



**HAL**  
open science

**UN LOGICIEL DE TRAITEMENT DES  
HISTOGRAMMES DE STRUCTURE DE  
POPULATION PAR LA MÉTHODE DES  
DIFFÉRENCES LOGARITHMIQUES DE  
BHATTACHARYA** A program for analysis of  
population structure histograms using Bhattacharya's  
method of logarithmic differences

M Lepetit, S Loranchet, P Gillet, J M Marion

► **To cite this version:**

M Lepetit, S Loranchet, P Gillet, J M Marion. UN LOGICIEL DE TRAITEMENT DES HISTOGRAMMES DE STRUCTURE DE POPULATION PAR LA MÉTHODE DES DIFFÉRENCES LOGARITHMIQUES DE BHATTACHARYA A program for analysis of population structure histograms using Bhattacharya's method of logarithmic differences. *Vie et Milieu / Life & Environment*, 1991, pp.127-131. hal-03039505

**HAL Id: hal-03039505**

**<https://hal.sorbonne-universite.fr/hal-03039505>**

Submitted on 3 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# UN LOGICIEL DE TRAITEMENT DES HISTOGRAMMES DE STRUCTURE DE POPULATION PAR LA MÉTHODE DES DIFFÉRENCES LOGARITHMIQUES DE BHATTACHARYA

*A program for analysis of population structure histograms using Bhattacharya's method of logarithmic differences*

M. LEPETIT <sup>(1)</sup>, S. LORANCHET <sup>(1)</sup>, P. GILLET <sup>(2)</sup>, J.M. MARION <sup>(1)</sup>

<sup>(1)</sup> Institut de Mathématiques Appliquées, 3 place André Leroy, 49008 Angers Cédex 01, France

<sup>(2)</sup> Laboratoire d'Ecologie Animale, 3 place André Leroy, 49008 Angers Cédex 01, France

DÉMÉCOLOGIE  
DYNAMIQUE DES POPULATIONS  
ANALYSE DE DONNÉES  
HISTOGRAMMES

**RÉSUMÉ** – Le logiciel « BHATTA », réalisé en Turbo Pascal sur IBM PC, permet le traitement des histogrammes de structure de population. La base de l'étude est l'analyse de la dynamique des populations à partir d'histogrammes de fréquences de taille. Le logiciel « BHATTA », élaboré à partir de la méthode des différences logarithmiques de Bhattacharya, permet un découpage des histogrammes afin d'isoler les composantes gaussiennes et d'en estimer les paramètres : effectif de l'échantillon, moyenne et écart-type de chaque classe d'âge ou cohorte pour établir les courbes de croissance de chaque cohorte.

DEMECOLOGY  
POPULATION DYNAMICS  
DATA ANALYSIS  
HISTOGRAMS

**ABSTRACT** – The software « BHATTA », achieved in Turbo Pascal on IBM PC makes possible histogram analysis for population structure. The base of the study is an analysis of population dynamics from histograms of size frequency. The program « BHATTA », made from the method of logarithmic differences of Bhattacharya, cuts the histograms in order to display gaussian components and to estimate the parameters : number (size of sample), mean and standard deviation for each generation or cohort to determine growth curves for each cohort.

## INTRODUCTION

La plupart des études démographiques ou de dynamique des populations se font à partir d'histogrammes de fréquence représentant la variation d'un critère biométrique tel que la taille, le poids, etc... (Barbault, 1981; Grant *et al.* 1987).

Pour séparer les différentes composantes gaussiennes d'une population, c'est-à-dire les différentes générations ou cohortes qui la composent, il est nécessaire d'utiliser une méthode mathématique. Cette méthode peut être graphique (Harding, 1949; Cassié, 1954) mais longue et fastidieuse ou informatisée et donc beaucoup plus performante tel que Elefan (Brey *et al.*, 1988).

La méthode de Bhattacharya (1967) avait déjà fait l'objet d'un programme informatisé, le programme NORMSEP d'optimisation par le maximum de vraisemblance (Hasselblad, 1966) en

Fortran IV sur Hewett Packard 9825 A (Tomlison, 1970). Ce programme avait été utilisé par Laurent et Lassere (1974) pour la détermination des classes d'âge de Daurades *Sparus aurata*, et par Gros et Cochard (1978) dans l'étude biologique du Crustacé *Nyctiphanes couchii*.

