BIOCHIMIE, 1979, 61, 263-273.

Dynamics of macromolecular populations : a mathematical model of the quantitative changes of RNA in the silkgland during the last larval instar.

Alain PAVÉ.

Laboratoire de Biométrie,
Université Claude Bernard Lyon I,
43, Boulevard du 11 Novembre 1918,
69621 Villeurbanne, France.
(and Greco : « Analyse des systèmes », C.N.R.S.).

Résumé.

Le modèle de Volterra-Kostitzin est utilisé pour l'analyse de l'évolution quantitative des RNA dans la glande séricigène de *Bombyx mori* au cours du dernier âge larvaire. On propose d'associer à ce modèle un mécanisme global traduisant notamment la synthèse et la dégradation de ces macromolécules. Les méthodes numériques et statistiques d'utilisation du modèle sont exposées en appendice.

Les variations des quantités de RNA total (essentiellement ribosomiques) ont été comparées après traitement (hormone juvénile) et entre différentes souches. On note l'importance du terme de dégradation, qui explique à lui seul la plupart des différences enregistrées, alors que la synthèse reste relativement stable.

Ces observations peuvent mener à une interprétation moléculaire de l'effet d'une sélection portant sur l'augmentation de la production de soie : plutôt qu'augmenter la productivité de la machinerie cellulaire, la dégradation aurait été limitée.

Summary.

The quantitative changes of RNA in the silkgland of *Bombyx mori* have been studied during the last larval instar by using a mathematical model (Volterra-Kostitzin model). This model can be associated with a global mechanism including synthesis and degradative processes. The numerical and statistical methods used for model analysis are described in an appendix.

Thus we have compared the accumulation of total RNA (essentially ribosomal) after a treatment (juvenile hormone) and between several strains. The importance of the degradative factor is denoted to explain the observed differences, whereas the synthesis rates remain relatively stable.

The last observation may lead us to an interpretation of the molecular effect of a selection to increase silk production : rather than an increase of the productivity of cellular machinery, the degradative process has been limited.

Key words : Mathematical model, Biochemistry, Biometry, Silkworm, RNA.

Introduction.

Very often, when mathematical modelling of dynamics of macromolecules is needed, approaches derived from Biochemical kinetics are attempted. However frequently the lack of data and the formal expression of the model itself (often nonlinear and complex with many unknown parameters) make them difficult to use. Furthermore many hypothesis have to be assumed on involved mechanisms, this last characteristic is in fact their principal property : they are considered as models of knowledge in the precise sense.

Thus we have thought that an other point of view is possible, even desirable, it consists to chose a class of dynamical models in the field of dynamics of populations. Such an approach is based on the assumption that macromolecules can be looked like individuals of populations, in the ecological sense [1]. If we consider on the one

hand the types of interactions between these molecules, which include eventual transitions, and on the other hand the number of involved « individuals », continuous time models seem to be adapted, at least for a first approach, particularly models chosen in the general framework defined by Lotka [2] and Volterra [3].

These models are related to a kind of mass action law (i.e. density dependent models, in ecology). That is to say that in a mixture of individuals of different species, the variations rates of quantities, or densities or concentrations, is proportionnal to theses quantities, or densities, or concentrations. We have shown in a previous paper that such models can be not only considered as representative (or descriptive) ones, but also how a « global » mechanism can be infered from these equations [4], and we have proposed to give a schematic picture of such mechanisms by using a « chemical representation », thus we call these scheme equivalent mechanism in the same sense that equivalent circuit in electronic science.

We consider also modelling as an approach defined on the basis of *objectives* and *data structures*. So a class of model, then a particular one, must be chosen on such considerations. In our case, both a global mechanism representation and comparisons between different experiments were needed, moreover if we consider experimental data (cf. section 2), it appears that the above mentionned class of models could provide a good tool in respect with such constraints, even it is not the only possible way. In fact we have developed some quantitative aspects to reach our objectives as it is shown in section 3 and in appendix.

We have proposed an integro-diffential model to analyse RNA accumulation on the basis of experimental data which show an increase of RNA quantities following by a decrease at the end of the last larval instar. This model has been previously proposed by Kostitzin [5] to study dynamics of some cellular populations, and it appears as a good descriptive model of such phenomena. We shall attempt to show that it can be also a tool for data analysis and for global mechanism analysis at the molecular level. This model can be included in the general framework of Volterra set of equations so we call it the Voleterra-Kostitzin model. We shall refer to the works of Fournier [6] and Prudhomme [7], where a first approach to the use of the Volterra-Kostitzin model in this field has been already attempted. We shall compare the quantitative changes of total RNA quantities between worms treated with juvenile hormone and untreated worms for the strain C124 × S124 (cf. section 2.1.), and finally the differences of RNA changes between several strains (cf. section 2.2) in relation with silk production. In spite of the limited number of data, some important phenomena are detected, particularly the significance of the degradative term, which appears as the most sensitive one to an external action.

