

## UTILISATION ET INTERPRETATION DU MODELE DE GOMPERTZ

### APPLICATION A L'ETUDE DE LA CROISSANCE DE JEUNES RATS MUSQUES (*Ondatra zibethica* L.)

A. PAVE (1), A. CORMAN (1), B. BOBILLIER-MONOT (2)

(1) Laboratoire de Biométrie, Université Claude BERNARD (LYON 1)  
69622 VILLEURBANNE Cedex, France

(2) Laboratoire de Toxicologie, Ecole Nationale Vétérinaire  
69260 CHARBONNIERE-LES-BAINS, France

#### RESUME

*Le modèle de GOMPERTZ est utilisé pour analyser les courbes de croissance de jeunes rats musqués. Une tentative d'explication de ce modèle en termes de mécanismes biologiques est proposée : croissance gouvernée par un facteur catalytique (facteur de croissance). Cette approche conduit à deux types de paramétrisation du modèle, l'une explicative (où apparaissent explicitement les paramètres du processus de croissance) mais peu utilisable numériquement pour l'identification du modèle (forte corrélation entre paramètres), l'autre, intermédiaire entre la formulation habituelle et la formulation explicative, ayant de meilleures propriétés pour l'estimation des paramètres. Au niveau biologique, l'analyse proposée permet, entre autres, de confirmer l'hypothèse de variation du taux de croissance suivant la date de naissance des jeunes animaux dans l'année.*

#### 1. INTRODUCTION

La caractérisation de la croissance d'organismes est un problème souvent posé en biologie des populations. La connaissance de ce processus permet notamment d'approcher certaines structures internes d'une population, par exemple la structure en âges à partir de mesures simples comme le poids des individus, ou encore d'étudier l'influence du milieu au niveau de l'individu ou à celui de la population.

Outre les aspects expérimentaux, la caractérisation de la croissance peut être affinée par des outils statistiques généraux, par exemple l'analyse longitudinale (ESTEVE et SCHIFFLERS, 1976). On peut également avoir recours à des modèles spécifiques. En plus de l'aspect descriptif, ce type de modèle a l'avantage d'inclure explicitement ou implicitement des hypothèses sur le déterminisme du phénomène étudié. En revanche, la mise en oeuvre pratique soulève souvent des problèmes techniques ou méthodologiques. L'objectif de ce travail est de montrer, sur l'exemple de la croissance de l'*Ondatra*, différents aspects de l'utilisation d'un modèle connu : le modèle de GOMPERTZ.

Au plan biologique, il s'agit d'approcher indirectement certaines caractéristiques populationnelles, liées à la démographie, peu connues car peu accessibles *in Natura*, en particulier la croissance des jeunes animaux. En effet, certains auteurs ont suggéré que le taux de croissance des jeunes rats musqués est variable selon la date de leur naissance au cours de la saison de reproduction (VINCENT et QUERE, 1972). Il est d'autre part intéressant d'étudier les variations des paramètres de la croissance en fonction des conditions d'élevage et du milieu naturel (Le BOULENGE, 1977).

## 2. DONNEES EXPERIMENTALES

Dans le cadre des problèmes relatifs à la biologie de l'Ondatra dans la DOMBES (région d'étangs au Nord-Est de Lyon), un élevage expérimental de ces animaux a été installé. Le protocole expérimental est le suivant :

-élevage : les animaux sont installés par couples dans des enceintes bétonnées (4m sur 2m), ils disposent d'une fosse leur assurant la disponibilité d'une eau constamment renouvelée, ils sont alimentés régulièrement (nourriture synthétique).

-mesures : après leur naissance onze jeunes de différents couples ont été suivis régulièrement. On dispose ainsi pour tous les animaux de la masse corporelle, de la longueur totale du corps, de la longueur de la queue et de la longueur de la patte arrière gauche (la figure 1 illustre le principe de la prise de données, le tableau 1 est un exemple de données obtenues sur un animal).

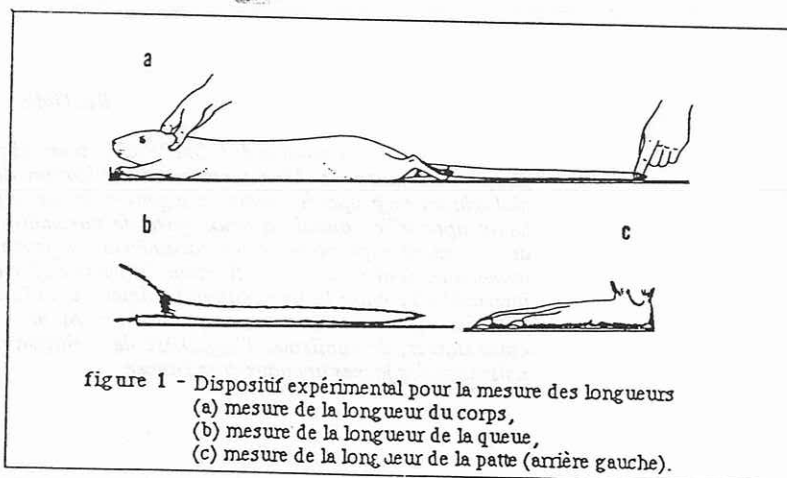


figure 1 - Dispositif expérimental pour la mesure des longueurs  
(a) mesure de la longueur du corps,  
(b) mesure de la longueur de la queue,  
(c) mesure de la longueur de la patte (arrière gauche).

Les animaux 1 à 10 sont nés au printemps 1982, le 11ème est né au printemps 1983.

## 3. CHOIX DU MODELE

Les modèles de croissance des organismes sont nombreux et anciens, dès 1936 TEISSIER présentait une synthèse qu'on peut encore considérer comme d'actualité. L'un des apports postérieurs les plus importants revient sans doute à BERTALANFFY (1949, 1960) qui propose d'exprimer la croissance des organismes supérieurs par le modèle général :

$$\frac{dy}{dt} = a y^m - b y^n \quad (1)$$

T	m	l.p.	l.q.	l.c.	T	m	l.p.	l.q.	l.c.
0	16	2.45	3.6	9.12	104	688	7.70	22.8	50.30
21	116	4.80	10.3	24.60	110	695	7.75	23.1	50.50
29	175	5.60	12.5	30.30	117	712	7.80	23.2	50.70
35	264	6.30	13.3	33.40	124	739	7.85	23.3	50.90
44	352	6.60	15.9	38.60	132	728	7.90	23.4	51.00
50	416	7.00	17.5	41.30	138	747	7.90	23.5	51.20
55	447	7.10	18.7	42.40	146	733	7.90	23.7	51.30
62	503	7.20	20.0	45.40	152	738	7.90	23.7	51.30
69	540	7.30	21.0	46.50	180	763	7.90	23.8	51.70
76	573	7.50	21.0	47.70	187	757	7.90	23.8	51.80
83	603	7.50	21.6	48.30	194	765	7.90	23.8	51.90
90	646	7.60	22.2	49.30	201	767	7.90	23.8	52.00
97	684	7.60	22.6	49.90					