Le logiciel « BHATTA », du nom de Bhattacharya (1967), auteur d'une méthode des différences logarithmiques pour séparer une distribution en composantes gaussiennes, a été réalisé en Turbo Pascal avec les utilitaires Grafix-Toolbox sur micro-ordinateur IBM PC. Ce logiciel permet le découpage des histogrammes de fréquences de taille afin d'isoler les différentes composantes et d'estimer les paramètres : effectif de l'échantillon, moyenne et écart type de chaque groupe ou cohorte en vue d'établir les courbes de croissance de chaque cohorte.

Le logiciel « BHATTA », plus convivial, grâce à une série de menus, permet la coloration des dif-



férentes cohortes et offre une plus large utilisation dans les études de dynamique des populations grâce à une compatibilité plus importante.

## MÉTHODE

La méthode des différences logarithmiques de Bhattacharya (1967), basée sur une résolution graphique, permet de séparer les composantes gaussiennes d'une population. La distribution des fréquences d'un caractère biométrique est supposée suivre une loi de Laplace-Gauss. On isole ainsi un certain nombre d'entités statistiques et/ou biologiques que l'on appelle sous-populations, classes d'âge, générations ou cohortes. La méthode permet de déterminer les paramètres moyenne et écart-type des composantes gaussiennes et calcule l'effectif théorique de chacune des composantes. Les fréquences résiduelles correspondent à des effectifs non pris en compte dans la sélection des droites, c'est-à-dire, des points non alignés qui correspondent, soit à des chevauchements de courbes, soit à des erreurs d'échantillonnage. La méthode se divise en deux parties : d'une part, la détermination du nombre  $k$  de sous-populations normales contenues dans la distribution, puis la détermination pour chacune d'elles, des paramètres : effectif, moyenne et écart type.

On note :

$y(x)$ , l'effectif observé pour la classe de point médian  $x$

$h$ , la longueur de l'intervalle de la classe

On trace alors le graphique des points de coordonnées  $(x, \delta \log y)$  avec  $\delta \log y = \log y(x+h) - \log y(x)$ , et on recherche les régions où le graphique ressemble à une droite à pente négative. Sous certaines hypothèses, le nombre de telles régions est le nombre des composantes. Si l'on note  $r$  le point d'intersection de la  $r^{\text{ième}}$  droite avec l'axe des abscisses,  $\theta_r$  l'angle de la  $r^{\text{ième}}$  droite avec l'axe des abscisses,  $b$  et  $d$  les échelles respectives pour  $x$  et  $\delta \log y$ , alors les paramètres gaussiens (moyenne et écart-type) pour la  $r^{\text{ième}}$  composante sont respectivement donnés par les formules suivantes :

$$\mu_r = \alpha r + h/2$$

$$\sigma^2 = (d^*h^* \cot \theta_r / b - (h^2/12))$$

On suppose que la fonction des fréquences de taille est un mélange de  $k$  distributions gaussiennes de paramètres  $(N_i, \mu_i, \theta_i)$   $i = 1, \dots, k$  et on se place dans l'hypothèse de distributions de composantes suffisamment séparées pour qu'il existe, pour chaque composante, un voisinage suffisant dans lequel l'influence des autres composantes est relativement négligeable.

## RÉSULTATS

Les résultats fournis par le logiciel « BHATTA » se présentent sous trois formes permettant une sortie graphique. Tout d'abord, le menu permet d'afficher l'histogramme (fig. 1a) où figure l'effectif de chaque classe et l'effectif de l'échantillon de la population. Dans un 2<sup>e</sup> temps, le logiciel affiche le graphique des différences logarithmiques (fig. 1b). A chaque segment coupant l'axe des abscisses correspond un groupe (cohorte). Il suffit alors d'entrer les points correspondants à chacune des régions ainsi mises en évidence. Le logiciel offre différentes options, en particulier, le traitement des fréquences résiduelles (cas des points non alignés). Enfin, le logiciel permet d'obtenir l'histogramme découpé (fig. 1c), où les cohortes

