

Le modèle de Volterra-Kostitzin en dynamique des populations. Ajustement et interprétation des paramètres

J. L. CHASSÉ, J. M. LEGAY et A. PAVÉ

Laboratoire de Biométrie,
Université Claude Bernard Lyon I,
43, Boulevard du 11-novembre-1918
69621 Villeurbanne, France

Résumé

L'usage des moyens modernes de calcul permet l'ajustement aux données expérimentales du modèle intégral-différentiel de Volterra-Kostitzin. Après l'étude mathématique de ce modèle, des considérations numériques sur le modèle, ses cas particuliers et ses approximations sont abordés, ainsi que l'effet de faibles modifications des paramètres sur la forme de la courbe. La méthode d'ajustement proposée est une méthode itérative consistant à minimiser une fonction des moindres carrés. Son utilisation nécessitant des estimations initiales des paramètres, deux méthodes simples d'obtention de ces estimations sont décrites : l'estimation directe sur la courbe et l'estimation sur la dérivée par la méthode des moindres carrés.

L'utilisation de ce modèle en Biologie est illustrée par trois exemples :

- évolution d'une population de *Paramecium caudatum*,
- croissance des follicules ovariens chez *Bombyx mori*,
- évolution du tRNA dans les glandes séricigènes de *Bombyx mori*.

Pour chaque exemple, l'intérêt du modèle est discuté. Outre la diversité des domaines d'application de ce modèle, on peut constater sa qualité descriptive et sa qualité d'outil au niveau de l'interprétation des données.

I. — Introduction

La construction de modèles mathématiques en dynamique des populations a, depuis longtemps, fait l'objet de nombreux travaux (comme le montre la revue de J. D. Lebreton (1973)). Nous nous proposons d'étudier et de donner des applications d'un modèle établi par Volterra :

$$\frac{dp}{dt} = ap - bp^2 - cp \int_0^t p(\gamma) d\gamma \quad (1)$$

p représente la taille de la population

t le temps ; a, b, c sont des constantes positives .

Ce modèle dérive de celui de Malthus (1798) $\frac{dp}{dt} = ap$, d'abord enrichi par Verhulst (1838), dans l'hypothèse d'une population fermée, par l'introduction d'un terme

limitatif de la croissance — bp^2 ; puis par Volterra par un terme intégral qui rend compte de « l'intoxication » du milieu par les individus de la population.

Les solutions de l'équation intégral-différentielle (1) ne peuvent être exprimées que sous forme paramétrique, ce qui rend son emploi malaisé et explique pourquoi l'utilisation de ce modèle a été si peu fréquente. L'usage des moyens modernes de calcul permet son ajustement à des données expérimentales ; nous exposerons précisément ici la technique que nous avons mise en œuvre avec succès dans plusieurs situations.

Par « population » on peut entendre un ensemble d'individus (comme en dynamique des populations) et ceux-ci peuvent être définis au niveau des organismes, des cellules, des molécules... ; la taille de la population peut s'exprimer en nombre d'individus, en dimensions, en activité spécifique, etc... De même, il faut concevoir le terme « intoxication » dans un sens large, c'est-à-dire un terme cumulatif résultant de l'évolution de la population elle-même ; ainsi, il pourra signifier, par exemple, modification du milieu, interaction cellulaire ou même dégradation de molécules.

II. — Etude du modèle

$$\frac{dp}{dt} = ap - bp^2 - cp \int_0^t p(\gamma) d\gamma.$$

Nous rappellerons les principaux résultats concernant cette équation différentielle et indiquerons les approximations qui peuvent être faites.

1. — Intégration de l'équation (1)

Posons

$$q = \int_0^t p(\gamma) d\gamma$$

(1) s'écrit alors :

$$q''(t) = aq'(t) - bq'^2(t) - cq'(t)q(t)$$

Nous devons chercher les solutions qui vérifient :

$$q(0) = 0 \quad q'(0) = p(0) = p_0.$$

On ne sait pas, malheureusement, expliciter les solutions de $q' = \omega(q)$, aussi est-on amené à fournir la solution sous forme paramétrique :

$$\left\{ \begin{aligned} p = \omega(q) &= \frac{a}{b} + \frac{c}{b^2} - \frac{cq}{b} - \left(\frac{a}{b} + \frac{c}{b^2} - p_0 \right) e^{-bq} \end{aligned} \right. \quad (2)$$

$$\left\{ \begin{aligned} t &= \int_0^q \frac{ds}{\omega(s)} \end{aligned} \right. \quad (3)$$

Si $p_0 < \frac{a}{b}$, p passe par un maximum :

$$p_m = \frac{a}{b} - \frac{c}{b^2} \operatorname{Log} \frac{ab + c - p_0 b^2}{c}$$

pour

$$q_m = \frac{1}{b} \operatorname{Log} \frac{ab + c - p_0 b^2}{c}$$

Dans l'intervalle $[0, q_1]$ qui correspond à des valeurs positives ou nulles de p , les variations sont les suivantes :

q	0	q_m	q_1
t	$0 \rightarrow t_m$	$t_m \rightarrow +\infty$	
p	$p_0 \rightarrow p_m$	$p_m \rightarrow 0$	

La détermination des points d'inflexion est intéressante pour la description du phénomène. Leur abscisse est solution de :

$$\frac{dt}{dq} \frac{d^2 p}{dq^2} - \frac{dp}{dq} \frac{d^2 t}{dq^2} = 0$$

c'est-à-dire

$$\omega(q) \omega''(q) + \omega'(q)^2 = 0 \quad (4)$$

(4) peut être résolue par une méthode numérique.

