées par leur taille que par leur appartenance à telle ou telle espèce. Il faut donc transformer ces mesures de manière à les rendre indépendantes de la taille et donc approximativement de l'âge des individus. Ce résultat est obtenu en divisant les mesures réalisées par l'une d'entre elles qui soit la plus représentative de la taille.

L'influence de la taille ou de l'âge est donc éliminée par la transformation suivante :

$$x'_{ij} = \frac{x_{ij}}{x_{ik}}$$

où  $x_{ij}$  est la quantité correspondant à la  $j^{\text{ème}}$  mesure effectuée sur le  $i^{\text{ème}}$  individu ;

$k$  est l'indice de la variable choisie comme la plus représentative de la taille.

Pour chacune de ces variables transformées on peut alors tester l'égalité des moyennes correspondant aux deux espèces (DAGNELIE, 1970, p. 27).

Si on est amené à accepter l'égalité des moyennes des deux espèces pour toutes les mesures corrigées, cela signifie que ces dernières n'ont aucun pouvoir discriminant et ne permettent pas de classer un individu dans l'une ou l'autre espèce. L'étude se termine donc sur un échec.

Si, par contre, on rejette l'hypothèse d'égalité des moyennes des deux espèces pour une série de variables, celles-ci doivent être retenues comme variables éventuellement « discriminantes ». Cela signifie que l'une ou l'autre d'entre elles peut être utilisée pour classer un individu suivant la valeur qu'il a pour cette variable

et selon une règle de classement à définir. Cette dernière consiste à calculer la quantité :

$$y = b_0 + b_1 x$$

où  $x$  est la valeur obtenue dans la variable étudiée par l'individu à classer,  $b_0$  et  $b_1$  sont les coefficients de cette fonction discriminante,

et à classer l'individu dans l'espèce n° 1 si  $y$  est positif et dans l'espèce n° 2 si  $y$  est négatif (DAGNELIE, 1975, p. 307). Les paramètres se calculent par les relations :

$$b_1 = \frac{\bar{x}_1 - \bar{x}_2}{\hat{\sigma}^2}$$

$$b_0 = -b_1 \frac{\bar{x}_1 + \bar{x}_2}{2}$$

où  $\bar{x}_1$  est la moyenne observée sur les individus de l'espèce n° 1 pour la variable étudiée,

$\bar{x}_2$  est la quantité correspondante pour l'espèce n° 2,

$\hat{\sigma}^2$  est la variance commune aux deux populations estimée par la relation :

$$\hat{\sigma}^2 = \frac{SCE_1 + SCE_2}{n_1 + n_2 - 2}$$

Dans cette dernière expression, le symbole SCE représente la somme des carrés des écarts entre les observations et leur moyenne.

Cette règle de classement comporte cependant un risque d'erreur, c'est-à-dire un risque de voir classer comme *lazera* un individu qui est de l'espèce *anguillaris* et réciproquement.

Cette probabilité de classement erroné (PCE) peut se calculer par l'expression :

$$PCE = 1 - \Phi (D/2)$$

où  $\Phi$  désigne la fonction de répartition d'une distribution normale réduite,

D se calcule par l'expression :

$$D^2 = \frac{(\bar{x}_1 - \bar{x}_2)^2}{\hat{\sigma}^2}$$

A chaque variable déclarée « discriminante » par le test d'égalité des moyennes, on peut donc associer une valeur de PCE.

Il se peut que l'une ou l'autre des variables discriminantes présente un risque de classement erroné acceptable. Dans ce cas, le problème du classement d'un individu est terminé.

S'il n'en est pas ainsi, c'est-à-dire si aucune des mensurations,

après correction, ne permet à elle seule de classer un individu avec une sécurité suffisante, on peut essayer d'améliorer la règle de classement en utilisant plusieurs mensurations simultanément ( $p$  par exemple). La fonction discriminante fera donc intervenir  $p$  variables.

Si on peut supposer que les deux populations, correspondant aux deux espèces, possèdent des distributions normales à  $p$  dimensions, de mêmes variances et covariances et de paramètres connus ou estimés de façon suffisamment précise, la fonction assurant le risque de classement erroné minimum est :

$$y = b_0 + \mathbf{b}'\mathbf{x}$$

où  $\mathbf{x}$  est le vecteur des valeurs obtenues dans les  $p$  variables retenues pour l'individu à classer,

$\mathbf{b}$  est le vecteur des coefficients de la fonction discriminante calculé par :

$$\mathbf{b} = \hat{\Sigma}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$$

où  $\bar{\mathbf{x}}_1$  est un vecteur contenant les  $p$  moyennes calculées sur l'échantillon extrait de la première population,

$\bar{\mathbf{x}}_2$  est le vecteur correspondant calculé sur le deuxième échantillon,

$\hat{\Sigma}$  est la matrice de variances et de covariances commune aux deux populations et estimée par la formule

$$\hat{\Sigma} = \frac{\mathbf{A}_1 + \mathbf{A}_2}{n_1 + n_2 - 2}$$

où  $\mathbf{A}_1$  est la matrice des sommes de produits et somme de carrés des écarts calculée sur le premier échantillon,

$\mathbf{A}_2$  est la matrice correspondante calculée sur le deuxième échantillon.