1. Characteristics of experimental results-model building and analysis.

Details concerning experimental features can be found in original papers [6, 7, 8]. Thus we just consider data properties in the goal of a modelling approach :

— the measures are the amounts of RNA at definite times along the 5th larval instar in glands, denoted $q_e(t_i)$ or q_{ei}. Note that we have only one variable observed in a complex system.

— for each experiment, data are not numerous (8 to 15 points).

— a datum is in fact an average obtained from several worms (approx. 20).

Then if we denote $q(t_i)$ the expected value of amount of RNA at time t_i we assume that

$$q_e(t_i) = q(t_i) + e_i$$

(i.e. the real measure is the expected value plus an error term e_i).

e_i is a gaussian variable of zero mean and of standard deviation σ, independent of time. In fact this last term is the sum of measure errors and of interindividual variability.

The experimental results are summarized in figure 1 and 2. If we consider the shape of the curves it is easy to see that a possible model to represent it, in the general framework of Volterra set of differential equations, in respect with the limited number of observed strate variable (only one), has been already proposed by Kostitzin [5] in dynamics of cellular populations (micro-organisms and embryonic development) :

$$\frac{dq(t)}{dt} = a\,q(t) - b\,q^2(t) - c\,q(t)\int_0^t q(t)\,dt \qquad (1)$$

a, b and c are constant.

When a and c are positive the solutions of that equation have the same shape that the observed

RNA evolution. However before to accept this model it is necessary to answer the following questions :

— Can this formula be related to a possible « mechanism » ?

— Is it possible to find estimations of parameters a, b and c from experimental data ?

1.1. *Global mechanism.*

The set of Volterra models is in fact bilinear, even multilinear, differential equations. But (1) can be written :

$$\begin{aligned} \frac{dq}{dt} &= aq - bq^2 - cqP \\ \frac{dP}{dt} &= q \end{aligned} \qquad (2)$$

Then a global mechanism can be proposed (4) :

$$\begin{aligned} &\text{(i)} \quad q \underset{b}{\overset{a}{\rightleftarrows}} 2q \\ &\text{(ii)} \quad q \xrightarrow[c_2]{c_1} q + P\ (+\ \ldots) \\ &\text{(iii)} \quad q + P \rightarrow P\ (+\ \ldots) \end{aligned} \qquad (3)$$

Note that c is a complex parameter which is in fact the product of two terms :

$$c = c_1 \times c_2$$

where c_1 is the rate of protein synthesis, and c_2 the rate of the degradation of RNA by the degradative factor P.

Then formula (2) can be written

$$\begin{aligned} \frac{dq}{dt} &= aq - bq^2 - c_2qP \\ \frac{dP}{dt} &= c_1q \end{aligned} \qquad (4)$$

Obviously that picture is not a set of real biochemical reaction but a symbolical scheme which gives some global transitions or interactions. Then it can be interpreted :

(i) is a picture of a limited autoreproduction of q (i.e. RNA). Obviously as these molecule are not known to be autoreproductive, this « reaction » can be seen as a summary of the limited autoreproduction of DNA and of a kind of feedback action of enzymes which catalyze RNA synthesis (cf. Fournier [6]).

(ii) concerns the production of a factor « P » which interacts with the RNA to degrade it in (iii). P can be interpreted as a degradative enzyme, in the circumstance the RNAase.

The last remark leads to consider that RNA accumulation can be explained, at this level, without considering complex mechanisms on regulation of RNAase synthesis (such as repression and derepression), the own dynamics of the process is sufficient. The detection of RNAase activity at the beginning of this instar confirms that last hypothesis [10].

In spite of the proposed formula is a rough model, we note that it summarizes globally the *synthetical activities relatively to RNA populations :* RNA production with reaction (i), and proteins synthesis with reaction (ii) particularly the synthesis of RNAase, and a degradative process represented by the reaction (iii).

1.2. *Model use : estimation and comparison of parameters.*

As one of our objectives is to compare results between different experiments, the problems of parameters estimation (i.e. to find values of parameters) and of comparison of these estimations arise. The details of the methods used can be found in appendix. Here we just give the general scheme of our work.

The estimations of parameters, from experimental data, have been obtained by minimizing the least squares function :

$$S(\theta) = \sum_{i=1}^{n} (q(t_i,\theta) - q_{ei})^2$$

θ is the vector of parameters (i.e. $\theta = \begin{pmatrix} a \\ b \\ c \\ q_0 \end{pmatrix}$)

n is the number of data, q_0 is the initial value of q.