Tableau 1 - Différentes mesures morphologiques effectuées (animal n° 8)  
( T : âge en jours, m : masse corporelle en grammes, l.p. longueur de la patte antérieure gauche, l.q. : longueur de la queue, l.c. : longueur du corps, longueurs en cm)

où  $y$  représente la biomasse,  $a$ ,  $b$ ,  $m$  et  $n$  sont des paramètres positifs. Suivant les valeurs de  $m$  et de  $n$ , on peut obtenir une grande variété de situations, par exemple pour  $m=1$  et  $n=2$  on retrouve le modèle logistique, ce qui fait d'ailleurs que de nombreux auteurs considèrent ce modèle comme une généralisation du modèle logistique (cf., par exemple, LEBRETON et MILLIER, 1982).

Le terme positif, selon BERTALANFFY représente les processus anaboliques, le terme négatif les processus cataboliques. Ce modèle a été largement discuté sur le fond (ZOTINA et ZOTIN, 1972), son utilisation numérique également. En particulier NELDER (1961) remarque que sous sa forme la plus générale, telle qu'elle est exprimée ici, on éprouve beaucoup de difficultés à estimer simultanément les paramètres à partir de données expérimentales.

Certains auteurs ont repris le modèle de GOMPERTZ pour représenter la croissance d'organismes ou de populations cellulaires, notamment les tumeurs cancéreuses (LAIRD, 1964, 1966<sup>a</sup>, 1966<sup>b</sup>; LAIRD *et al.*, 1965, 1968; KENDAL, 1985).

On peut l'écrire sous la forme d'une équation différentielle :

$$\frac{dy}{dt} = a y (\ln K - \ln y) \quad (2)$$

où  $a$  et  $K$  sont des paramètres positifs. On notera que cette expression est la limite de l'équation de BERTALANFFY lorsque  $m=1$  et  $n$  tend vers 1 (NELDER, 1962).

La solution analytique de (2) est

$$y = K e^{\left(\frac{\ln y_0}{K}\right) e^{-at}} \quad (3)$$

Pour  $y_0$  inférieur à  $K$ , le graphe de (3) est une courbe sigmoïde croissante d'asymptote  $K$  ayant un point d'inflexion de coordonnées

$$t_{\text{infl.}} = -\frac{1}{a} \ln \left( \frac{1}{\ln \frac{K}{y_0}} \right), \quad y_{\text{infl.}} = \frac{K}{e}$$

cette solution est utilisée pour représenter la croissance des organismes,  $K$  s'interprète comme la valeur de la variable  $y$  à l'âge adulte. Remarquons que si  $y_0$  est supérieur à  $y_{\text{infl.}}$  le point d'inflexion n'est pas observé et la courbe pour  $t > 0$  est concave. Le paramètre  $a$ , généralement interprété comme un taux de croissance, est invariant par allométrie : on montre aisément que si  $y$  et  $z$  sont deux variables liées par la relation

$$z = A y^\mu$$

$y$  suivant un modèle de GOMPERTZ de paramètres  $(a, K)$ , alors  $z$  suit aussi un modèle de GOMPERTZ de paramètres  $(a, A K^\mu)$ .

Ce modèle a été retenu pour décrire la croissance de l'Ondatra, pour des raisons à la fois empiriques (les propriétés énoncées ci-dessus sont compatibles avec les données expérimentales, comme le montre la figure 2) et théoriques (interprétation en termes de mécanismes biologiques) comme on le verra dans la discussion.

#### 4. ESTIMATION DES PARAMETRES

L'exploitation quantitative de ce modèle demande de mettre au point une méthode d'estimation des paramètres à partir des données expérimentales (ajustement ou identification) puis d'exploiter les résultats obtenus en fonction des propriétés de ces estimations.

##### 4.1. Méthode d'ajustement

L'ajustement du modèle aux données expérimentales a été réalisé sur la base du critère classique des moindres carrés ordinaires. Rappelons que ce critère correspond à un modèle d'erreur du type:

$$y_i = f(\pi^*, t_i) + e_i$$

où  $y_i$  représente la mesure expérimentale au temps  $t_i$ ,  
 $\pi^*$  est le vecteur des vraies valeurs des paramètres  $K$  et  $a$ , c'est-à-dire :

$$\pi^* = \begin{pmatrix} K^* \\ a^* \end{pmatrix}$$

enfin  $e_i$  est un terme d'erreur, de fluctuation aléatoire. On considère qu'il s'agit de la valeur d'une variable aléatoire de moyenne nulle et d'écart-type  $\sigma_i$ . Enfin, si les erreurs sont indépendantes, de même variance et si leur loi est normale, la méthode des moindres carrés est identique à la méthode du maximum de vraisemblance. Dans le cas présent nous supposons ces hypothèses vérifiées.

L'expression (3) étant non-linéaire en fonction des paramètres  $K$  et  $a$ , nous avons utilisé une procédure itérative du type GAUSS-MARQUARDT, il s'agit de la procédure VA07AD

de la bibliothèque d'HARWELL (FLETCHER, 1970). Les aspects statistiques associés à ce genre de méthode sont bien exposés dans la synthèse de VILA (1982), on notera en particulier, qu'en plus des estimations fournies par la procédure de minimisation, il est possible d'obtenir une estimation de la matrice des covariances des paramètres :

$$\hat{\Sigma}_{\pi} = \frac{s(\hat{\pi})}{n-p} (X^T X)^{-1}$$

$s(\pi)$  est la valeur du critère au minimum (i.e. la somme des carrés des écarts correspondant aux valeurs estimées des paramètres),  $n$  est le nombre de données expérimentales (i.e. le nombre de couples  $(t_i, y_i)$ ),  $p$  est le nombre de paramètres (ici  $p=2$ ), enfin  $X$  est la matrice de sensibilité ( $n \times p$ ), calculée pour les valeurs estimées des paramètres,  $X^T$  est sa transposée. On a

$$X = \left( \frac{\partial f(t_i, \pi)}{\partial a} \Big|_{a=\hat{a}} \quad \frac{\partial f(t_i, \pi)}{\partial K} \Big|_{K=\hat{K}} \right)$$

pour  $1 \leq i \leq n$ ,  $f(t_i, \pi)$  est ici l'expression analytique (3) du modèle de GOMPERTZ, ainsi dans notre exemple  $X$  est de dimension  $n \times 2$  ( $n$  données, 2 paramètres).