La constante  $b_0$  se calcule par l'expression :

$$b_0 = \left( \frac{\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2}{2} \right)' \mathbf{b}$$

La probabilité de classement erroné se calcule comme énoncé ci-dessus avec toutefois pour la quantité  $D$ , l'expression :

$$D = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathbf{b}$$

Cette règle de classement étant ainsi établie, il reste à choisir les variables ou mensurations à introduire dans la fonction discriminante.

Les premières variables à introduire dans la fonction discriminante sont évidemment celles pour lesquelles les moyennes des deux populations sont les plus significativement différentes. Il ne faut cependant pas s'arrêter à celles-là. En effet, une variable

dont les moyennes des populations ne sont pas différentes peut se révéler utile dans une fonction discriminante lorsqu'elle est combinée avec d'autres variables (COCHRAN, 1964; HÉBRANT, 1973, et LUBISCHEN, 1962) surtout si la corrélation interspécifique liant cette variable aux autres est de signe opposé à leur corrélation intraspécifique. En pratique, nous avons retenu cinq variables parmi celles qui présentaient des moyennes supposées différentes et parmi les autres toutes celles qui présentaient des corrélations significatives avec ces cinq précitées. Cela nous conduit à sélectionner un nombre important de variables qui rendrait la fonction discriminante peu pratique et dénuée de sens.

La solution adoptée est d'introduire successivement chacune de ces variables dans la fonction discriminante en essayant, à chaque étape, de diminuer le plus possible la probabilité de classement erroné. Cela pose évidemment des problèmes de calcul assez importants et suppose la définition de procédures adéquates (MC CABE, 1975, et MC KAY, 1976 et 1977). Signalons cependant que cette procédure ne garantit pas l'obtention de la meilleure fonction discriminante (MURRAY, 1977) sans que beaucoup d'autres possibilités simples ne soient présentées.

Pour résoudre facilement ce problème de calcul nous avons utilisé un programme de régression multiple progressive, en présentant comme variables explicatives les variables susceptibles de composer la fonction discriminante et comme variable expliquée, une variable artificielle valant  $n_2/(n_1 + n_2)$  pour les individus de la première population et  $-n_1/(n_1 + n_2)$  pour les individus de la deuxième population.

Dans ces conditions, les coefficients de la fonction discriminante peuvent s'obtenir par l'expression :

$$\mathbf{b} = \frac{(n_1 + n_2) (n_1 + n_2 - 2)}{n_1 n_2 (1 - \mathbf{c}' (\bar{X}_1 - \bar{X}_2))} \mathbf{c}$$

où  $\mathbf{c}$  est le vecteur des coefficients de régression partielle, le coefficient  $b_0$  étant calculé comme ci-dessus en fonction des coefficients  $\mathbf{b}$ .

Il faut également signaler que le calcul de la probabilité de classement erroné est facilité par la relation (DAGNELIE, 1975, p. 310) :

$$D^2 = \frac{(n_1 + n_2) (n_1 + n_2 - 2)}{n_1 n_2} \frac{R^2}{(1 - R^2)}$$

où  $R^2$  est le coefficient de détermination.

La fonction discriminante définitive est ainsi facilement choisie parmi les équations de régression multiple en cherchant celle qui conduit à une probabilité de classement erroné acceptable avec un minimum de variables. Il convient également de tenir compte de la difficulté et de la fiabilité des mesures à réaliser dans le choix de cette fonction discriminante.

L'utilisation de la fonction discriminante pour le classement d'un individu dans l'une ou l'autre espèce peut être précédée par la réalisation d'un test confirmant que l'individu en question appartient bien à l'une ou l'autre des deux espèces et non à une troisième non identifiée (Mc DONALD *et al.*, 1976).

Signalons enfin que le fait de devoir estimer les paramètres des deux espèces (moyennes et matrices de variance et de covariance) conduit à une légère sous-estimation des risques de classement erroné (Mc LACHLAN, 1973 et 1975 ; MORAN, 1975, et MURRAY, 1977).