Si $p_0 > \frac{a}{b}$ il n'y a pas de maximum, $p(t)$ est négatif et comme dans le cas précédent la population disparaît.

2. — Cas particuliers de l'équation (1)

Si un ou plusieurs paramètres du modèle (1) sont nuls, on obtient évidemment les « sous-modèles » suivants :

$$b = c = 0 \quad (1) \text{ s'écrit } \frac{dp}{dt} = ap \text{ c'est la loi de Malthus}$$

$$c = 0 \quad (1) \text{ s'écrit } \frac{dp}{dt} = ap - bp^2 \text{ c'est l'équation de Verhulst}$$

mais aussi

$$b = 0 \quad (1) \text{ s'écrit } \frac{dp}{dt} = ap - cp \int_0^\gamma p(\gamma) d\gamma$$

Cette équation intégral-différentielle admet des solutions explicites :

$$p = \frac{4 p_0 (a^2 + 2 c p_0) e^{\lambda t}}{c^2 (q_2 - q_1 e^{\lambda t})^2} \quad (5)$$

avec

$$\lambda = \frac{c}{2} (q_2 - q_1)$$

$$q_2, q_1 \text{ étant racines de } p_0 + a q - \frac{c}{2} q^2 = 0.$$

Ce cas particulier est intéressant car les variations de p en fonction du temps ont une allure semblable à celles obtenues à partir de (1), dans le cas où $p_0 < \frac{a}{b}$. En effet, l'étude de (5) montre l'existence :

$$\text{— d'un maximum } t_m = \frac{-2}{c(q_2 - q_1)} \text{Log} \frac{-q_1}{q_2}; \quad p_m = \frac{a^2}{2c} + p_0 \quad (6)$$

— de 2 points d'inflexion

$$t_{i1} = \frac{2}{c(q_2 - q_1)} \text{Log} \left[\frac{-q_2}{q_1} (2 + \sqrt{3}) \right]; \quad p_{i1} = \frac{2}{3} p_m$$

$$t_{i2} = \frac{2}{c(q_2 - q_1)} \text{Log} \left[\frac{-q_2}{q_1} (2 - \sqrt{3}) \right]; \quad p_{i2} = \frac{2}{3} p_m$$

Dans ce cas, il existe également un axe de symétrie, la droite d'équation $t = t_m$. Notons enfin que (5) peut être exprimé en fonction de p_m :

$$p = \frac{8 p_0 p_m e^{\lambda t}}{c(q_2 - q_1 e^{\lambda t})^2} \quad (7)$$

3. — Approximation de l'équation (1)

Le fait que la solution de (1) s'exprime sous forme paramétrique et que (3) ne soit intégrable que numériquement limite l'utilisation de ce modèle ; d'où l'idée d'obtenir une forme explicite pour la solution, au prix d'approximations. Ces approximations vont porter soit sur les paramètres (par exemple on pourra supposer b négligeable (voir cas particuliers)), soit sur e^{-bq} qui sera approché par un développement limité.

$$\bullet \quad \underline{e^{-bq} \simeq 1 - bq.}$$

A partir de (2) il vient :

$$p = p_0 + (a - p_0 b) q$$

soit

$$p' = (a - p_0 b) p$$

c'est-à-dire

$$p = p_0 e^{(a - p_0 b)t}.$$

On retrouve le modèle exponentiel.

$$\bullet \quad e^{-bq} \simeq 1 - bq + \frac{b^2 q^2}{2}.$$

A partir de (2) il vient :

$$p = \frac{p_0(q_2 - q_1)^2 e^{B(q_2 - q_1)t}}{[q_2 - q_1 e^{B(q_2 - q_1)t}]^2} \quad (8)$$

avec

$$\left\{ \begin{array}{l} B = \frac{1}{2} (ab + c - p_0 b^2) \\ q_1, q_2 \text{ racines de : } p_0 + (a - bp_0)q - \frac{1}{2} (ab + c - p_0 b^2) q^2 = 0. \end{array} \right.$$

Les variations de (8) sont « semblables » à celles de (1) ; dans ce cas les coordonnées du maximum sont :

$$t_m = \frac{1}{B(q_2 - q_1)} \text{Log} \frac{-q_2}{q_1}$$

$$p_m = p_0 + \frac{(a - bp_0)^2}{2(ab + c - b^2 p_0)}$$

Puisqu'on a supposé b petit pour approcher e^{-bq} par un développement limité, l'évolution de p est très proche de celle observée dans le cas où $b = 0$.

4. — Approche de Kostitzin

Kostitzin propose d'étudier le phénomène en 2 phases :

1^{re} phase : $t \in [0, t_m]$

2^e phase : $t \in [t_m, \infty[$

Dans la 1^{re} phase p pourrait être approché par la loi logistique ayant comme limite p_m :

$$p = \frac{p_0 p_m}{p_0 + (p_m - p_0) e^{-at}} \quad t < t_m.$$

Dans la 2^e phase Kostitzin utilise le développement limité, à l'ordre 2, de $e^{-b(a - q_m)t}$. Il vient :

$$p = \frac{p_m}{\text{ch}^2 \left[\gamma \frac{c}{2} (t - t_m) \right]} \quad \text{pour } t > t_m \quad (9)$$

$$\text{avec } \gamma = \sqrt{\frac{2p_m}{c}}$$

Notons que l'expression indiquée dans l'ouvrage de Kostitzin est incorrecte (le terme en γ est absent) ; l'estimation du paramètre c proposée ensuite est donc fautive. Remarquons que (9) est identique à (5) qui représente l'évolution de la population à partir de l'instant initial quand $b = 0$.