La connaissance de l'estimation de la matrice des covariances des paramètres permet d'apprécier la précision des estimations obtenues et la corrélation entre estimateurs. Théoriquement ce résultat n'est utilisable que lorsque le critère est peu différent d'un critère quadratique (ce qui est sensiblement vérifié dans notre cas comme le montre la figure 3). On conviendra d'utiliser ce résultat pour définir un indice de précision utile pour comparer des estimations obtenues dans différentes situations, ici les cinétiques de croissance résumées par les paramètres du modèle pour les différents animaux.

#### 4.2. Problème des estimations initiales

Une des difficultés rencontrées dans l'utilisation de ce type d'algorithme itératif est celui des estimations initiales des paramètres pour lancer la procédure itérative de recherche du minimum de la fonction critère. A cette fin, nous avons adapté une méthode utilisée pour d'autres modèles dont l'expression sous forme d'équation différentielle dépend linéairement des paramètres, ou qui peuvent se ramener à ce cas (modèle de VOLTERRA-KOSTITZIN : PAVE, 1979, modèle exponentiel : PAVE, 1982, modèle de MONOD : CORMAN et PAVE, 1983, une présentation générale peut se trouver dans PAVE, 1980).

Reprenons (2), en développant on a

$$\frac{dy}{dt} = (a \ln K) y - a y \ln y \quad (6)$$

la solution peut s'écrire sous la forme

$$y = y_0 + A \int_0^t y(u) du - a \int_0^t y(u) \ln(y(u)) du \quad (7)$$

avec  $A = a \ln K$ .

A partir des données expérimentales, c'est-à-dire  $n$  couples  $(t_i, y_i)$ , on peut

évaluer par intégration numérique (méthode des trapèzes, méthode de SIMPSON...) les quantités

$$X_1(t_i) = \int_0^{t_i} y(u) du \quad \text{et} \quad X_2(t_i) = \int_0^{t_i} y(u) \ln(y(u)) du$$

alors (7) s'écrit

$$y = y_0 + A X_1 - a X_2 \quad (8)$$

Ainsi  $a$  et  $A$  (donc  $\ln K$  et  $K$ ) peuvent être estimés par régression linéaire sur les deux variables  $X_1$  et  $X_2$ . Cependant les estimations obtenues par cette méthode sont biaisées, en effet l'évaluation par quadrature des valeurs  $X_1(t_i)$  et de  $X_2(t_i)$  reporte les erreurs faites sur chaque mesure et introduit un terme d'autocorrélation qui dépend, en outre, de la méthode d'intégration utilisée, c'est pourquoi il est ensuite nécessaire d'améliorer cette estimation par une méthode itérative du type de celle présentée en 4.1.

## 5. RESULTATS

Nous ne présentons en détail que les résultats des ajustements relatifs aux données de masse corporelle (tableau 2). La figure 2 montre un exemple de comparaison entre les données expérimentales et la courbe calculée à partir des estimations obtenues sur ces données (exemple de l'animal n° 8).

n°	d.n.	s	n.d.	$\hat{K}$	$\hat{a}$	$\hat{\sigma}_K$	$\hat{\sigma}_a \cdot 10^4$	$\hat{\rho}_{K,a}$
1	8/05	♀	12	747	286	35	15	-0.91
2	8/05	♂	32	708	292	7	8	-0.61
3	10/05	♂	25	778	225	10	6	-0.77
4	10/05	♀	25	803	262	6	5	-0.73
5	15/07	♂	25	759	347	6	5	-0.67
6	15/07	♂	25	715	325	5	5	-0.69
7	15/07	♂	25	716	364	7	7	-0.65
8	15/07	♂	25	750	357	4	4	-0.66
9	15/07	♀	25	697	366	4	5	-0.65
10	1/07	♀	24	737	342	6	6	-0.66
11	11/06	♂	25	728	265	11	19	-0.75

Tableau 2 - Résultats des ajustements effectués sur les données pondérales.  
(les animaux 1 à 10 sont de 1982, le n° 11 est de 83)

n° : numéros des individus, d.n. : dates de naissance, s : sexe (♂ : mâles, ♀ : femelles),  
n.d. : nombre de données. Estimation du poids adulte ( $K$ ), du taux de croissance ( $a$ ),  
des écarts types et de la corrélation entre paramètres.

(1) Le bon comportement de la méthode d'ajustement peut être expliqué par

- la corrélation relativement peu élevée entre les estimateurs des paramètres (cf. tableau 2 : les corrélations sont toutes, à une exception près, inférieures à 0.9 en valeur absolue),

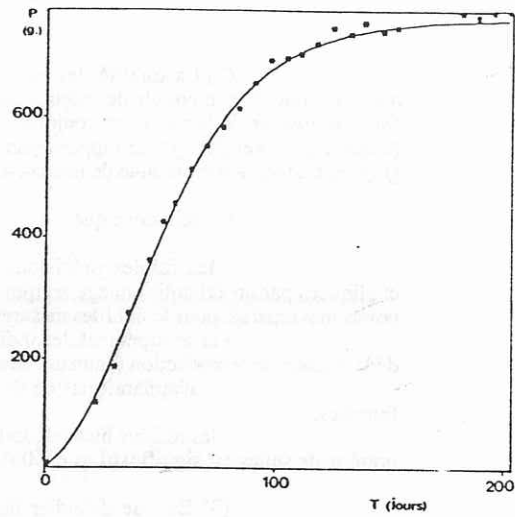


Figure 2 - Exemple d'ajustement du modèle de Gompertz aux mesures de croissance pondérale (animal n° 8). Paramètres :  $a = 0.036 \text{ j}^{-1}$  et  $K = 760 \text{ g}$ , écart-type résid. = 10.7 g.

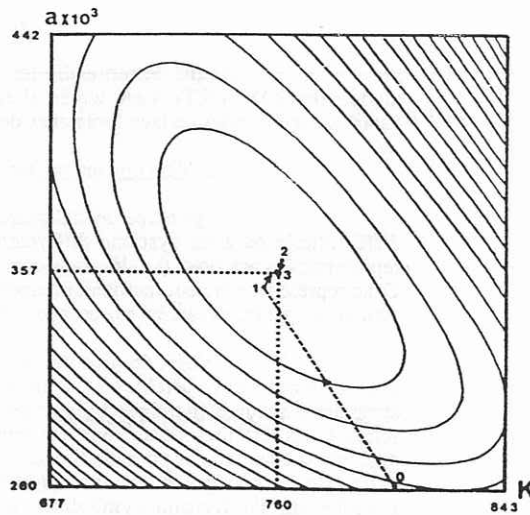


Figure 3 - Courbes isocritères (exemple des données pondérales de l'animal n°8).  
La procédure d'estimation a convergé en 3 itérations:  
0 : estimations initiales  
3 : estimations finales.