It was assumed that $P_0 = 0$. The importance of this last constraint has been studied, we have found that it introduces a bias in estimation of parameter *a*, but this systematical error can be neglected when we consider the variability of the related estimator.

To solve the problem of minimization of S we have used a special procedure adapted to dynamical systems where explicit solutions are not known (cf. Appendix § 1). We have also conside-

red statistical properties of estimations in the goal of testing parameter values between different experiments (cf. Appendix § 2).

2. Analysis of experimental results.

We have analysed two sets of experimental data. The first one involves quantitative changes of total RNA (where ribosomal RNA represents 90 per cent of the total) induced by a juvenile hormone treatment, and the second one collects results found in literature with a goal of comparing several strains of silkworm.

2.2.1. *Action of an analogue of the juvenile hormone.*

The experiment led us to compare treated and untreated worms to see if that hormone has an action at the studied molecular level (note that the treatment has been applied during the three first days of the last instar). The results, summarized in the figure 1, show an increase of the RNA production correlated with an increase of the duration of this instar for treated worms [8].

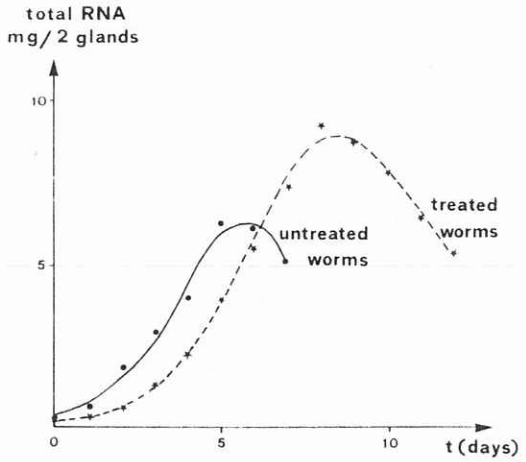

Fig. 1. — *Total RNA.*
Effect of juvenile hormone on dynamics of total RNA.

The analysis of data, on the basis of the Volterra-Kostitzin model, shows (cf. tables I and II) :

- that the rate of synthesis remains remarkably constant,
- no significant difference between the values of the autolimitation term **b**, in spite of a great variation of the absolute values.
- only the degradative term **c** seems to be important to explain the difference of the accumulation of total RNA.

As the estimation of parameters has included data on the treatment period, we have done an other computation starting at the end of this period, then the previous results have been confirmed : the found values do not differ significantly (cf. table I [5]).

Thus we can conclude that the juvenile hormone action has a wide effect, even out of the treatment period, which seems to be associated essentially with the degradative process. However, to have a more precise discussion we shall examine these results with the following ones.

2.2. *Comparison of total RNA variations between several strains.*

To compare sufficiently numerous data, we have used results from three origins : [7, 9], and the above mentioned ones on treated and untreated worms. Figure 2 depicts a simultaneous representation of all data. It is obvious that a qualitative approach to compare them would be uneasy.

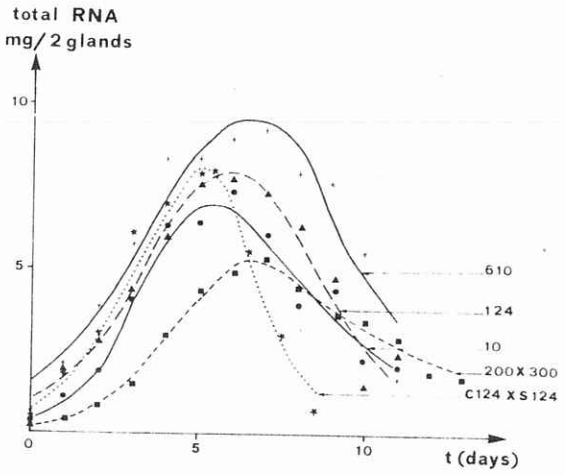

Fig. 2. — *Changes of total RNA during the last instar for several strains.*

Experimental data :

610	
10	
c 124 × s 124	Kurata *et al.* [9]
124	
200 × 300	Prudhomme [7].