III. — Considérations numériques sur l'équation (1), ses cas particuliers et ses approximations

1. — Intégration numérique

Nous avons procédé à l'intégration numérique de l'équation (1) grâce au système MILADIE (A. Pavé et J. D. Lebreton (1973)) pour les valeurs suivantes des paramètres :

$$a = 0,5 \quad b = 0,001 \quad c = 0,000\ 03 \quad p_0 = 10.$$

Ainsi, la courbe 0 (fig. 1) exprime les variations de p en fonction de t ($t \in [0,30]$). Sur

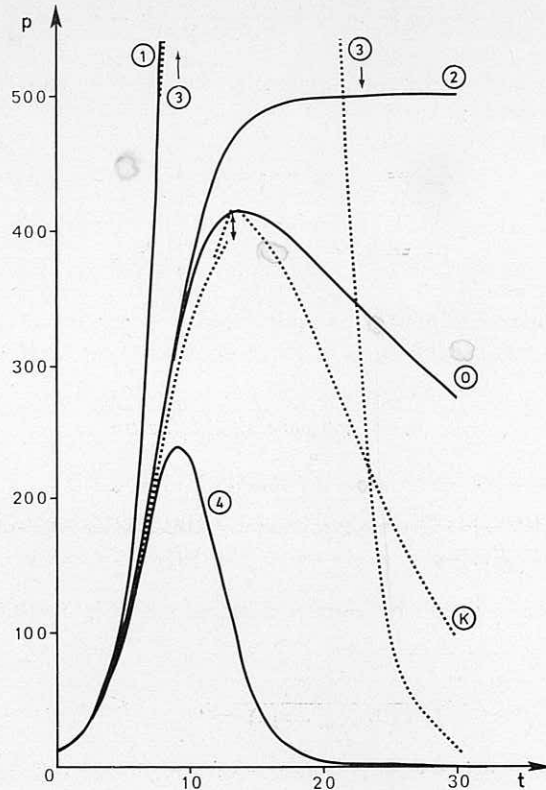


FIG. 1. — Evolution de p en fonction de t

- Modèle de Volterra-Kostitzin $a = 0,5$ $b = 0,001$ $c = 0,00003$
 $p_0 = 10$ Courbe 0
- Cas particuliers : $b = c = 0$ Courbe 1
 $c = 0$ Courbe 2
 $b = 0$ Courbe 3
- Approximations : $e^{-bq} = 1 - bq$ Courbe 1
 $e^{-bq} = 1 - bq + \frac{b^2 q^2}{2}$ Courbe 4
- Approche de Kostitzin Courbe K

le même graphique sont représentées les courbes correspondant aux cas particuliers du modèle (1) ou à ses approximations :

$$\begin{array}{llll}
 b = c = 0 & \text{courbe 1} & e^{-bq} \simeq 1 - bq & \text{courbe 1} \\
 c = 0 & \text{courbe 2} & e^{-bq} \simeq 1 - bq + \frac{b^2 q^2}{2} & \text{courbe 4} \\
 b = 0 & \text{courbe 3} & &
 \end{array}$$

L'examen de la figure 1 nous montre que, bien sûr, seules les courbes 3, 4 et K ont même allure que la courbe 0, c'est-à-dire présentent un maximum. Cependant, pour les courbes 3 et 4, l'écart par rapport à la courbe 0 que l'on peut apprécier sommairement par la différence des maximum est très important. Ces courbes correspondent aux hypothèses $b = 0$ ou b très petit, le désaccord observé pourrait donc s'expliquer, en partie, par la valeur trop élevée de b ($b = 0,001$). Il apparaît donc intéressant de faire varier ce paramètre et de calculer la valeur maximum de p pour les différentes courbes (a , c et p_0 demeurant constants). Les résultats sont les suivants :

b	0	3	4
10^{-3}	414	4 177	241
10^{-4}	2 061	4 177	1 568
10^{-5}	3 763	4 177	3 580
10^{-6}	4 131	4 177	4 108

L'examen de ce tableau montre que l'approximation 4 est la plus intéressante.

2. — Influence d'une faible modification des paramètres

La figure 2 montre l'influence d'une modification de ± 10 p. 100 de l'un des paramètres, ceci pour chacun des paramètres et pour la courbe initiale suivante :

$$p_0 = 10 \quad a = 0,5 \quad b = 10^{-3} \quad c = 3 \cdot 10^{-5}.$$

L'examen de ce graphe montre que la sensibilité des paramètres à une même modification relative croît suivant l'ordre c , b , a .

IV. — Ajustement au modèle de Volterra

Nous disposons de n couples de valeurs expérimentales (t_i, p_i) , le problème est de trouver à partir de ces valeurs les estimations \hat{a} , \hat{b} , \hat{c} des paramètres a , b , c (on suppose que l'estimation de p_0 ne présente pas de difficultés).