## 2. UTILISATION DE L'ANALYSE DISCRIMINANTE.

### 2.1. Matériel.

Les exemplaires utilisés (64 *lazera* et 21 *anguillaris*) ainsi que les mesures réalisées ont été décrits précédemment (MARQUET *et al.*, 1980 ; REIZER *et al.*, 1972). Cette étude avait permis de diviser les poissons du genre *Clarias* du Sénégal moyen en deux espèces suivant le rapport mentionné ci-dessus mais aucun des autres indices étudiés n'avait alors confirmé la validité de cette division.

On trouvera à la figure 1 la représentation de la plupart des mesures réalisées.

Le traitement manuel des mensurations étant le principal facteur limitant, nous avons pu augmenter le nombre des indices étudiés en utilisant l'ordinateur.

### 2.2. Résultats.

Trois mesures ont été choisies comme dénominateur afin d'éliminer l'effet de la taille. Il s'agit de la longueur standard pour son caractère universel, de la longueur de la tête et de l'espace inter-orbital pour la fiabilité de leur mesure (os crâniens incompressibles).

Les informations relatives aux rapports des mesures réalisées

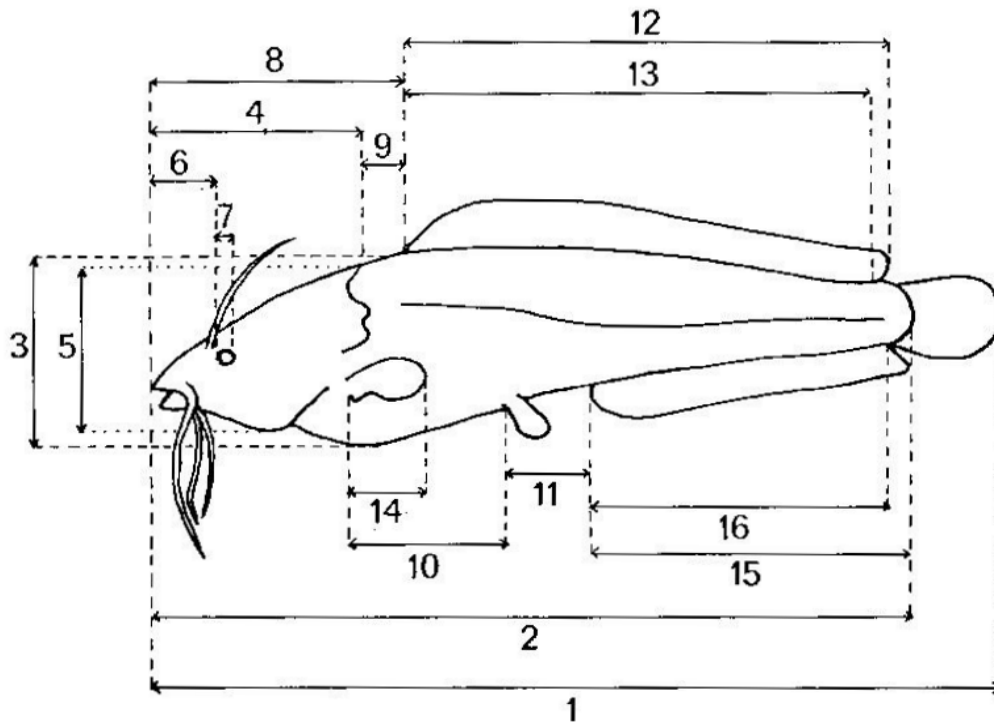


FIG. 1. — Mesures effectuées sur les *Clarias*.

- |                             |                                  |
|-----------------------------|----------------------------------|
| 1. Longueur totale.         | 9. Distance tête-dorsale.        |
| 2. Longueur standard.       | 10. Distance pectorale-ventrale. |
| 3. Hauteur du corps.        | 11. Distance ventrale-anale.     |
| 4. Longueur de la tête.     | 12. Dorsale, longueur maximum.   |
| 5. Hauteur de la tête.      | 13. Dorsale, longueur de base.   |
| 6. Longueur du museau.      | 14. Pectorale, longueur maximum. |
| 7. Diamètre de l'œil.       | 15. Anale, longueur maximum.     |
| 8. Distance museau-dorsale. | 16. Anale, longueur de base.     |

sur ces dénominateurs figurent respectivement aux tableaux I, II et III. Pour chaque mesure transformée on trouvera les moyennes observées des deux espèces, la statistique du test d'égalité de deux moyennes énoncé ci-dessus et, dans les cas de rejet d'égalité des moyennes des deux espèces, la probabilité de classement erroné correspondante. Rappelons que cette dernière correspond à l'utilisation du rapport concerné comme fonction discriminante.