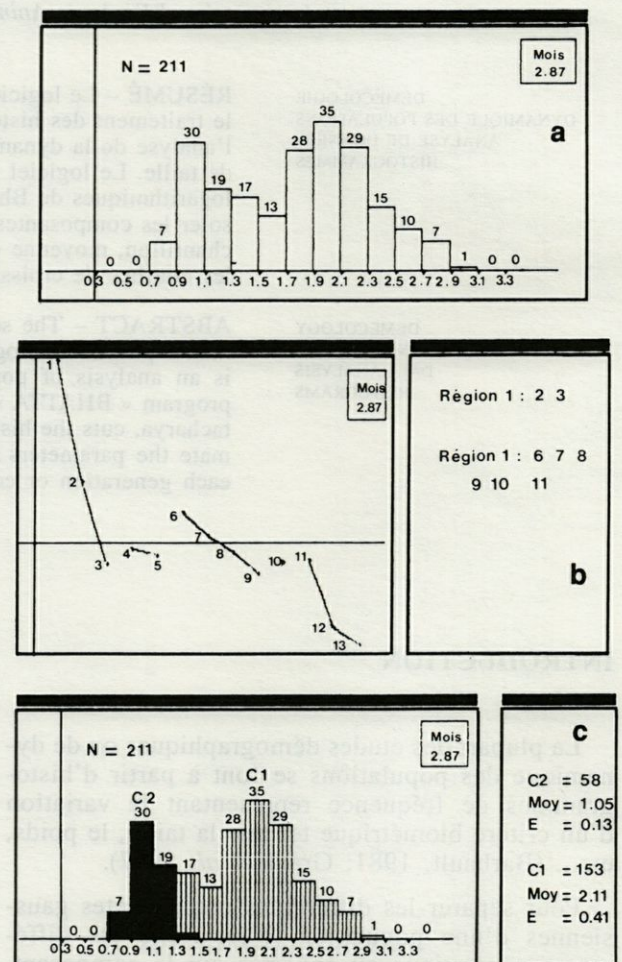


Fig. 1. — Structure d'une population de *Nereis diversicolor* (Annélide Polychète) de l'estuaire de la Loire (1987). a, histogramme des fréquences de taille; b, graphique des différences logarithmiques; c, histogramme découpé et paramètres.

Structure of a population of *Nereis diversicolor* (Annelida Polychaeta) from Loire estuary (1987). a, histogram of size frequency; b, graphic of logarithmic differences; c, cut histogram and parameters.



isolées sont coloriées de façon différente. Les paramètres effectif, moyenne et écart type sont affichés dans une fenêtre à droite de l'histogramme découpé.

Le premier exemple (fig. 1) est l'échantillon d'une population de *Nereis diversicolor* (Annélide Polychète) de l'estuaire de la Loire de février 1987 (Gillet, 1990). L'histogramme présente 2 classes modales de, respectivement 30 et 35 individus. Le graphique des différences logarithmiques met en évidence 2 segments : les segments formés des points 1, 2, 3 et 6, 7, 8, 9 qui coupent l'axe des abscisses. L'utilisateur retiendra donc l'hypothèse de 2 régions. Pour la première, on choisira les points 2 et 3, excluant le point 1 afin de ne pas avoir un effectif trop important surestimant la cohorte C2 par rapport à la cohorte C1. A l'inverse, en plus des points 6, 7, 8, 9, on traitera les fréquences résiduelles en tenant compte des points 10 et 11 afin que l'effectif de la cohorte C1 soit suffisant. On peut noter d'ailleurs qu'en tenant compte des points 10 et 11 on obtient un segment avec les points 12 et 13; tous ces points faisant partie de la même région. L'histogramme découpé est alors tracé avec la cohorte C1 en hachures et la cohorte C2 en noir. Les paramètres sont affichés dans une fenêtre. Les effectifs sont de 58 pour le groupe 1 (C2) et 153 pour le groupe 2 (C1) soit un effectif total de 211 individus. La taille moyenne est de 1.05 mm avec un écart type de 0.13 pour C2 et 2.11 mm avec un écart type de 0.41 pour C1.

**DISCUSSION**

Il nous a paru intéressant de tester le logiciel à partir d'un histogramme correspondant à des fréquences calculées à partir de la densité de la loi normale centrée réduite. Ceci afin d'examiner le problème de l'importance des effectifs, puis de tester la validité des résultats obtenus par le traitement des fréquences résiduelles. Nous avons ainsi construit 2 histogrammes avec 2 précisions différentes. Les valeurs des fréquences calculées pour un effectif de 70 étaient les suivantes : 20,2 - 11 - 3,3 - 0,5 - 0,5 - 3,3 - 11 - 20,2

On a alors construit les 2 histogrammes à partir des fréquences suivantes : 1 - 3 - 11 - 20 - 20 - 11 - 3 - 1 et 5 - 33 - 110 - 202 - 202 - 110 - 33 - 5

Les résultats obtenus (fig. 2) sont plus corrects lorsque la précision est plus grande; en effet, la droite est parfaite pour le second histogramme, alors que pour le premier histogramme les points 1 et 7 en sont exclus (précision de 1/70 pour le cas 1 et 1/700 pour le cas 2).