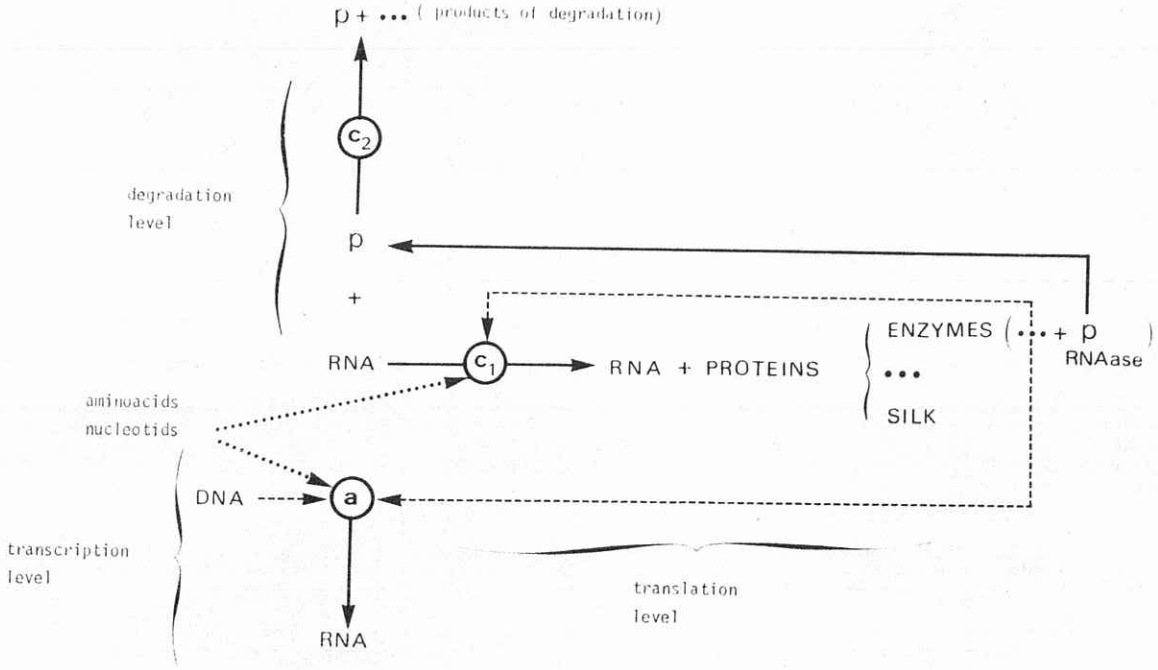

FIG. 3. — *Scheme of molecular system associated to our model.* (a, c_1, c_2 have the same meaning that in 1.1).

TABLE I.

Volterra-Kostitzin model : estimations of parameters for the total RNA in some strains.

Strain	Number of data	Parameters estimations				
		$\hat{a}$	$\hat{b}.\ 10^2$	$\hat{c}.\ 10$	$\hat{q}_0$	silk product.
Untreated worms (1)	8	0.704	0.110	0.412	0.385	350
Treated worms (1)	13	0.665	2.620	0.141	0.178	550
Treated worms (5)	10	0.657	2.526	0.141	—	—
200 × 300 (3)	15	0.791	6.643	0.279	0.187	380 (4)
610 (2)	12	0.498	0.707	0.123	1.470	565
10 (2)	12	0.940	6.830	0.246	0.368	310
124 (2)	12	0.534	0.116	0.198	1.010	435
C 124 × S 124 (2)	14	0.669	− 1.780	0.393	0.842	390

origins of data
(1) Kurata et Daillie [8]
(2) Kurata *et al.* [9]
(3) Prudhomme [7]
(4) Fayard [20]
(5) without considering data of the three first days.

TABLE II.

Comparisons of parameters values for total RNA between several strains by a test on a Fisher-Snedecor variable F.

			values of variable F for each par.			
STRAIN(1)	STRAIN(2)	d.f	F(a)	F(b)	F(c)	$F(q_o)$
untreated worms	*treated worms*	4 ; 13	0.036	0.160	4.456 ///////	1.46
	200 x 300	4 ; 15	0.178	1.276	1.270	1.46
	610	4 ; 12	0.346	0.007	4.570 ///////	1.42
	10	4 ; 12	0.387	0.717	1.400	0.003
	124	4 ; 12	0.278	$0.6\ 10^{-6}$	2.440	1.08
	c124 x s124	4 ; 14	0.179	0.08	0.21	1.90
treated worms	*200 x 300*	4 ; 20	1.40	14.8	170.3 ///////	0.01
	610	4 ; 17	0.447	0.353	0.43	2.94
	10	4 ; 17	0.977	1.094	6.23 ///////	0.50
	124	4 ; 17	0.349	0.574	2.92	2.91
	c124 xs124	4 ; 19	0.006	2.482	32.4 ///////	6.72 ///////
200 x 300	*610*	4 ; 19	1.420	4.030 ///////	39.13 ///////	3.23 ///////
	10	4 ; 19	0.300	0.003	0.71	0.50
	124	4 ; 19	1.350	4.620 ///////	6.96 ///////	3.16 ///////
	c124 xs124	4 ; 21	0.490	10.8 ///////	7.93 ///////	7.09 ///////
610	*10*	4 ; 16	1.395	1.431	6.10 ///////	1.80
	124	4 ; 16	0.012	0.017	3.16 ///////	0.25
	c124 x s124	4 ; 18	0.358	0.381	28.03 ///////	0.66