A la lecture de ces tableaux on constate que les probabilités de classement erroné atteignent des niveaux fort élevés. Il convient donc de calculer une fonction discriminante faisant intervenir plusieurs rapports conformément à la méthodologie exposée au paragraphe 1.

Pour chacun des dénominateurs les rapports susceptibles d'entrer dans la fonction discriminante figurent au tableau IV. L'ordre de présentation de ces rapports correspond à l'ordre



TABLEAU I.

Moyennes observées des rapports à la longueur standard pour les *Clarias lazera* et les *C. anguillaris*, statistique  $u_{\text{obs}}$  du test de leur égalité et probabilité de classement erroné correspondante.

NUMÉRATEURS DES RAPPORTS A LA LONGUEUR STANDARD	MOYENNES OBSERVÉES		$u_{\text{obs}}$	PCE en %
	<i>lazera</i>	<i>anguillaris</i>		
Poids (g/mm).....	0,770	0,777	0,08	
Longueur totale.....	1,172	1,152	7,55	23
Hauteur du corps.....	0,145	0,139	1,63	
Longueur de la tête.....	0,293	0,285	2,32	36
Largeur de la tête.....	0,185	0,185	0,28	
Hauteur de la tête.....	0,112	0,109	1,35	
Longueur du museau.....	0,066	0,073	4,19	28
Diamètre de l'œil.....	0,026	0,026	1,00	
Espace inter-orbital.....	0,117	0,117	0,00	
Distance museau-dorsale.....	0,356	0,355	0,34	
Distance tête-dorsale.....	0,064	0,070	1,57	
Distance pectorale-ventrale.....	0,213	0,220	2,31	
Distance ventrale-anale.....	0,066	0,068	0,69	
Dorsale, longueur maximum.....	0,097	0,084	6,30	21
Dorsale, longueur de base.....	0,626	0,622	0,95	
Pectorale, longueur maximum.....	0,129	0,118	6,67	21
Pectorale, longueur de base.....	0,107	0,094	4,81	25
Ventrale, longueur maximum.....	0,101	0,096	2,19	28
Anale, longueur maximum.....	0,077	0,065	4,79	26
Anale, longueur de base.....	0,435	0,420	4,41	30
Longueur du barbillon nasal.....	0,162	0,141	3,38	31
Longueur du barbillon maxillaire.....	0,307	0,292	1,32	
Longueur du barb. mand. extérieur....	0,243	0,224	2,65	34
Longueur du barb. mand. intérieur....	0,177	0,164	2,33	36

TABLEAU II.

Moyennes observées des rapports à la longueur de la tête pour les *Clarias lazera* et les *C. anguillaris*, statistique  $u_{obs}$  du test de leur égalité et probabilité de classement erroné correspondante.

NUMÉRATEURS DES RAPPORTS A LA LONGUEUR DE LA TÊTE	MOYENNES OBSERVÉES		$u_{obs}$	PCE en %
	<i>lazera</i>	<i>anguillaris</i>		
Poids (g/mm).....	2,620	2,790	0,19	
Longueur totale.....	4,004	4,051	0,91	
Longueur standard.....	3,416	3,517	2,25	36
Hauteur du corps.....	0,494	0,489	0,38	
Largeur de la tête.....	0,632	0,649	1,61	
Hauteur de la tête.....	0,384	0,384	0,06	
Longueur du museau.....	0,224	0,255	6,33	18
Diamètre de l'œil.....	0,090	0,090	0,05	
Espace inter-orbital.....	0,401	0,413	1,97	40
Distance museau-dorsale.....	1,214	1,247	1,89	
Distance tête-dorsale.....	0,218	0,248	1,74	
Distance pectorale-ventrale.....	0,728	0,773	2,75	35
Distance ventrale-anale.....	0,227	0,238	1,61	
Dorsale, longueur maximum.....	0,332	0,296	6,75	25
Dorsale, longueur de base.....	2,141	2,188	1,31	
Pectorale, longueur maximum.....	0,441	0,413	3,88	30
Pectorale, longueur de base.....	0,367	0,331	4,16	28
Ventrale, longueur maximum.....	0,344	0,338	0,70	
Anale, longueur maximum.....	0,264	0,229	4,42	30
Anale, longueur de base.....	1,486	1,479	0,31	
Longueur du barbillon nasal.....	0,552	0,493	3,09	34
Longueur du barbillon maxillaire.....	1,049	1,028	0,63	
Longueur du barb. mand. extérieur....	0,831	0,783	2,09	38
Longueur du barb. mand. intérieur....	0,604	0,574	1,65	

TABLEAU III.