- la qualité des estimations initiales qui les place dans une zone sensiblement quadratique du critère, assurant donc une convergence rapide et sûre de la procédure itérative (c'est ce qu'illustre la figure 3).

(2) La qualité des estimations peut être visualisée par le tracé des ellipses d'inertie pour chaque couple de valeurs des paramètres  $a$  et  $K$  (figure 4). L'aire de ces ellipses est fonction inverse de la précision conjointe obtenue pour les deux estimations, et leur allongement (mesuré, par exemple, par le rapport grand axe sur petit axe, ou tout simplement visualisé sur le graphe) est fonction croissante de leur corrélation.

On remarque que

- les faibles précisions enregistrées pour les animaux 1 et 11 peuvent être expliquées par un échantillonnage temporel mal conçu (pour le n° 1 on ne dispose que de peu de points mal répartis, pour le n° 11 les mesures ont été effectuées tardivement),
- le groupement des individus 5 à 10 correspondant à des jeunes nés vers la fin de la période de reproduction (animaux de deuxième portée),
- il n'apparaît pas de différence décelable entre la croissance des mâles et des femelles,
- les résidus bien que faibles paraissent systématiquement autocorrélés (le test du nombre de suites est significatif,  $p < 0.05$ ).

(3) En vue d'étudier les relations d'allométrie entre les différentes variables morphologiques mesurées et le poids des ajustements sur ces données ont été réalisés (pour l'animal n° 8). Dans chaque cas, les résultats obtenus ont été très satisfaisants comme le montrent le tableau 3 et la figure 6.

## 6. DISCUSSION

Avant de reprendre les résultats présentés et de les discuter, une analyse du modèle de GOMPERTZ a été tentée, il en ressort notamment que l'interprétation classique du paramètre  $a$  en termes de taux croissance doit être affinée sinon révisée.

### 6.1. Vers une interprétation biologique du modèle de Gompertz

Le principe de la démarche est d'associer un schéma fonctionnel à une équation différentielle ou à un système différentiel, schéma analogue à ceux utilisés en chimie pour représenter les réactions (i.e. les transformations monospécifiques et interactions interspécifiques). Cette représentation peut faciliter l'interprétation biologique du modèle mathématique. Les aspects techniques généraux ont été exposés par ailleurs (PAVE, 1977, 1979, 1980).

Il s'agit de réécrire, si nécessaire, les équations différentielles, en introduisant éventuellement des variables auxiliaires (nouvelles variables d'état ?) de façon à obtenir des systèmes à seconds membres bilinéaires (voire multilinéaires) en fonction des variables d'état. Alors sous certaines conditions, on peut associer à ces systèmes un schéma fonctionnel analogue à ceux utilisés pour décrire les réactions chimiques, on parle alors de représentation type chimique: dans un précédent article (PAVE et RECHENMANN, 1986), on trouvera un exposé complet sur l'utilisation symbolique de ces schémas, la traduction en termes d'équations différentielles décrivant la dynamique du système, et inversement l'inférence de schémas à partir d'une expression mathématique, c'est d'ailleurs ce second point qui nous intéresse ici. Dans le cas de l'équation de GOMPERTZ, cette approche est relativement simple à exposer :

Reprenons l'équation (2) :

$$\frac{dy}{dt} = a y (\ln K - \ln y)$$



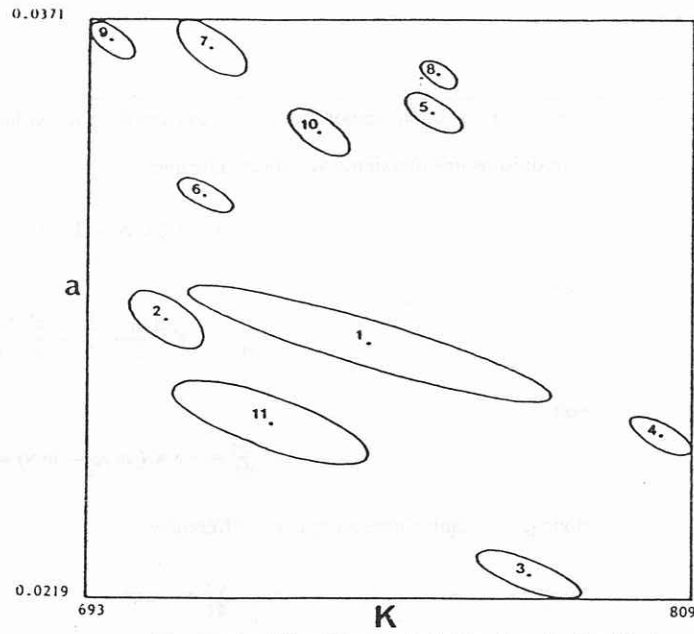


Figure 4 - Répartition des individus dans le plan (K,a).  
Les ellipses d'inertie associées à chaque point permettent d'apprécier la précision relative des estimations obtenues.

$$(\pi - \hat{\pi})^T \hat{\Sigma}_{\pi}^{-1} (\pi - \hat{\pi}) = 1 \quad \text{avec} \quad \pi = \begin{pmatrix} K \\ a \end{pmatrix} \quad \hat{\pi} = \begin{pmatrix} \hat{K} \\ \hat{a} \end{pmatrix}$$

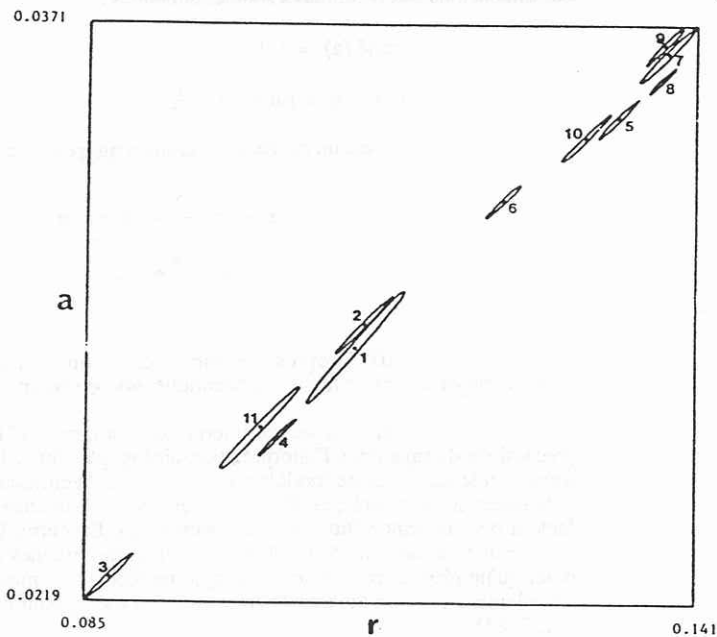


Figure 5 - Répartition dans le plan (r,a).