Nous allons distinguer 3 types d'approche. La 1^{re}, empirique, consiste à effectuer des approximations des paramètres, « à la main » sur des portions de la courbe expérimentale. La 2^e approche permet d'obtenir des estimations en utilisant une méthode des moindres carrés sur la courbe dérivée. Enfin, la 3^e consiste à minimiser une fonction des moindres carrés sur les points expérimentaux par une procédure itérative. Cette dernière méthode est plus « précise » que les deux autres, elle en est

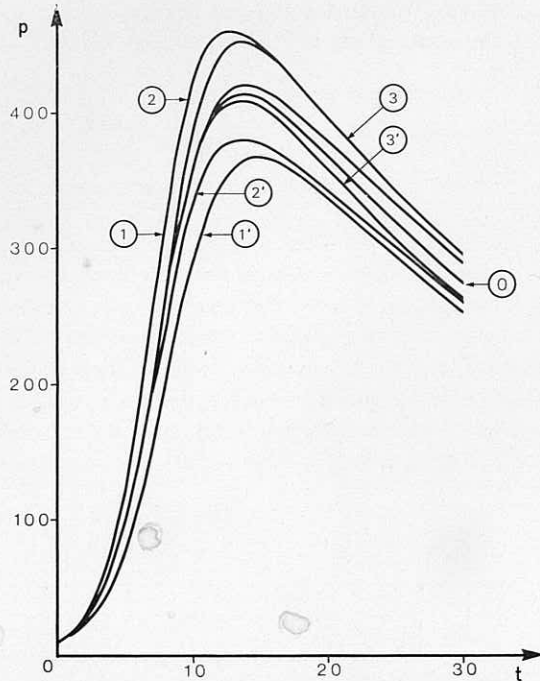


FIG. 2. — Evolution de p en fonction de t . Influence d'une faible modification des paramètres :

0 : courbe de référence	$a = 0,5$	$b = 0,001$	$c = 0,0003$	$p_0 = 10$
1 : $a = 0,55$	2 : $b = 0,0009$	3 : $c = 0,000033$		
1' : $a = 0,45$	2' : $b = 0,0011$	3' : $c = 0,000027$		

complémentaire dans la mesure où elle nécessite des estimations initiales des paramètres ; il est également à noter qu'elle est coûteuse en temps machine.

1. — Estimation directe sur la courbe

La pratique nous a amenés à distinguer deux cas, selon que \underline{b} est petit ou non (on peut juger que \underline{b} est petit si la courbe $p = f(t)$ est peu dissymétrique) :

● b non petit.

Dans ce cas, la courbe (0) est donc dissymétrique (fig. 1), alors :

— pour t très petit il y a accord avec l'exponentielle (1), ce qui permet une estimation de \underline{a} .

— pour estimer \underline{b} on suppose que le maximum observé pour (0) diffère peu de la limite de la loi logistique (2), c'est-à-dire :

$$p_m \simeq \frac{a}{b}$$

En fait $\frac{a}{b}$ est sous-estimé mais comme, de fait, il y a aussi tendance à sous-estimer a , la valeur obtenue pour b n'est pas aberrante.

— c peut être obtenu à partir de la relation (9), comme le fait Kostitzin, en prenant une valeur de p pour $t > t_m$ (il y a intérêt à prendre une valeur de t pas trop éloignée de t_m).

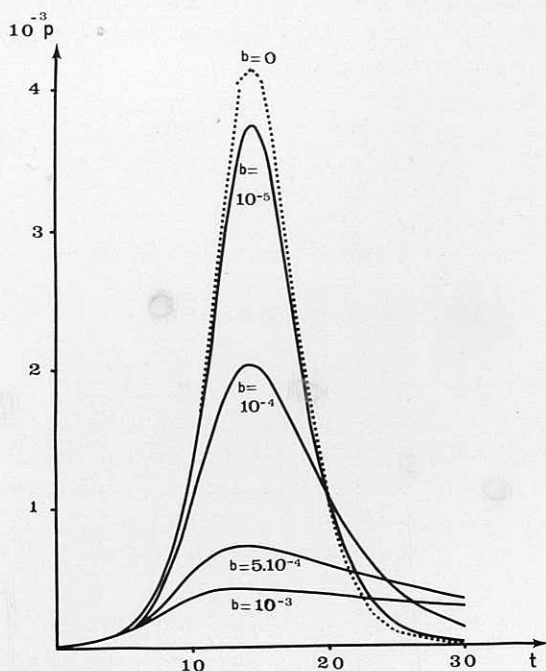


FIG. 3. — Evolution de p en fonction de t pour différentes valeurs de b .
 $a = 0,5$ $c = 0,00003$ $p_0 = 10$.

• b petit.

Dans ce cas, la courbe (0) est peu dissymétrique (fig. 3), alors les paramètres peuvent être estimés ainsi :

- a comme précédemment,
- c comme précédemment ou à partir de la relation $p_m = p_0 + \frac{a^2}{2c}$
- b peut être estimé à partir de la relation :

$$p_m = p_0 + \frac{(a - bp_0)^2}{2(ab + c - b^2 p_0)}$$

mais comme b est très petit, il y a instabilité numérique. Nous proposons donc dans ce cas de choisir comme première estimation de b une valeur proche de 0 ; l'examen de la figure 3 suggère de choisir une valeur de l'ordre de 10^{-4} ou 10^{-5} .

Application.