So the authors have attempted a statistical approach (regression, correlations... cf. for instance Kurata *et al.* [9]), linking some particular characteristics of the average curve (i.e. position of the maximum, total area...) with others results on DNA quantities or silk production, for example. But, despite some correlations, such approaches are limited particularly from a point of view of

			values of variable F for eacn par.			
STRAIN (1)	STRAIN(2)	d.f	F(a)	F(b)	F(t)	F(q_0)
10	*124*	4 ; 16	1.308	1.682	0.81	1.27
	c124 x s124	4 ; 18	0.765	3.410 ///////	6.07 ////////	1.67
124	*c124 x s124*	4 ; 18	0.263	0.215	12.94 ////////	0.10

The significant differences at the critical point $\alpha = 0.05$ are hachured. The column d.f contains the degrees of freedom of the F variable.

mechanism analysis. Thus, once again we have attempted an analysis of data on the basis of the proposed model.

After the estimation of parameters values, we have first compared globally the results (i.e. the four components vectors of parameters values), it appears significant differences between all the strains except between untreated worms of the previous experiment and the strain C124 × S124 of Kurata's measurements ($F \underset{\sim}{} 3.9$ whereas F ($\alpha = 0{,}05$) $\underset{\sim}{} 4.2$) particularly if we consider the other values of the calculated F variable which are all greater than 10. In fact this result is a kind of a confirmation of the statistical approach since these worms come from the same strain, the place of experiments but just differs.

If we look at the table 2 we note, once again, the importance of parameter **c** to discriminate between the RNA accumulation. The synthesis parameter **a** is never significant, and parameter **b** presents few differences in spite of absolute large variations. In fact this last parameter is badly estimated (great variance), that is to say that the response of the model is not very sensitive to a variation of this parameter. We suggest that **b** can be explained more as a measure of spatial constraints (i.e. the number of ribosomes is limited in space, it certainly depends of the size of gland cells).

2.3. *Discussion.*

Consider the proposed model of global mechanism [3], it can be more explicitly written with some assumed interactions as shown in figure 3.

In the following discussion we refer to this global scheme.

A first conclusion deals with the apparent stability of parameter **a** which can be seen either as a stable production of RNA polymerases, or as a limited synthesis due to the limited amounts of DNA : the amount of DNA is at most multiplied by 4 during the last instar (Gillot and Daillie [12]), while the amounts of total RNA is on average increased by 25. That involves quasi stability of transcription rates.

Now consider the variations of parameter **c**, to complete the results summarized in table 1 and 2, we have computed the Spearman rank coefficient of correlation between the production of silk and the values of parameter **c**, we have found $r_s = -0.75$ whereas the critical value is -0.71 for the critical point $\alpha = 0.05$, then a negative correlation can be reasonnably assumed : the increase of the production of silk seems to be linked with a decrease of the value of parameter **c**.

Two alternative hypothesis can be enounced about the localization of the variation of this parameter ; for example if we consider an increase of **c**, it may be related to :

i) An increase of the rate of the protein synthesis c_1, with a constant RNAase activity (c_2 = cst), and a feed-back effect on total amount of RNAase which would imply a decrease of RNA quantities by an increased degradation.

ii) An increase of RNAase activity c_2, with a constant rate of the synthesis of proteins (c_1 = cst).

Now, on the one hand according to our scheme it appears that the quantity of P can be considered roughly as proportionned to the total production of proteins.

On the other hand if we assume that the production of silk is a good picture of total proteins synthesis, it becomes possible to check these hypothesis by considering first that c_2 is constant and arbitrary fixed to 1, and second that c_1 is constant

and arbitrary fixed to 1. Obviously in the two cases the variations of q are the same for each values of c, but not the amounts of proteins.

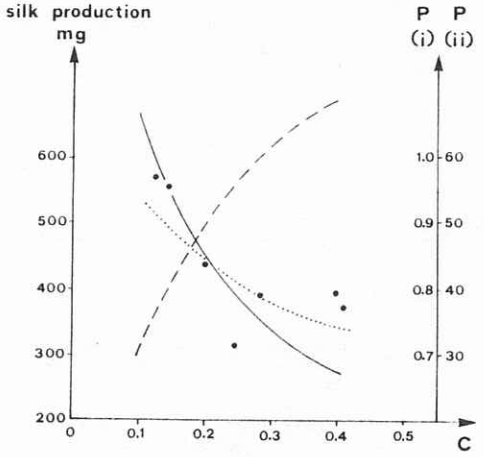

FIG. 4. — *Comparison between silk production and variations of parameter* **c**. *test of hypothesis* (*i*) *and* (*ii*).

● points obtained from experimental results average curve for experimental data simulated curve (**a** = 0.7, **b** = 0.06).

----- with hypothesis (i) : increase of the rate of proteins synthesis.