Ellipses d'inertie associées à ces paramètres  $(B - \hat{B})^T \hat{\Sigma}_B^{-1} (B - \hat{B}) = 1$

$$B = \begin{pmatrix} r \\ a \end{pmatrix} \quad \hat{B} = \begin{pmatrix} \hat{r} \\ \hat{a} \end{pmatrix}$$

posons  $a = a' b$ , on supposera en outre que ces deux nouvelles constantes sont positives, introduisons une deuxième variable  $x$  telle que

$$x = a' (\ln K - \ln y)$$

alors

$$\frac{dx}{dt} = -a' \frac{d \ln y}{dt} = -\frac{a'}{y} \frac{dy}{dt}$$

soit

$$\frac{dx}{dt} = -a a' (\ln K - \ln y) = -a x$$

donc (2) est équivalente au système différentiel

$$\begin{cases} \frac{dx}{dt} = -a x \\ \frac{dy}{dt} = b x y \end{cases} \quad (9)$$

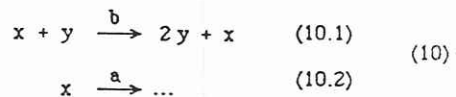
avec les conditions initiales :  $x = x_0$  et  $y = y_0$

Les dimensions des paramètres sont les suivantes :

$$\text{unité (a)} = t^{-1}$$

$$\text{unité (b)} = (\text{unité (x)})^{-1} \cdot t^{-1}$$

(9) est un système différentiel qui peut être déduit du schéma suivant :



(10.1) représente une autoreproduction de  $y$  "catalysée" par un facteur  $x$ , ce même  $x$  disparaissant de façon exponentielle indépendamment de  $y$  (10.2).

On peut penser interpréter  $x$  comme un "facteur de croissance" avec toutes les précautions d'usage dans l'interprétation biologique d'un schéma aussi global. Cependant, on sait d'une part le succès de ce modèle pour représenter la croissance d'un certain nombre d'organismes, notamment les vertébrés, et d'autre part que cette croissance est sous la dépendance d'au moins un facteur de croissance (hormone de croissance). En outre, LAIRD (1964) avait remarqué que le modèle de Gompertz traduisait bien l'évolution de certaines tumeurs, à ce sujet il est intéressant de noter qu'un résultat récent de la biologie moléculaire a montré que dans ce processus intervenait globalement aussi un facteur de croissance, au moins pour certaines d'entre elles (DOOLITTLE et col., 1983).

## 6.2. Reformulation du modèle de Gompertz

En intégrant le système (9), il vient :

$$x = x_0 e^{-a t} \quad (11)$$

$$y = y_0 e^{\frac{r}{a}(1 - e^{-a t})} \quad (12)$$

où

$$r = b x_0 \quad (13)$$

on a évidemment

$$K = y_0 e^{\frac{r}{a}} \quad (14)$$

(i) K dépend explicitement de a,

(ii) on peut montrer que  $y(a)$  est décroissante sur  $\mathbb{R}^+$ , c'est-à-dire que pour  $y_0$ ,  $x_0$  et  $b$  donnés (i.e.  $y_0$  et  $r$  donnés), en prenant deux valeurs de  $a$  telles que  $a_1 < a_2$ , alors pour  $t > 0$  on a  $y(a_1) > y(a_2)$ .

Ainsi, si on retient l'interprétation classique de  $a$  comme taux de croissance, le résultat précédent n'est pas intuitivement satisfaisant : à un "taux de croissance" plus élevé correspond une croissance moins rapide et moins importante. En fait, le paramètre  $a$  apparaît clairement dans le schéma (10) comme taux de dégradation du facteur de croissance, ce qui explique le sens de (ii) ci-dessus : à une valeur élevée de  $a$ , donc à une vitesse de croissance plus lente et à une taille finale plus petite, correspond une disparition rapide des facteurs de croissance, toutes choses égales par ailleurs.

La confusion sous-jacente dans l'interprétation traditionnelle du modèle de Gompertz revient donc à associer au paramètre  $a$  une signification de taux de croissance et de déconnecter  $a$  et  $K$  alors que (14) montre leur dépendance fonctionnelle, dans l'interprétation que nous proposons.

En reprenant les formules (9) (11) et (13), on voit que :

$$\frac{dy}{dt} = r y e^{-a t} \quad (15)$$

à  $t = 0$ , on a

$$\left. \frac{dy}{dt} \right|_{t=0} = r y_0$$

de fait  $r$  représente le taux de croissance "exponentiel" en début de croissance.

Il est clair que les formules (12) et (3) moyennant la relation (14) sont formellement équivalentes, mais elles ne le sont pas sémantiquement. En effet (12) présente l'avantage de décrire la croissance sans tenir compte explicitement de  $K$  (condition finale), mais seulement de la condition initiale  $y$  et des constantes de temps  $r$  et  $a$ .

### 6.3. Reparamétrisation : matrice des covariances des paramètres $r$ et $a$ .

(12) constitue une reparamétrisation "fonctionnelle" du modèle de Gompertz (au sens de l'interprétation biologique). Le paramètre  $r$  peut se calculer facilement à partir de (14) :

$$r = a \ln \frac{K}{y_0}$$

De façon à approcher les propriétés des estimateurs conjoints de  $r$  et de  $a$ , on peut estimer leur matrice des covariances à partir de celle du couple  $(K, a)$ . En effet, très généralement, on peut montrer que si on connaît :

$$\hat{\Sigma}_{\pi} = \frac{s(\hat{\pi})}{n-p} (X_{\pi}^T X_{\pi})^{-1}$$

et les relations de reparamétrisation

$$\begin{aligned} \beta_1 &= f_1(\pi_1, \dots, \pi_p) \\ &\vdots \\ \beta_p &= f_p(\pi_1, \dots, \pi_p) \end{aligned} \quad (16)$$

où  $\pi_1, \dots, \pi_p$  sont les composantes du vecteur de paramètres  $\pi$ , alors

$$\hat{\Sigma}_{\beta} = J \hat{\Sigma}_{\pi} J^T \quad (17)$$

où  $J$  est la matrice jacobienne de (16). Dans notre cas, on a

$$J = \begin{pmatrix} \frac{a}{K} & \ln \frac{K}{y_0} \\ 0 & 1 \end{pmatrix}$$

La figure 5 (à comparer avec la figure 4) montre la répartition des animaux dans le plan des paramètres  $(r, a)$ . On remarque immédiatement que le groupement des animaux est conservé, sinon accentué, mais que la corrélation entre ces nouveaux paramètres est très élevée (en moyenne de 0.99). Ceci résulte de la définition même de la relation entre les deux paramètres.