Les estimations précédentes ont été mises en œuvre sur les 2 types de courbes (fig. 1 et 3), les points étant choisis par une personne ne connaissant pas les vraies valeurs des paramètres. Les résultats ont été les suivants :

$$\text{Figure 1 : } \hat{a} = 0,473 \quad \hat{b} = 0,001\ 15 \quad \hat{c} = 0,000\ 015 .$$

Seule l'estimation c s'éloigne de la vraie valeur (0,000 03) mais pas de façon « dramatique » si l'on a à l'esprit que ces estimations sont les valeurs initiales du calcul itératif qui sera mis en œuvre ultérieurement.

$$\text{Figure 3 : } \hat{a} = 0,499 \quad \hat{c} = 0,000\ 030\ 7 \left(p_m = \frac{a^2}{2c} + p_0 \right) \\ \hat{c} = 0,000\ 032 \text{ (Kostitzin).}$$

2. — *Estimation sur la dérivée par la méthode des moindres carrés*

Considérons l'équation (1) écrite sous la forme :

$$d = ap - bp^2 - cpq \quad \text{avec } q = \int_0^t p(\gamma) d\gamma . \quad (10)$$

Connaissant les valeurs observées de p , q et d il est possible de réaliser un ajustement, par la méthode des moindres carrés, sur la dérivée, en minimisant la fonction :

$$S(a, b, c) = \sum_{i=1}^n (\tilde{d}_i - d_i)^2$$

où d_i représente la valeur « observée » de la dérivée et \tilde{d}_i la valeur calculée suivant (10) au point t_i . Ce calcul nécessite l'estimation des valeurs d_i de la dérivée et du terme intégral q_i en chaque point. d_i est obtenu en dérivant numériquement la courbe expérimentale par des fonctions splines d'ordre 3 (Laurent et Carasso, 1970), q_i est obtenu par intégration numérique par la méthode de Simpson, avec $q_0 = 0$.

S est minimisée par l'annulation des dérivées partielles par rapport aux paramètres (i. e. par annulation du gradient de S) :

$$\frac{\partial S}{\partial a} = 0 ; \quad \frac{\partial S}{\partial b} = 0 ; \quad \frac{\partial S}{\partial c} = 0 .$$

Ainsi les estimations \hat{a} , \hat{b} , \hat{c} sont solutions du système linéaire :

$$\begin{pmatrix} \sum p_i^2 & -\sum p_i^3 & -\sum p_i^2 q_i \\ -\sum p_i^3 & \sum p_i^4 & \sum p_i^3 q_i \\ -\sum p_i^2 q_i & \sum p_i^3 q_i & \sum p_i^2 q_i^2 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} \sum d_i p_i \\ -\sum d_i p_i^2 \\ -\sum d_i p_i q_i \end{pmatrix} .$$

Les valeurs obtenues sont satisfaisantes dans le cas de faibles erreurs expérimentales, sinon l'instabilité numérique de la dérivation peut conduire à des valeurs aberrantes (par exemple des valeurs fortement négatives des paramètres). Néanmoins, il faut signaler que cette approche peut être utilisée pour toute équation différentielle (ou

intégrable, ou système différentiel) qui dépend linéairement des paramètres (soulignons que de tels systèmes sont importants car ils recouvrent une grande partie des modèles déterministes de la cinétique chimique et de la dynamique des populations).

3. — Estimation sur les points par une méthode des moindres carrés itérative

A partir des estimations obtenues par l'une des deux méthodes précédentes, nous calculons, par intégration numérique, les \hat{p}_i correspondant aux différents t_i . L'écart entre les résultats expérimentaux et le modèle est apprécié par la distance :

$$Q = \sum_{i=1}^n (p_i - \hat{p}_i)^2.$$

Nous cherchons ensuite à minimiser la distance Q en modifiant légèrement les estimations des paramètres a , b , c . Plus précisément, nous fixons un pas de variation pour chaque paramètre, soit h_a , h_b , h_c . A chaque étape j du calcul, nous disposons donc, pour chaque paramètre, de 3 valeurs :

$$\begin{cases} a_j + h_a \\ a_j \\ a_j - h_a \end{cases} \quad \begin{cases} b_j + h_b \\ b_j \\ b_j - h_b \end{cases} \quad \begin{cases} c_j + h_c \\ c_j \\ c_j - h_c \end{cases}$$

et nous calculons pour chacune des 27 combinaisons de 3 paramètres la distance Q . Soit Q_j la plus petite de ces 27 distances :

— si $Q_j < Q_{j-1}$ alors le calcul se poursuit, en prenant comme valeurs a_{j+1} , b_{j+1} , c_{j+1} la combinaison à laquelle correspond Q_j .

— si $Q_j = Q_{j-1}$ alors le calcul est poursuivi en diminuant les pas de variation des paramètres.

Dans les deux cas, le calcul est terminé dès que la distance Q_j est inférieure ou égale à la distance Q_A calculée en supposant un écart moyen de 5 p. 100 sur chaque point expérimental, c'est-à-dire :

$$Q_A = \sum_{i=1}^n (0,05 p_i)^2.$$

Le choix de ce critère d'arrêt, outre la réduction du temps de calcul, est fondé sur le fait que nous n'estimons pas nécessaire de trop « raffiner » l'ajustement car, d'une part les valeurs expérimentales (t_i , p_i) sont entachées d'erreurs et d'autre part la méthode d'exploration de la surface $Q = f(a, b, c)$ utilisée ne permet probablement d'obtenir qu'un minimum local.

V. — Utilisation du modèle en biologie

Nous avons confronté le modèle de Volterra à trois situations expérimentales relevant de la dynamique de population, de la croissance d'organes et de la synthèse d'acides nucléiques.

1. — Evolution d'une population de *Paramecium caudatum*

La courbe d'évolution d'une population de paramécies (fig. 4) (Gause, 1934), présente une allure assez fréquemment rencontrée dans ce type d'observations. Généralement, les auteurs décrivent ce type de courbes par le modèle logistique, faisant implicitement l'hypothèse que l'existence d'un maximum est due à des fluctuations aléatoires, ce maximum ferait partie de la suite des valeurs asymptotiques fournies par le modèle. A cette hypothèse implicite nous opposons explicitement la suivante : il existe 2 phases dans l'évolution de la population et nous proposons comme modèle de la 1^{re} phase, que l'on pourrait appeler phase de croissance, le modèle de Volterra. La 2^e phase, dite d'équilibre, commencerait seulement après le maximum atteint vers la fin de la 1^{re} phase ; ce maximum dépasserait systématiquement le point d'équilibre, si bien que la 2^e phase représenterait alors un réajustement logistique.