Moyennes observées des rapports à l'espace inter-orbital pour les *Clarias lazera* et les *C. anguillaris*, statistique  $u_{obs}$  du test de leur égalité et probabilité de classement erroné correspondante.

NUMÉRATEURS DES RAPPORTS A L'ESPACE INTER-ORBITAL	MOYENNES OBSERVÉES		$u_{obs}$	PCE en %
	<i>lazera</i>	<i>anguillaris</i>		
Poids (g/mm).....	6,56	6,62	0,06	
Longueur totale.....	10,02	9,82	1,80	
Longueur standard.....	8,55	8,53	0,21	
Hauteur du corps.....	1,23	1,18	1,89	
Longueur de la tête.....	2,50	2,43	2,08	41
Largeur de la tête.....	1,58	1,57	0,47	
Hauteur de la tête.....	0,96	0,93	1,31	
Longueur du museau.....	0,56	0,62	4,49	30
Diamètre de l'œil.....	0,23	0,22	0,59	
Distance museau-dorsale.....	3,04	3,02	0,48	
Distance tête-dorsale.....	0,55	0,60	1,49	
Distance pectorale-ventrale.....	1,82	1,87	1,64	
Distance ventrale-anale.....	0,57	0,58	0,55	
Dorsale, longueur maximum.....	0,83	0,72	6,43	26
Dorsale, longueur de base.....	5,36	5,30	0,71	
Pectorale, longueur maximum.....	1,10	1,00	5,89	27
Pectorale, longueur de base.....	0,92	0,80	4,94	26
Ventrale, longueur maximum.....	0,86	0,82	2,45	38
Anale, longueur maximum.....	0,66	0,56	4,44	29
Anale, longueur de base.....	3,72	3,59	2,55	40
Longueur du barbillon nasal.....	1,38	1,20	3,39	32
Longueur du barbillon maxillaire.....	2,62	2,50	1,32	
Longueur du barb. mand. extérieur....	2,08	1,91	2,70	34
Longueur du barb. mand. intérieur....	1,51	1,40	2,24	37

TABLEAU IV.

Probabilités de classement erroné associées à l'introduction successive dans la fonction discriminante de huit mesures rapportées à trois grandeurs différentes.

VARIABLE AJOUTÉE : rapport à la longueur standard de	PCE en %	VARIABLE AJOUTÉE : rapport à la longueur de la tête de	PCE en %	VARIABLE AJOUTÉE : rapport à l'espace inter-orbital de	PCE en %
pectorale, longueur maximum.	21	longueur du museau.....	18	dorsale, longueur maximum...	26
longueur du museau.....	16	dorsale, longueur maximum...	15	longueur du museau.....	16
longueur de la tête.....	14	diamètre de l'œil.....	13	poids.....	14
anale, longueur maximum....	11	pectorale, longueur de base...	12	longueur barb. mand. ext.....	12
diamètre de l'œil.....	10	longueur standard.....	11	distance ventrale-anale.....	10
longueur du barb. nasal.....	8	anale, longueur de base.....	9	anale, longueur maximum.....	8
anale, longueur de base.....	7	poids.....	8	longueur standard.....	8
poids.....	6	longueur totale.....	7	longueur de la tête.....	6

d'entrée dans la fonction discriminante défini par le processus progressif de régression multiple. Les rapports sont ainsi approximativement présentés dans l'ordre décroissant de leur pouvoir discriminant. La probabilité de classement erroné figurant en regard d'un rapport correspond au résultat de l'adjonction, dans la fonction discriminante, de ce rapport à ceux qui le précèdent.

### 2.3. Discussions.

La procédure de calcul progressif de la fonction discriminante qui a été utilisée présente les mesures dans un ordre différent suivant le dénominateur choisi. Cela nous amène à considérer certains rapports de valeurs moins fiables dans le cas de l'espace inter-orbital. Par exemple, la longueur des barbillons peut être altérée par une éventuelle repousse après mutilation.

En terme de probabilité de classement erroné, on constate également que les résultats fournis par les trois dénominateurs sont sensiblement comparables. En conséquence, pour son caractère plus universel, nous choisirons la longueur standard comme dénominateur c'est-à-dire comme indice de la taille.