La reparamétrisation proposée satisfaisante sur le plan de l'interprétation, c'est-à-dire les aspects fonctionnels et conceptuels, l'est beaucoup moins numériquement et statistiquement. En particulier vouloir identifier la forme (12) serait risqué, connaissant les corrélations élevées entre les paramètres  $r$  et  $a$ . Un bon compromis réside dans la formulation suivante :

$$y = y_0 e^{R(1 - e^{-at})}$$

où  $R$  (sans dimension) est le rapport des constantes  $r$  sur  $a$ . Dans ce cas on obtient

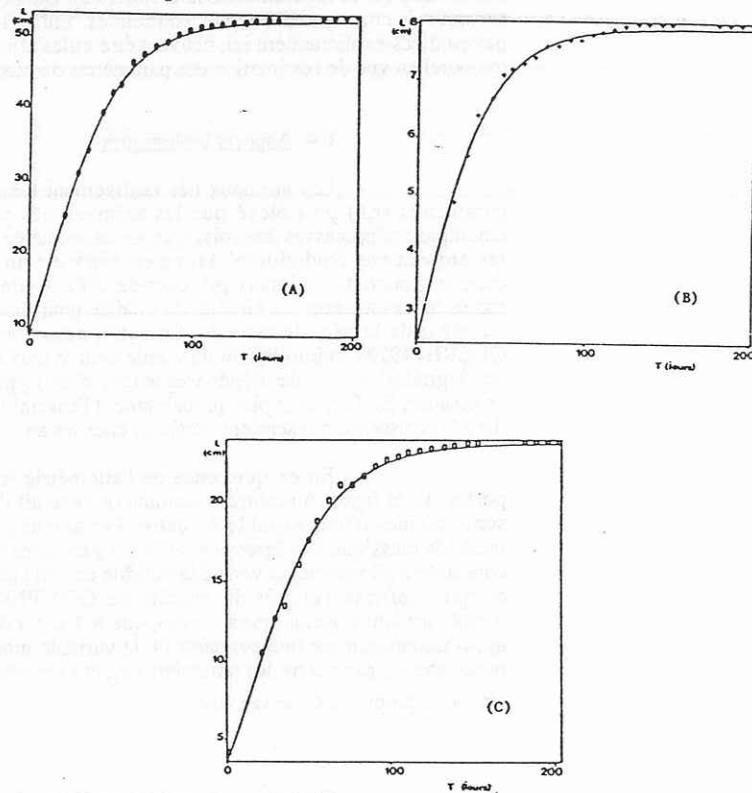
$$J = \begin{pmatrix} \frac{1}{K} & 0 \\ 0 & 1 \end{pmatrix}$$

d'où, d'après (17),

z	Long. pied	Long. queue	Long. corps
A	1.18	0.97	2.93
$\mu$	0.29	0.48	0.44
$\rho$	0.993	0.998	0.997
K'	8.05	23.95	52.94
K	7.86	23.90	51.60
a	0.043	0.036	0.040

**Tableau 3** - Etude de la relation d'allométrie entre la masse corporelle ( $m$ ) et différentes variables morphologiques ( $z$ );  $z = a m^{\mu}$ . Exemple de l'animal n° 8.

K : plateau estimé par ajustement du modèle de GOMPERTZ, K' : plateau estimé par la relation d'allométrie à partir de celui obtenu pour la variable  $m$ , a : taux de croissance estimé par ajustement du modèle de GOMPERTZ,  $\rho$  : coefficient de corrélation entre variables linéarisée pour l'allométrie.



**Figure 6** - Ajustement du modèle de GOMPERTZ à différentes variables morphologiques (animal n°8).

- (A) longueur totale du corps ( $a = 0.040$  j et  $K = 51.6$  cm)
- (B) longueur de la patte antérieure gauche ( $a = 0.040$  j et  $K = 7.86$  cm)
- (C) longueur de la queue ( $a = 0.036$  j et  $K = 23.9$  cm).

$$\hat{\Sigma}_{(R, a)} = \begin{pmatrix} \frac{\text{Var}(K)}{K^2} & \frac{\text{Cov}(a, K)}{K} \\ \frac{\text{Cov}(a, K)}{K} & \text{Var}(a) \end{pmatrix} = \begin{pmatrix} \text{Var}(R) & \text{Cov}(a, R) \\ \text{Cov}(a, R) & \text{Var}(a) \end{pmatrix}$$

il apparaît clairement que  $\text{Corr}(R, a) = \text{Corr}(K, a)$

La représentation dans le plan (R,a) sera donc très proche de celle dans le plan (K, a) présentée en figure 4.

En conclusion, les reparamétrisations présentées pour le modèle de GOMPERTZ ont un intérêt essentiellement théorique et permettent une discussion originale des mécanismes pouvant conduire à retenir ce modèle dans le cadre d'études de croissance. Le cas de la paramétrisation (r, a) montre que l'écriture la plus explicite d'un modèle, du point de vue de l'interprétation, ne correspond pas nécessairement à la forme la plus commode à utiliser, bien qu'en simulation elle puisse permettre des ouvertures intéressantes (par exemple, pour tester des hypothèses sur la régulation de la croissance). Quant à la paramétrisation (R, a), elle offre un bon compromis entre les différentes contraintes. Enfin, les fonctions de sensibilité que nous n'avons pas étudiées explicitement ici, peuvent être utiles afin de définir le meilleur plan d'échantillonnage temporel en vue de l'estimation des paramètres du modèle pour une future campagne de mesures.