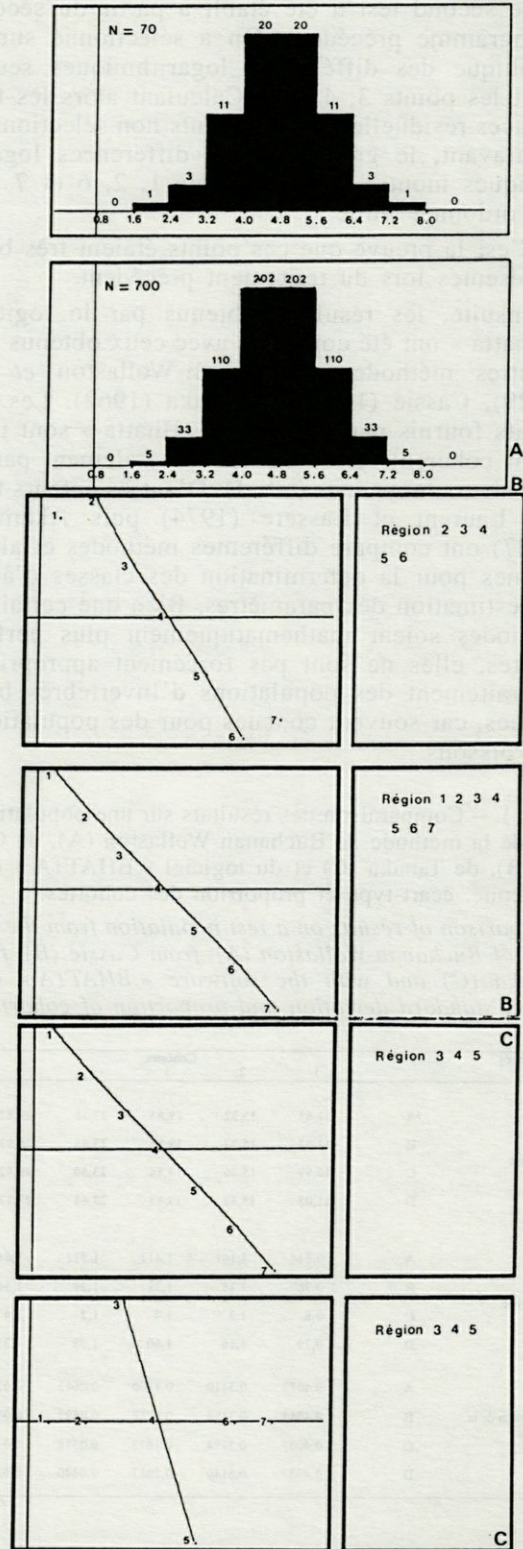


Fig. 2. - Histogrammes des fréquences de taille et graphiques des différences logarithmiques d'une même population avec des effectifs de 70 à 700 individus.

*Histograms of size frequency and graphics of logarithmic differences for the same population with total number from 70 to 700 individuals.*



Le second test a été établi à partir du second histogramme précédent. On a sélectionné sur le graphique des différences logarithmiques seulement les points 3, 4 et 5. Calculant alors les fréquences résiduelles sur les points non sélectionnés auparavant, le graphique des différences logarithmiques montre que les points 1, 2, 6 et 7 ont une ordonnée nulle.

C'est la preuve que ces points étaient très bien représentés lors du traitement précédent.

Ensuite, les résultats obtenus par le logiciel « Bhatta » ont été comparés avec ceux obtenus par d'autres méthodes : Buchanan-Wollaston *et al.* (1929), Cassié (1954) et Tanaka (1962). Les résultats fournis par le logiciel « Bhatta » sont tout à fait cohérents et se situent généralement parmi les valeurs moyennes (tab. I). D'autres auteurs tels que Laurent et Lassere (1974) puis Akamine (1987) ont comparé différentes méthodes et algorithmes pour la détermination des classes d'âges et l'estimation des paramètres. Bien que certaines méthodes soient mathématiquement plus performantes, elles ne sont pas forcément appropriées au traitement des populations d'Invertébrés benthiques, car souvent conçues pour des populations de Poissons.

Tabl. I. – Comparaison des résultats sur une population-test de la méthode de Buchanan-Wollaston (A), de Cassié (B), de Tanaka (C) et du logiciel « BHATTA » (D). Moyenne, écart-type et proportion des cohortes.