De plus, il nous paraît justifié de nous limiter à une fonction discriminante de cinq variables. En effet, après celles-ci, la diminution de PCE ne justifie plus l'introduction de variables supplémentaires qui risquent de compromettre l'intérêt pratique de la fonction discriminante. Dans ces conditions cette dernière s'écrit :

$$y = 189 x_1 - 321 x_2 + 144 x_3 + 116 x_4 - 174 x_5 - 38,644$$

où  $x_1$  : longueur maximum de la nageoire pectorale divisée par la longueur standard,

$x_2$  : longueur du museau divisée par la longueur standard,

$x_3$  : longueur de la tête divisée par la longueur standard,

$x_4$  : longueur maximum de la nageoire anale divisée par la longueur standard,

$x_5$  : diamètre de l'œil divisé par la longueur standard.

Une valeur positive de  $y$  traduit l'appartenance de l'individu concerné à l'espèce *Clarias lazera* ( $y$  négatif correspond à *Clarias anguillaris*). Rappelons que la probabilité de classement erroné correspondante est de un poisson sur dix.

Ce résultat n'est guère satisfaisant et justifie la poursuite de nos investigations.

4. TEST VISUEL.

3.1. Matériel.

Les particularités morphologiques des poissons diffèrent fortement d'un genre à l'autre. C'est la raison pour laquelle nous avons pensé qu'il serait préférable de choisir les mensurations à effectuer après examen des exemplaires à différencier plutôt que d'adopter un système de mensuration automatique.

Ce test visuel consistait à prendre tous les exemplaires disponibles des *Clarias* en provenance du bassin du Sénégal, à les répartir en *Clarias anguillaris* (60 individus) et *Clarias lazera* (60 individus également) sur base du rapport nombre de branchiospines / longueur standard, à les associer par paire de taille croissante (un exemplaire de chaque espèce et de taille approximativement égale dans chaque paire) et enfin à comparer taille par taille les deux espèces.

Après un tel examen comparé de chaque partie du corps, nous avons retenu les mensurations suivantes :

- longueur et largeur de la fontanelle frontale,
- largeur et hauteur du processus occipital,
- distance dorsale-caudale.

Les mensurations relatives à la tête sont schématisées sur la figure 2.

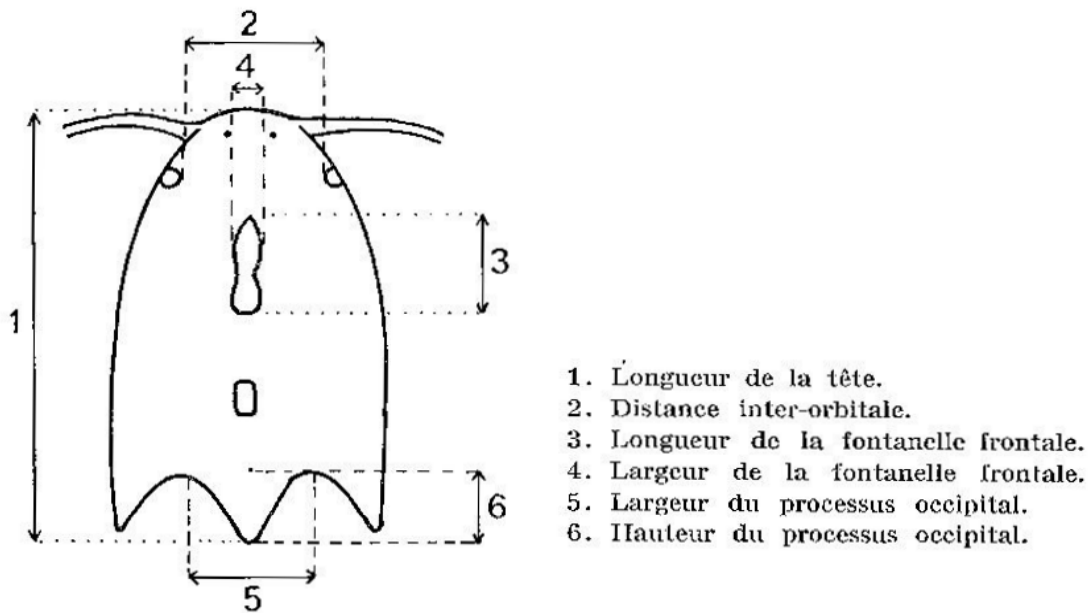


FIG. 2. --- Mesures effectuées sur la tête des *Clarias*.

Nous avons également mesuré la longueur standard et la longueur de la tête pour servir de références aux différentes mensurations.

### 3.2. Résultats.

Quatre mesures ont été choisies comme dénominateur afin d'éliminer l'effet de la taille. Il s'agit de la longueur standard, de la longueur de la tête, de la longueur de la fontanelle frontale et de la largeur du processus occipital. Les informations relatives à ces quatre séries de rapports sont présentées dans le tableau V d'une manière analogue à celle des tableaux I, II et III.