..... simulated curve for the same values of parameters with hypothesis (ii) : increase of the RNAase activity.

Thus we have computed the asymptotical values of P(c) with fixed values of **a** and **b** for each hypothesis. The results are plotted in figure 4. It appears that the obtained curves have complete opposite shapes which leads to accept the second hypothesis and to reject the first one.

Therefore, in the framework of our model, the variations of parameter c seems to be associated to changes at the RNAase level, either of the rate of RNAse production or of the affinity of that enzyme for RNA. The increase of protein synthesis would be more a consequence of an increase of the amount of RNA (regulated by degradative activity), than an increase of the synthesis rate of proteins.

However, in the proposed model **a**, **b** and **c** are constant parameters, then it doesn't include possible changes governed by the stopping of feeding of the silkworm before spinning cocoon. Such a phenomenon should lead to a relative decrease of cell metabolites, such as aminoacids and nucleotides, thus to limit the synthesis rate after this stopping. A more accurate model whould to take in account a such peculiarity, for example the Volterra-Kostitzin model with variable parameters function of « inputs » such as feeding.

But, on the one hand, it seems that the starvation doesn't provide the same effect on the RNA kinetics if the feeding is experimentaly stopped or if the silkworm stops itself its alimentation [6] : in the first case it appears a discontinuity of RNA changes while in the second case the variation seems to be continuous. In fact just before and during spinning, it seems that some tissues of the worm are partly degraded, but not, or less, if starvation is experimentally induced [13]. Then it is possible to consider that such a degradation provides amino acids and nucleotides and then limit the decrease of parameter **a** (i.e. to remain almost constant the protein synthesis at least for a limited time). Thus to consider this parameter constant is a reasonable constraint although it is, obviously, a simplification.

On the other hand, the exact time of *natural* stopping of feeding is not well known and *never* indicated in the studied experimental results.

Therefore we have limited our analysis to the use of the Volterra-Kostitzin model with constant parameters. However we think that in the future it will be interesting and perhaps necessary, to take in account *variable parameters*.

3. Conclusion.

Our results lead to suggest some hypothesis on the studied system of protein synthesis, particularly the apparent stability of the translation and transcription levels. An increase of the production of proteins seems to be more relevant of *RNA quantity* than an increase of the rates of the above mentionned mechanisms. The RNA quantity itself seems to be governed essentially by a degradative activity where the involved operator would be the RNAse. It appears that this last process has a greater sensitivity to an external action than purely anabolic ones, it could be related to the development of the silkworm along the last instar where degradative processes are probably the most sensitive ones because they change before to reach the larvo-nymphal molt, while the synthesis processes are involved since the beginning of the worm life.

Then an interpretation of the effects of selection can be attempted : the selection has been applied on silkworm with the purpose of increasing silk

production. It seems that it has induced a decrease of degradative rates rather than an increase of synthesis rates.

We think that such results enhance the interest of the study of dynamics of molecular processes, that is to say to take up the work of cell machinery. Our approach based on global models of dynamics of populations could provide a possible framework for studying molecular systems. However, it is important to note that a such data analysis gives more hypothesis on involved mechanisms, which can be tested by experimental ways, than an absolute answer to Biologist's questions : this type of approach cannot provide more informations than the input data do, it is only a possible tool to deal with experimental informations and to propose further development of experimental approaches.

APPENDIX.

We just summarize how results enounced in section 1.2. are obtained more details can be found in Boisvieux [14], Fletcher [15], Harwell [16], Morrisson [17], Pontriaguine [18] and Vila [19].

1. Solution of minimization problem of the least squares function.

Consider the least squares function

$$S(\theta) = \sum_{i=1}^{n} (q(t_i, \theta) - q_{ei})^2 \qquad (A1)$$

Where $q(t_i, \theta)$ *are implicit functions of parameters defined by the differential system :*

$$\begin{aligned} \frac{dq}{dt} &= aq - bq^2 - cqP \\ \frac{dP}{dt} &= q \end{aligned} \qquad (A2)$$

for time values $t_0, t_1 \ldots, t_i, \ldots, t_n$ *:*

q_{ei} *is the experimental value obtained at time* t_i *and* θ *is the vector of parameters*

$$\theta = \begin{pmatrix} a \\ b \\ c \\ q_o \end{pmatrix}$$

A minimum of $S(\theta)$ *corresponds to a solution of the system :*

$$\frac{\partial S(\theta)}{\partial \theta} = 0$$

and

$$\frac{\partial S(\theta)}{\partial \theta_j} = 2 \sum_i (q(t_i, \theta) - q_{ei}) \frac{\partial q(t_i, \theta)}{\partial \theta_j} \qquad (A3)$$

Where θ_j *is one of the parameters :* **a**, **b**, **c**, or $\mathbf{q_o}$.