#### 6.4. Aspects biologiques

Les animaux nés tardivement (animaux 5 à 10) ont un couple de valeurs des paramètres (r,a) plus élevé que les animaux nés plus tôt dans la saison de reproduction. La cinétique des processus de croissance serait accélérée suivant la date de naissance dans l'année. On remarquera que l'individu n° 11, né en 1983, est un animal de première portée et qu'il est placé dans le groupe des animaux précoces de 1982. Cette observation constitue, sinon une preuve, du moins une indication sur l'intérêt du modèle pour discriminer les animaux du "printemps" de ceux du début de la période estivale. En outre, nous complétons ainsi la remarque de VINCENT et QUERE (1972) en impliquant non seulement le taux de croissance (r) mais également le processus de dégradation, ou de dégénérescence, d'un hypothétique, mais vraisemblable, facteur de croissance. En fait, c'est plus probablement l'ensemble des cinétiques des phénomènes intervenant dans la croissance qui seraient accélérés chez les animaux tardifs.

En ce qui concerne l'allométrie (cf. l'exemple de l'animal n° 8), les résultats portés sur la figure 6 montrent, comme on pouvait l'attendre, que les estimations du paramètre a sont voisines d'une variable à l'autre. Par ailleurs, sur les données du tableau 1, et suivant la méthode classique de régression sur les logarithmes des variables morphologiques, on a estimé la constante d'allométrie, et vérifié la validité de cette relation pour les variables retenues (tableau 3), ce qui confirme l'intérêt du modèle de GOMPERTZ pour décrire l'ensemble des variables morphologiques. Remarquons, enfin, que le facteur de croissance et son processus de dégradation apparaissent comme indépendants de la variable morphologique mesurée (d'où "l'invariance" du paramètre a), par contre les paramètres  $y_0$  et r caractérisent l'état initial, le dimensionnement, et la cinétique propre de cette variable.

On peut se demander quelle est la variable la plus "fiable" de celles observées pour l'ajustement au modèle, avec évidemment une idée à développer de prédiction sur le terrain d'autres variables (par exemple l'âge d'un animal). Il est possible de comparer la qualité des ajustements pour les différentes variables par le calcul de l'écart relatif moyen (E.R.M.) :

$$\text{E.R.M.} = \frac{\sum |y_{\text{obs.}} - y_{\text{calc.}}|}{n \cdot y_{\text{obs.}}} \cdot 100$$

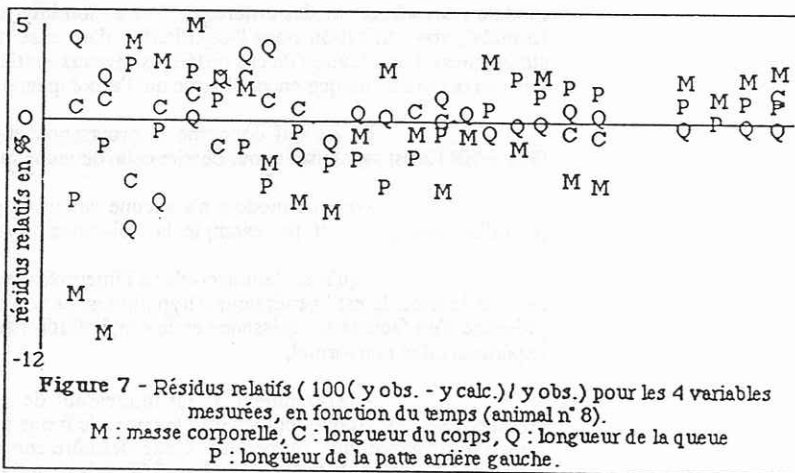
n° animal	masse corp.	long. corps	long. queue	long. patte arr. g.
2	4.22	1.02	1.91	0.93
3	1.33	0.20	1.04	0.62
4	1.97	0.27	0.84	0.56
5	2.71	0.57	0.99	0.82
6	1.95	0.50	1.29	1.03
7	3.32	0.60	1.56	1.05
8	2.27	0.42	0.98	0.93
9	2.28	0.58	0.48	0.92
10	2.39	0.54	1.51	0.44
11	4.47	0.75	0.90	0.30
E.R.M. moyen	2.69	0.54	1.15	0.76

Tableau 4 - Ecart relatif moyen (E.R.M.) pour les quatre variables ajustées par le modèle de GOMPERTZ (en p.100). pour l'ensemble des animaux, sauf l'animal n° 1 car on ne dispose pas pour lui de mesure de longueurs.

On observe (tableau 4) que :

- quelle que soit la variable ou l'animal cet écart est inférieur à 5 p. 100,
- en moyenne, sur l'ensemble des animaux, c'est la masse corporelle qui présente l'E.R.M. le plus élevé. La valeur la plus faible de cet indice est obtenu pour la variable "longueur du corps".

L'intérêt de cette dernière variable est confirmé par l'étude de la figure 7, sur laquelle on a reporté les écarts relatifs :  $100 | y_{obs.} - y_{calc.} | / y_{obs.}$  pour les 4 variables, aux différents temps de mesure dans le cas de l'animal n° 8. Les résidus relatifs de la variable "longueur



du corps" sont plus faibles que ceux des autres variables, notamment pour les premières semaines de la croissance. On note également sur ce graphique une légère tendance à l'autocorrélation des résidus et ce quel que soit la variable considérée. Les fluctuations autour du modèle, bien que faibles, ne sont pas entièrement aléatoires.

Quelle est la variable la plus performante en tant qu'indicateur d'âge ? La masse corporelle est, on l'a vu, une donnée relativement fluctuante, dont la précision relative, bien que bonne, est nettement plus faible que les autres. Elle a cependant l'avantage de pouvoir être utilisée sur une période assez longue (200 jours), car sa valeur initiale est faible comparée à la valeur adulte. Ce n'est pas le cas d'une autre variable comme la "longueur de la patte arrière gauche", qui, bien qu'ayant une bonne précision relative, atteint pratiquement en 50 jours sa valeur adulte, si bien qu'au delà de cette date toute estimation de l'âge basée sur cette variable serait très imprécise (cf. tableau 1). Il reste deux variables, la longueur du corps dont nous avons déjà discuté, mais aussi la longueur de la queue qui est un bon indice d'âge suivant les critères de notre discussion, ce qui va dans le même sens que l'une des conclusions de Le BOULANGE (1977).

Pratiquement, remarquons que le modèle permet d'obtenir une estimation de la date de naissance d'un animal suivi séquentiellement :

- 5 à 6 points peuvent être suffisants pour ajuster le modèle,
- il est ensuite possible d'estimer le temps (par rapport à  $t_0$ , temps de début des mesures) pour lequel la variable prend une valeur donnée par le calcul de la fonction réciproque de (3). Ce peut être la valeur moyenne de la variable à la naissance, caractéristique que l'on peut supposer assez stable d'un individu à l'autre (Le BOULANGE, 1977).