Dans l'exemple choisi, la 1^{re} phase se déroulerait de 0 à 3 jours. L'estimation des paramètres conduit aux valeurs suivantes $\hat{a} = 2,0295$; $\hat{b} = 0,00542$; $\hat{c} = 0,01252$. Nous avons représenté sur la même figure le modèle logistique ayant même vitesse de croissance initiale ($a = 0,295$) et tendant vers une valeur d'équilibre choisie comme égale à 105 (ce qui conduit à $b = 0,0193$).

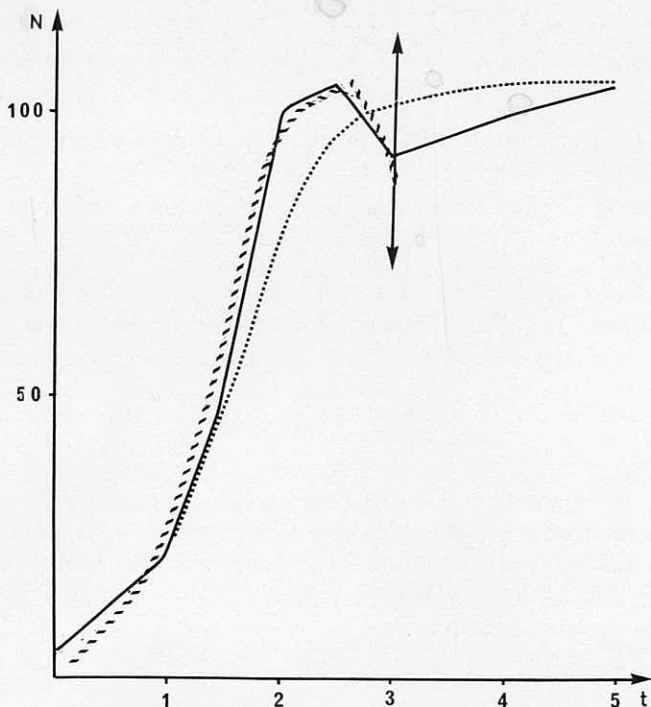


FIG. 4. — Nombre de paramécies/0,5 cm³ (N) en fonction du temps (en jour).

— — — Données expérimentales

- - - - - Modèle de Volterra ; $\hat{a} = 2,0295$; $\hat{b} = 0,00542$; $\hat{c} = 0,01252$

· · · · · Modèle logistique : $\hat{a} = 2,0295$; $\hat{b} = 0,0193$.

Ainsi, la discussion des divers modèles de la même famille (équation (1)) nous permet de proposer une interprétation nouvelle en dynamique de populations d'organismes.

2. — Croissance des follicules ovariens chez *Bombyx mori*

Dans une étude sur la croissance séquentielle de l'ovocyte de Ver à soie pendant le stade nymphal, Legay (1974) a mesuré la longueur et la largeur des ovocytes. De l'ensemble des données, nous avons extrait les résultats concernant un follicule (fig. 5). Ce graphique illustre une des conclusions de Legay : « la croissance individuelle en fonction du temps ne peut être assimilée à une loi logistique. Par contre, le modèle de Volterra appliqué, comme l'avait proposé Kostitzin, à la croissance embryonnaire décrit parfaitement les données expérimentales. » L'aspect descriptif du modèle est le suivant :

— le 1^{er} terme décrit une croissance malthusienne au début du développement embryonnaire.

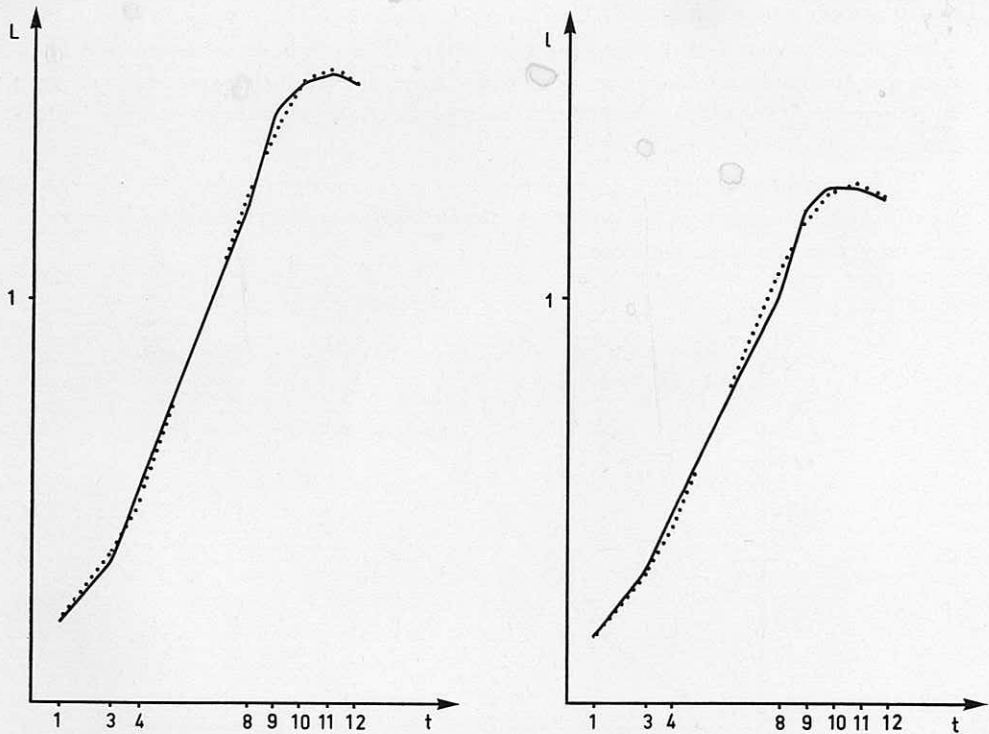


FIG. 5. — Croissance d'un follicule de rang 10, en fonction du temps (en jour).