*Comparison of results on a test population from the method of Buchanan-Wollaston (A), from Cassié (B), from Tanaka (C) and with the software « BHATTA » (D). Mean, standard deviation and proportion of cohorts.*

Paramètres		Cohortes				
		1	2	3	4	5
Moyenne	A	11,05	15,32	19,85	23,58	26,82
	B	11,02	15,33	19,85	23,46	26,92
	C	10,99	15,26	19,84	23,50	26,82
	D	11,05	15,32	19,88	23,64	27,12
Ecart-type	A	0,844	1,161	1,412	1,212	1,443
	B	0,76	1,15	1,32	1,29	1,54
	C	0,8	1,2	1,4	1,2	1,4
	D	0,79	1,16	1,60	1,09	1,23
Proportion de la cohorte	A	0,4072	0,3110	0,1860	0,0642	0,0316
	B	0,4049	0,3164	0,1788	0,0693	0,0307
	C	0,4007	0,3194	0,1873	0,0598	0,0328
	D	0,4037	0,3140	0,2087	0,0440	0,0290

## CONCLUSION

Le logiciel « Bhatta » réalisé en Turbo Pascal sur IBM PC à partir de la méthode des différences logarithmiques de Bhattacharya (1967) permet le traitement des histogrammes de fréquence utilisés

dans les études de dynamique des populations. Les utilisations sont nombreuses et les exemples sur la dynamique des populations d'Annélides Polychètes peuvent être étendus à d'autres groupes zoologiques : Mollusques, Crustacés... mais également aux Vertébrés ou à toute étude de structure démographique utilisant des critères biométriques : taille, poids...

L'utilisation est facilitée grâce à une série de menus et au coloriage qui augmentent la convivialité du logiciel et le rend accessible à un utilisateur n'ayant aucune connaissance particulière en informatique.

## BIBLIOGRAPHIE

- AKAMINE T., 1987. Comparison of algorithms of several methods for estimating parameters of a mixture of normal distributions. *Bull. Jpn Sea Reg. Fish. Res. Lab.* **37** : 259-277.
- BARBAULT R., 1981. Ecologie des populations et des peuplements. Masson, Paris, 200 p.
- BHATTACHARYA, C.G., 1967. A simple method of resolution of a distribution into Gaussian components. *Biometrics* **23** (1) : 115-135.
- BREY T., M. SORIANO & D. PAULY, 1988. Electronic length frequency analysis. A revised and expanded user's guide to elefan 0,1 and 2. *Ber. Inst. Meeresk.* **177** : 31 p.
- BUCHANAN-WOLLASTON H.G. & W.C. HODGSON, 1929. A new method of treating frequency curves in fishery statistics with some results. *J. Cons.* **4** : 207-225.
- CASSIE R.M., 1984. Some uses of probability paper for the graphical analysis of polymodal frequency distributions. *Aust. J. Mar. Freshw. Res.* **5** : 513-522.
- GILLET P., 1990. Biomasse, production et dynamique des populations de *Nereis diversicolor* (Annélide polychète) de l'estuaire de la Loire. *Oceanologica Acta* **13** (3) : 361-371.
- GRANT A., P.J. MORGAN & P.J.W. OLIVE, 1987. Use made in marine ecology of methods for estimating demographic parameters from size/frequency data. *Mar. Biol.* **95** (2) : 201-208.
- GROS P. & J.C. COCHARD, 1978. Biologie de *Nyctiphanes couchii* dans le secteur Nord du Golfe de Gascogne. *Ann. Inst. Oceanogr. Paris* **54** (1) : 25-46.
- HARDING J.F., 1949. The use of probability paper for graphical analysis of polymodal frequency distributions. *J. Mar. Biol. Ass. U.K.* **28** : 141-153.
- HASSELBLAD V., 1966. Estimation of parameters for a mixture of normal distributions. *Technometrics* **8** (3) : 431-446.
- LAURENT M. & G. LASSERRE, 1974. Détermination des classes d'âge dans trois échantillons de Daurade (*Sparus aurata* L.). Comparaison de la méthode sca-



limétrique et de la méthode des différences logarithmiques. *Ann. Hydrobiol.* 5 (1) : 73-76.

TANAKA S., 1962. A method of analysing of polymodal frequency distribution and its application to the length distribution of Porgy, *Tarus tumifrons* (J. and S.). *J. Fish. Res. Bd. Can.* 19 (6) : 1143-1159.

TOMLISON P.K., 1970. Program NORMSEP - Normal distribution separator 11 (1) 2. 1 to 11 (1) 2. 10 in Abramson N.J. F.A.O. Fish Tech. Pap., 101 p.

Reçu le 17 janvier 1990; received January 17, 1990  
Accepté le 25 octobre 1990; accepted October 25, 1990