If $q(t_i, \theta)$ *is explicitly defined (analytical solution of the associated differential system), then partial derivatives relatively to parameters can be calculated, and it is possible to use a numerical procedure to find a solution of the non linear system* (A3). *In that case a Gauss-Marquart method is well adapted (cf. Fletcher in Harwell Lib.* [16]).

A same procedure can be used in our case because if no explicit solution is kown, however numerical evaluation of $q(t_i, \theta)$ *and of partial derivative can be obtained by numerical integration for a current value of parameters (we have used a Runge-Kutta method of order 4). If this procedure is classical for the function* $q(t_i, \theta)$ *it is not yet the case for partial derivatives, so we shall briefly show how simultaneous computing of these terms is possible.*

If we denote ϕ *the right hand side of the differential equation related to* q, *we have by derivation of each term* (*)

$$\frac{\partial}{\partial \theta_j}\left(\frac{dq}{dt}\right) = \frac{\partial \phi}{\partial \theta_j} + \frac{\partial \phi}{\partial q}\frac{\partial q}{\partial \theta_j}$$

And for each parameter we have

$$\frac{d}{dt}\left(\frac{\partial q}{\partial a}\right) = q + (a - 2bq - cP)\frac{\partial q}{\partial a}$$

$$\frac{d}{dt}\left(\frac{\partial q}{\partial b}\right) = -q^2 + (a - 2bq - cP)\frac{\partial q}{\partial b}$$

$$\frac{d}{dt}\left(\frac{\partial q}{\partial c}\right) = -qf + (a - 2bq - cP)\frac{\partial q}{\partial c}$$

$$\frac{d}{dt}\left(\frac{\partial q}{\partial q_0}\right) = (a - 2bq - cP)\frac{\partial q}{\partial q_0}$$

This is an ordinary differential system for partial derivatives of q *relatively to parameters. Then these partial derivatives can be obtained by numerical integration simultaneously with values of* $q(t_i, \theta)$, *with the initial conditions :*

$$\left.\frac{\partial q}{\partial a}\right|_{t=t_o} = \left.\frac{\partial q}{\partial b}\right|_{t=t_o} = \left.\frac{\partial q}{\partial c}\right|_{t=t_o} = 0$$

And

$$\left.\frac{\partial q}{\partial q_o}\right|_{t=t_o} = 1 \text{ (cf. Pontriaguine [18] p. 202).}$$

(*) Only variable q is measured so, in a first approximation, we have considered P as a constant function of parameters (p(t) fixed, and determined from initial values of parameters).

Note that all numerical procedures of minimization of S(θ) need initial values for parameters. In our case it is easy to find such initial conditions :

From (A2) *we can write :*

$$q(t_i) = a\int_{t_o}^{t_i} q(t)\ dt + b\int_{t_o}^{t_i} q^2(t)\ dt + c\int_{t_o}^{t_i} q(t)\ f(t)\ dt + q_o$$

The different integral terme can be estimated from experimental data by numerical integration, rather than classical trapeze or Simpson methods, we have used spline functions, which ensure good numerical properties, to solve this problems. Then an initial value can be obtained by a linear regression.

2. *Statistical point of view of parameter estimation.*

The previous procedure gives estimations of parameters, confidence region can be obtained for the vector θ in R^p *(in our case* $p = 4$*, it is the number of parameters).*

Under assumptions on measure errors (i.e. Gaussian errors), the least squares method is identical to the maximum likelihood method, then an estimation of covariance matrix can be found at the end of minimization procedure above defined. Theorically, we have following results, as the estimator $\hat{\theta}$ *is assymptotically distributed fallowed a gaussian distribution* $N(\theta^*, \Sigma\,\hat{\theta})$

$$(\theta^* - \hat{\theta})'\ \Sigma^{-1}_{\hat{\theta}}\ (\theta^* - \hat{\theta}) = \chi^2_p \qquad (A5)$$

Where θ^* *is the « real » value of the parameters vector,* χ^2_p *is the chi-square variable with* p *degrees of freedom. And the covariance matrix of parameters is*

$$\Sigma\,\hat{\theta} = \sigma^2\ (D'\ D)^{-1} \qquad (A6)$$

σ^2 *is the variance of errors and* D *is the matrix of partial derivatives of functions* ϕ *relatively to parameters, for parameters values obtained at the end of the minimization procedure.* D *is an* $n \times p$ *matrix* (n *is the number of experimental data and* p *the number of parameters) :*

$$D = \left[\left.\frac{\partial q_i}{\partial \theta}\right|_{\hat{\theta}}\right] ;\ (q_i = q(t_i, \theta))$$

On the one hand, from (A5) *and* (A6), *it comes*

$$\frac{1}{p\sigma^2}\ (\theta^* - \hat{\theta})'\ D'D\ (\theta^* - \hat{\theta}) = \frac{\chi^2_p}{p} \qquad (A7)$$