— Données expérimentales

- - - - - Modèle de Volterra : pour la longueur (L en mm) :

$$\hat{a} = 0,343 ; \hat{b} = 0,0525 ; \hat{c} = 0,0291$$

pour la largeur (l en mm) :

$$\hat{a} = 0,353 ; \hat{b} = 0,0750 ; \hat{c} = 0,0354.$$

— le 2^e terme ($-bp^2$) peut rendre compte d'une certaine concurrence entre follicules au cours de la croissance aussi bien au niveau nutritionnel qu'au niveau purement mécanique (augmentation de la pression avec la croissance).

— le terme intégral, lui aussi limitatif, serait lié à l'accumulation de produits métaboliques défavorables soit au niveau de l'organisme maternel, soit au niveau du follicule, ou des deux.

Ajoutons que cet exemple, outre la qualité descriptive, illustre parfaitement la qualité d'outil du modèle (Legay (1973)). En effet, il a permis d'expliquer certains résultats concernant la croissance séquentielle et de rendre compte de certaines irrégularités observées lors de travaux précédents.

3. — Etude de l'évolution du tRNA chez *Bombyx mori*

Fournier (1974) a étudié l'évolution de différents tRNA dans les glandes séricigènes du Ver à soie en fonction de l'âge. La figure 6 illustre les résultats obtenus pour le tRNA total, des figures analogues sont obtenues pour les tRNA correspondant aux principaux acides aminés étudiés.

Fournier a confronté ses données avec le modèle de Volterra et la figure 6 nous montre la qualité descriptive de ce modèle (des ajustements satisfaisants ont également été obtenus pour les tRNA particuliers). Le rôle d'outil du modèle est ici dans l'interprétation des différents paramètres.

— Le paramètre a , qui correspond à la croissance exponentielle des tRNA apparaît comme commun à l'ensemble des tRNA, il serait, selon l'auteur, lié à la vitesse de transcription de la polymérase.

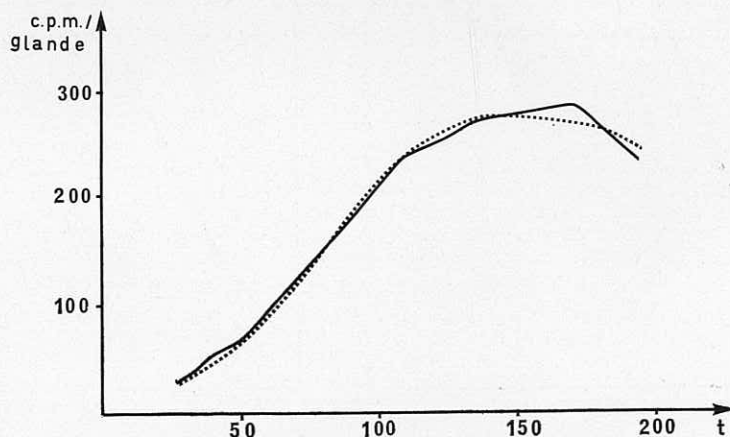


FIG. 6. — Evolution du tRNA total (en coups/minute/glande) dans les glandes séricigènes du Ver à soie, en fonction du temps (en heure).

— — — — — Données expérimentales
 - - - - - Modèle de Volterra $\hat{a} = 0,411$ $\hat{b} = 0,001155$
 $\hat{c} = 0,000049$.

— Le paramètre b qui exprime une diminution du taux de synthèse, varie selon les tRNA. Ce simple constat permet d'éliminer un des deux types d'hypothèses pouvant expliquer cette diminution, et ainsi de réduire le champ d'investigation.

— Le paramètre c exprime une activité de dégradation à effet cumulatif, puisqu'il est lié au terme intégral. Le fait qu'on ne connaisse qu'une ribonucléase chez le Ver à soie et que le paramètre c varie très largement selon les tRNA conduit à admettre que l'attaque par la ribonucléase, des différents tRNA, est modulée par d'autres facteurs.

V. — Conclusions

La méthode d'ajustement au modèle intégro-différentiel de Volterra que nous venons de proposer revêt un caractère empirique et elle doit être susceptible d'amélioration. Pour l'instant nous pensons que sa qualité est de permettre l'utilisation (dans le cas où des approximations ne sont pas possibles), grâce aux moyens de calcul dont nous disposons, d'un modèle dont Kostitzin a montré, il y a déjà longtemps, l'intérêt théorique mais dont l'usage avait été arrêté pour des raisons techniques. En outre, la très bonne qualité des ajustements obtenus suggère une conclusion importante ; à savoir que l'accroissement d'une population n'est pas seulement une fonction simple du temps, indépendante de tout autre facteur. Le fait que cette population puisse atteindre un maximum supérieur au point d'équilibre qui ne sera acquis qu'ultérieurement (et éventuellement peut ne pas l'être) démontre la complexité des interactions dont la courbe de croissance n'est qu'un bilan.