### 3.3. Discussion.

A la lecture de ce tableau on remarque de suite une probabilité de classement erroné très faible pour le rapport hauteur du processus occipital sur largeur de ce processus et, dans une moindre mesure, pour la distance dorsale-caudale sur la largeur du processus occipital.

De plus, à une exception près, les moyennes de ces rapports sont toutes significativement différentes d'une espèce à l'autre. Cette deuxième méthode semble donc beaucoup plus efficiente au niveau du pouvoir discriminant que la précédente.

Il faut également signaler qu'une fonction discriminante réunissant la hauteur du processus occipital divisée par sa largeur et le rapport distance dorsale-caudale sur la largeur du processus ne diminue pas sensiblement la probabilité de classement erroné par rapport à celle déduite du seul rapport hauteur sur largeur du processus occipital (1,1 % contre 1,3 %). Dans ces conditions nous retenons la fonction discriminante suivante :

$$y = 16,1 x_1 - 22,8$$

où  $x_1$  est le rapport de la hauteur du processus occipital sur sa largeur et 16,1 et 22,8 sont les coefficients calculés selon les formules exposées ci-dessus.

Une valeur positive de  $y$  classe l'individu testé dans l'espèce *lazera* et, inversement, une valeur négative de  $y$  le classe dans l'espèce *anguillaris*. Le risque d'erreur associé à cette méthode est particulièrement bas puisqu'il est compris entre un et deux poissons sur cent.

TABLEAU V.

Moyennes observées pour les *Clarias lazera* et les *C. anguillaris* de différents rapports, statistique  $u_{obs}$  du test de leur égalité et probabilité de classement erroné correspondante.

NUMÉRATEURS DES RAPPORTS	MOYENNES OBSERVÉES		$u_{obs}$	PCE en %
	<i>lazera</i>	<i>anguillaris</i>		
Numérateurs des rapports à la longueur standard				
Longueur de la tête.....	0,320	0,314	2,33	42
Longueur de la fontanelle frontale.....	0,072	0,082	5,41	31
Largeur de la fontanelle frontale.....	0,024	0,022	2,59	41
Largeur du processus occipital.....	0,066	0,061	4,11	35
Hauteur du processus occipital.....	0,047	0,049	1,75	
Distance dorsale-caudale.....	0,029	0,035	4,37	34
Numérateurs des rapports à la longueur de la tête				
Longueur de la fontanelle frontale.....	0,226	0,261	7,00	26
Largeur de la fontanelle frontale.....	0,075	0,070	2,03	43
Largeur du processus occipital.....	0,206	0,196	3,19	40
Hauteur du processus occipital.....	0,148	0,156	2,96	39
Distance dorsale-caudale.....	0,090	0,110	4,46	34
Numérateurs des rapports à la longueur de la fontanelle frontale				
Largeur de la fontanelle frontale.....	0,332	0,272	6,59	27
Largeur du processus occipital.....	0,924	0,758	7,13	26
Hauteur du processus occipital.....	0,664	0,608	3,07	39
Numérateurs des rapports à la largeur du processus occipital				
Hauteur du processus occipital.....	2,023	0,806	24,35	1
Distance dorsale-caudale.....	1,225	0,575	12,25	12



## 4. GÉNÉRALISATION ET CONCLUSION.

Si l'on compare les deux méthodes de mensuration suivies et les résultats correspondants, on constate qu'il est plus intéressant de chercher des mesures adaptées aux poissons étudiés plutôt que d'utiliser un schéma uniforme de mensurations.

En effet, ce dernier, plus facile à étendre et à vulgariser, ne permet cependant pas une utilisation rationnelle des mesures en vue d'un classement objectif.

Nous disposons donc maintenant d'une technique de mesures rapide, simple et efficace pour différencier *Clarias lazera* de *Clarias anguillaris*. De plus, contrairement au système de classification précédent, il n'est plus nécessaire de sacrifier l'animal pour le déterminer. Nous avons donc atteint le but que nous nous étions fixé au départ.

Il convient maintenant de proposer un ensemble de règles à suivre pour généraliser la méthode de manière optimale aux différents genres de poissons dont la détermination présente de grandes difficultés (notamment les poissons non écaillés).

Précisons à nouveau que la méthode n'est avantageusement appliquée que pour des espèces exclusivement déterminables *post mortem*; le ou les indices retenus pour une détermination sur le terrain seront choisis sur base d'un échantillon suffisant, préalablement déterminé en laboratoire.

Une fois déterminés, les espèces seront comparées deux à deux, taille par taille, et on recherchera des différences morphologiques à mesurer. Cette opération est essentielle, car de la qualité des observations dépend le succès de cette méthode. En pratique, toute proportion différente d'un animal à l'autre doit être reprise dans les mesures.