En tout état de cause l'utilisation d'une seule courbe de référence comme "standard" est à proscrire étant donnée la variabilité inter-individuelle du paramètre  $a$ .

## 7. CONCLUSION

Cette contribution illustre plusieurs facettes de la sélection et de l'utilisation d'un modèle : son choix sur des critères de forme, son interprétation en termes fonctionnels (aspects formels), son utilisation pour l'exploitation d'un ensemble de données (aspects numériques et statistiques). L'intégration de ces différents niveaux participe du processus de modélisation et de la mise en oeuvre technique encore lourde qui l'accompagne.

En ce qui concerne la croissance elle-même, il est clair que le modèle de GOMPERTZ est satisfaisant pour décrire celle du jeune rat musqué. Retenons cependant

- que ce modèle n'a aucune valeur de généralité comme modèle de croissance pour d'autres espèces (cf. par exemple, la croissance humaine),

- qu'il ne faut accorder à l'interprétation proposée qu'une valeur heuristique. En ce sens le modèle est "générateur d'hypothèses, et non de fait démontré, la démonstration de la présence d'un facteur de croissance et de son évolution au cours du temps ne peut être que d'ordre expérimental et non formel,

- inversement, il est intéressant de noter que ce modèle, souvent considéré comme "exotique" (comment justifier le terme de freinage logarithmique ?), contient implicitement plus d'information que soupçonné. Cette dernière remarque est sans doute vraie pour d'autres modèles, d'ailleurs J. MONOD, dès 1942, faisait le commentaire suivant : "le contenu d'une expression mathématique est toujours plus riche que ne le croit en général son auteur".



USE AND INTERPRETATION OF THE GOMPERTZ MODEL.  
APPLICATION TO THE STUDY OF THE GROWTH OF YOUNG MUSKRATS  
(*Ondatra zibethica* L.)

## SUMMARY

The GOMPERTZ model is used to analyze muskrats growth curves. We attempt to explain this model in terms of biological mechanisms : growth governed by a catalytic factor (a growth factor). Our approach leads to propose two parametrizations of the model, the first explanatory (parameters of the growth process are explicit) but not usable for parameter estimation (parameters are very correlated), the other, intermediary between the classical mathematical expression and the explanatory one which have better properties for parameter estimation. At biological level the analysis leads particularly to confirm and illustrate the hypothesis of the growth rate variations of young animals relatively to their birth date in the year.

## REFERENCES

- Bertalanffy L.V. (1949). Problems of organic growth. *Nature* 163, 156-158.
- Bertalanffy L.V. (1960). Principles and theory of growth. In *Fundamental aspects of normal and malignant growth* (Nowinsky, ed) Elsevier, Londres, 137-259.
- Corman A., Pavé A. (1983). On parameter estimation of Monod's bacterial growth model from batch culture data. *J. Gen. Appl. Microbiol.* 29, 91-101.
- Doolittle R.F., Hunkapiller M.W., Hood L.E., Devare S.G., Robbins K.C., Aaronson S.A., Antoniades H.N. (1983). Simian Sarcoma virus onc gene, v-sis, is derived from the gene (or genes) encoding a platelet-derived growth factor. *Science* 221, 275-277.
- Esteve J., Schifflers E. (1976). Discussion et illustration de quelques méthodes d'analyse longitudinale. In *Proceeding of the 9th International Biometric Conf.*, the Biometric International Society, 1, 463-480.
- Fletcher R. (1970). A new approach to variable metric algorithm. *Comp. Journal* 13, 317-322.
- Kendal W.S. (1985). Gompertzian growth as a consequence of tumor heterogeneity. *Mathematical Biosciences* 73, 103-107.
- Laird A.K. (1964). Dynamics of tumor growth. *British Journal of Cancer* 18, 490-502.
- Laird A.K. (1966). Dynamics of embryonic growth. *Growth* 30, 263-275.
- Laird A.K. (1966). Post-natal growth of birds and mammals. *Growth* 30, 349-363.
- Laird A.K. (1967). Evolution of the human growth curve. *Growth* 31, 345-355.
- Laird A.K., Tyler S.A., Barton A.D. (1965). Dynamics of normal growth. *Growth* 29, 249-263.
- Laird A.K., Howard A. (1967). Growth curve in inbred mice. *Nature* 213, 786-788.
- Laird A.K., Barton A.D., Tyler S.A. (1968). Growth and time: an interpretation of allometry. *Growth* 32, 334-347.

- Le Boulangé E. (1977). Two ageing methods for muskrats : live or dead animals. *Acta Theriologica*. 22, 509-520.
- Lebreton J.D., Millier C. (1982). Modèles dynamiques déterministes définis par des équations différentielles. In *Modèles dynamiques déterministes en biologie* (Lebreton et Millier, Ed) Masson, Paris, 13-57.
- Nelder J.A. (1961). The fitting of a generalization of the logistic curve. *Biometrics* 17, 89-110.
- Nelder J.A. (1962). An alternative form of a generalized logistic equation. *Biometrics* 18, 614-616.
- Pavé A., Pagnotte Y. (1977). An approach to computer aided design, a tool for mathematical modelling in biology and ecology. *Computer in biology and medicine* 7, 301-310.
- Pavé A. (1979). Dynamics of molecular populations : the evolution of RNA quantities in the silkgland during the last larval instar. *Biochimie* 61, 263-273.
- Pavé A. (1980). Contribution à la théorie et à la pratique des modèles mathématiques pour l'analyse dynamique des systèmes biologiques. Thèse d'Etat, Université Claude Bernard, Lyon, 147 pp.
- Pavé A. (1982). Modèles à compartiments linéaires. In *modèles dynamiques déterministes en biologie* (Lebreton et Millier ed.). Masson, Paris, 99-134.
- Pavé A. Rechenmann F. (1986). Computer-aided modelling in biology : an artificial intelligence approach. In *AI applied to simulation* (Kerckhoffs et al., ed), the Society for Computer Simulation, simulation series 18, 1, 52-70.
- Teissier G. (1936). Les lois quantitatives de la croissance. *Annales de physiologie* 12, 527-545.
- Vila J.P. (1982). Méthodes d'identification des modèles dynamiques. In *modèles dynamiques déterministes en biologie* (Lebreton et Millier ed.). Masson, Paris, 171-195.
- Vincent J.P., Quere J.P. (1972). quelques données sur la reproduction et la dynamique des populations du rat musqué *Ondatra zibethica* L. dans le nord de la France. *Annales de zoologie et d'écologie animales* 4, 395-415.
- Zotina R.S. Zotin A.I. (1972). Towards a phenomenological theory of growth. *Journal of theoretical biology* 35, 213-225.