Les 3 exemples que nous avons donnés montrent la diversité des domaines d'utilisation de ce modèle. Ils permettent de constater sa qualité descriptive et sa qualité d'outil au niveau de l'interprétation des données.

Nous pensons qu'un certain nombre de résultats expérimentaux décrits ou interprétés à l'aide du modèle logistique mériteraient d'être repris en utilisant le modèle de Volterra, et pourraient conduire à une meilleure connaissance des phases de croissance en dynamique des populations.

Reçu pour publication en juin 1977.

Summary

Volterra-Kostitzin model. Fitting and interpretation of parameters. Biological examples.

1. — Introduction

Mathematical modelling in population dynamics has been for a long time an important field of investigation. We examine the Volterra-Kostitzin model (1) which derives from the Malthus and Verhulst models by adding an integral term interpreted as a degradation factor :

$$\frac{dp}{dt} = ap - bp^2 - cp \int_0^t p(\gamma) d\gamma.$$

The analytical solutions of the integro-differential equation (1) are parametrical, so their use is not easy, particularly for model fitting to experimental data. We propose numerical procedures to solve this problem, using a computer.

Further, the term « population » can be understood as a set of organisms, cells, molecules, etc., hence its application is varied.

2. — Analytical study

A parametrical solution has been previously proposed ((2) and (3)) by Volterra and Kostitzin. Some particular cases are studied with one or two parameters of (1) null, so explicit solutions are known :

- (i) $b = c = 0$: exponential (Malthus law)
- (ii) $c = 0$: logistic (Verhulst law)
- (iii) $b = 0$: in this case the shape of the curve looks like that of (1) and $t = t_m$ is an axis of symmetry.

Explicit solutions are also obtained when a limited Mac Laurin series is used as an approximation or the exponential term of (1).

The two step approach of Kostitzin gives two distinct analytical solutions, the first one is logistic ($t < t_m$), the second is a solution of the decreasing part of the curve.

3. — Numerical study

Using a numerical procedure for integration of model (1), a numerical study is possible. Figure 1 summarizes results obtained. One can note the spread between real values and the approximations proposed. Also, we show how a family of curves can be generated from small parameter variations (fig. 2).

4. — Model fitting

From experimental data we propose estimates of parameters a , b , c , using three procedures :

The first one consists of obtaining estimates from partitioning of the curve.

The second consists of a least squares method on the derivative. So the problem is to have a good procedure for numerical derivation of the experimental data.

The third is also a least squares method but iterative, directly on the data. It requires initial estimates ; one of the two previous methods can be used for this.

This method gives good estimates, in the sense that the calculated points are near the experimental ones, but cannot estimate parameter precision (i. e. estimation of the covariance matrix of parameters).

5. — Biological examples

The Volterra-Kostitzin model has been applied on experimental data :

(i) In population dynamics : population growth of *Paramecium caudatum* (fig. 4). Using the model it is possible to assume the occurrence of two successive phases in the evolution of this population. The first one can be considered as a solution of Volterra-Kostitzin model (the size of the population at first exceeds the equilibrium state and then decreases), second, a logistic phase, can be considered as a slow evolution toward the equilibrium state. This phenomenon could be seen as a simple autocorrective process.

(ii) In organismic growth : growth of ovarian follicles of *Bombyx mori* (fig. 5). The sequential growth of ovarian follicles can be described using the model in the same way as in the studies of Kostitzin on embryological development. Besides the descriptive interest, the model permits us to explain certain results of sequential growths and particularly irregularities observed in previous studies.

(iii) Evolution of molecular species (molecular biology) : the evolution of tRNA in the silk gland of *Bombyx mori* (fig. 6). Evolution of tRNA in the last instar of the development of the silk-worm can

also be described using the Volterra-Kostitzin model. In this case the significance of the parameters can be related to molecular mechanisms and allows us to reduce experimental fields of investigation.

6. — Conclusion

The interest of this model is illustrated by biological examples as a descriptive tool and as an inference tool. In spite of the imperfect model fitting, it gives better estimates than the Kostitzin empirical method. We think previous studies could be reexamined using this model.

Références bibliographiques

- FOURNIER, A., 1974. Contribution à l'étude de l'adaptation fonctionnelle quantitative des tRNA à la biosynthèse protéique. Thèse 3^e cycle, Lyon.
- GAUSE, G. F., 1934. Untersuchungen über den Kampf ums Dasein bei Protisten. *Biol. Zentralbl.*, **51**, 536.
- KOSTITZIN, V. A., 1937. Biologie mathématique. A. Colin. Paris.
- LAURENT, P. J., CARASSO, C. 1970. Approximation : obtention de la dérivée d'une fonction donnée par points. Procédures Algol en analyse numérique. Ed. CNRS, p. 300-301.
- LEBRETON, J. D., 1973. Introduction aux modèles mathématiques de la dynamique des populations. *Informatique et Biosphère*. Paris.
- LEGAY, J. M., 1973. La méthode des modèles, état actuel de la méthode expérimentale. *Informatique et Biosphère*. Paris.
- LEGAY, J. M., 1974. Sur la croissance séquentielle de l'ovocyte dans le tube ovarique du *Bombyx mori* pendant le stade nymphal. *Ann. Embr. morph.*, **7**, 3, 277-282.
- PAVÉ, A., LEBRETON, J. D., 1973. Miladie : un minilangage pour le traitement numérique d'équations différentielles et de récurrence. *R.A.I.R.O.*, B-2, 73-79.
-