On prendra les mesures sélectionnées sur un échantillon suffisant et le plus représentatif possible dans chaque espèce selon les possibilités (50 à 100 exemplaires de toutes tailles par espèce). On éliminera le facteur taille associé à ces mesures en utilisant un ou des dénominateurs choisis parmi les mesures les plus fiables, les mieux corrélées à la taille et les plus utilisées dans la littérature.

La comparaison des moyennes de ces rapports par un test d'égalité des moyennes indiquera les indices susceptibles d'avoir une valeur discriminante.

Pour chaque rapport dont la différence des moyennes est signi-

ficative, on calculera le risque d'erreur associé à une détermination basée sur la seule valeur du rapport étudié.

On choisira alors le ou les rapports permettant une détermination comportant un risque d'erreur minimum et on comparera ce risque à ceux calculés pour une combinaison linéaire de plusieurs de ces rapports.

Le choix final dépendra du niveau d'erreur atteint par indice et du nombre de mesures nécessaire à l'établissement de la fonction discriminante.

On calculera enfin la fonction discriminante et on vérifiera son efficacité sur d'autres échantillons de mêmes espèces.

#### REMERCIEMENTS.

Les auteurs remercient le Professeur C. REIZER pour le matériel mis à leur disposition, ainsi que le Professeur D. THYS VAN DEN AUDENAERDE, ses assistants et son personnel, pour l'hospitalité et l'aide qu'ils leur ont dispensées lors de la constitution du matériel à analyser.

#### BIBLIOGRAPHIE

- COCHRAN, W. G. (1964). — On the performance of the linear discriminant function. *Technometrics*, **6**, p. 179-190.
- DAGET, J. (1954). — Les poissons du Niger Supérieur. *Dakar*, Mém. IFAN, n° 36, p. 1-391.
- DAGNELIE, P. (1970). — Théorie et Méthode statistiques, t. 2. *Gembloux*, Presses Agronomiques, 451 p.
- (1975). — Analyse statistique à plusieurs variables. *Gembloux*, Presses Agronomiques, 362 p.
- HÉBRANT, F. (1973). — Problèmes de discrimination dans le cas de deux populations. *Biométrie-Praximétrie*, **13**, p. 15-32.
- HUISSE, E. DE (1974). — Bijdrage tot een morfologisch-systematische studie van *Clarias* vissen uit west of centraal Afrika (Mémoire). *Gent*, Rijks Universiteit, p. 1-122.
- LUBISHEN, A. A. (1962). — On the use of discriminant function in taxonomy. *Biometrics*, **18**, p. 455-477.
- MARQUET, J.-P., REIZER, C. & MATTEI, X. (1980). — Contribution à l'étude de la faune ichtyologique du bassin du fleuve Sénégal. IV. *Clariidae*. *Bull. IFAN*, **41**, A, n° 3, p. 596-626.
- MC CABE, G. P. (1975). — Computations for variable selection in discriminant analysis. *Technometrics*, **17** (1), 103-109.

- Mc DONALD, L. L., LOWE, V. W., SMIDT, R. K. & MEISTER, K. A. (1976). — A preliminary test for discriminant analysis based on small samples. *Biometrics*, **32**, p. 417-422.
- Mc KAY, R. J. (1976). — Simultaneous procedures in discriminant analysis involving two groups. *Technometrics*, **18** (1), 47-53.
- (1977). — Simultaneous procedures for variable selection in multiple discriminant analysis. *Biometrika*, **64** (2), 283-290.
- Mc LACHLAN, G. J. (1973). — An asymptotic expansion of the expectation of the estimated error rate in discriminant analysis. *Austral. J. Statist.*, **15** (3), 210-214.
- (1975). — Confidence intervals for the conditional probability of misallocation in discriminant analysis. *Biometrics*, **31** (1), p. 161-167.
- MORAN, M. A. (1975). — On the expectation of errors of allocation associated with a linear discriminant function. *Biometrika*, **62** (1), p. 141-148.
- MURRAY, G. D. (1977). — A cautionary note on selection of variables in discriminant analysis. *Appl. Statist.*, **26** (3), p. 246-250.
- REIZER, C., MATTEI, X. & CHEVALIER, J.-L. (1972). — Contribution à l'étude de la faune ichthyologique du bassin du fleuve Sénégal. I. *Polypteridae*. *Bull. IFAN*, XXXIV, A, n° 1, p. 111-126.
- VIBERT, R. & LAGLER, K. F. (1961). — Pêches continentales. Paris, Dunod, p. 56-69.
-