On the other hand we have

$$\frac{1}{\sigma^2}\ S\ (\hat{\theta}) = \sum_i \frac{1}{\sigma^2}\ (q(t_i, \hat{\theta}) - q_{ei})^2 = \chi^2_{n-p}$$

then

$$\frac{1}{(n-p)\sigma^2}\ S(\hat{\theta}) = \frac{1}{n-p}\ \chi^2_{n-p} \qquad (A8)$$

dividing (A7) *by* (A8) *it comes*

$$\frac{(n-p)}{pS(\hat{\theta})}\ (\theta^* - \theta)'\ D'D\ (\theta - \hat{\theta}) = \frac{\chi^2_p\ /p}{\chi^2_{n-p}\ /(n-p)} = F^p_{n-p}$$

then the confidence domain in R^p *at the level* $1-\alpha$ *is given by*

$$\frac{(n-p)}{p\ S(\hat{\theta})}\ (\theta^* - \hat{\theta})'\ D'D\ (\theta^* - \hat{\theta}) = F^p_{n-p\ ;\alpha} \qquad (A9)$$

where $F^p_{n-p;\alpha}$ *is the value of the Fisher-Snedecor variable with* p *and* n-p *degrees of freedom, for the critical point* α *(generally one takes* $\alpha = 0.05$*), and* $\hat{\theta}$ *has the found value of parameters at the end of the minimization procedure. From* (A8) *it comes*

$$S(\hat{\theta}) = \sigma^2\ \chi^2_{n-p}$$

then

$$E(S(\hat{\theta})) = \sigma^2\ (n-p)$$

$$\sigma^2 = \frac{E(S(\hat{\theta}))}{n-p}$$

Thus an unbiaised estimator of σ^2 *is*

$$\sigma^2 = \frac{S\ (\hat{\theta})}{n-p}$$

and an estimation of the covariance matrix of parameters can be proposed at the minimum of the last squares function (i.e. $\hat{\theta}$ *has the found value of parameters, as above mentionned) :*

$$\hat{\Sigma}\hat{\theta} = \frac{S(\hat{\theta})}{n-p}\ (D'D)^{-1} \qquad (A10)$$

(A9) *can be written*

$$\frac{1}{p}\ (\theta - \hat{\theta})'\ \hat{\Sigma}^{-1}_{\hat{\theta}}\ (\theta - \theta) = F^p_{n-p\ ;\alpha}$$

From this last formula it is easy on the one hand to propose a test for comparison of two experiments globally, on the other hand to built a test to compare two estimations of a parameter θ_i *(theses tests are derived by analogy with tests of mean vectors and of linear functions of means).*

i) Global test

We have

$$\frac{1}{p}\ (\hat{\theta}_1 - \hat{\theta}_2)'\ \hat{\Sigma}^{-1}_{\hat{\theta}_1 - \hat{\theta}_2}\ (\hat{\theta}_1 - \hat{\theta}_2) = F^p_{n_1 + n_2\ -2p} \qquad (A11)$$

where $\Sigma \hat{\theta}_1 - \hat{\theta}_2$ *is the estimation of the covariance matrix of the difference between the estimation* $\hat{\theta}_1$ *of parameters vector for the first experiment (with data number* n_1*) and the estimation of parameters vector* $\hat{\theta}_2$ *obtained in the second experiment (where the number of data is* n_2*) :*

$$\hat{\Sigma}_{\hat{\theta}_1 - \hat{\theta}_2} = \frac{(n_1 - p)\,\hat{\Sigma}_{\hat{\theta}_1} + (n_2 - p)\,\hat{\Sigma}_{\hat{\theta}_2}}{n_1 + n_2 - 2p}$$

ii) individual test for parameters.

The test to compare two values of a paremeter is

$$\frac{(\hat{\theta}_{j1} - \hat{\theta}_{j2})^2\,(n_1 + n_2 - 2p)}{p\,((n_1 - p)\,s_{j1}^2 + (n_2 - p)\,s_{j2}^2)} = F^{p}_{n_1 + n_2 - 2p} \qquad \text{(A12)}$$

where

$\hat{\theta}_{j1}$ *(resp.* $\hat{\theta}_{j2}$*) is the estimation of parameter* θ_j *for the 1st (resp. 2d) set of data.*

n_1 *(resp.* n_2*) is the number of data in the 1st (resp. 2d) set of data.*

s_{j1}^2 *(resp.* s_{j2} *is the estimation of variance of* $\hat{\theta}_{j1}$ *(resp.* $\hat{\theta}_{j2}$*) that is to say the* jjth *element of matrix* $\hat{\Sigma}\hat{\theta}_1$ *(resp.* $\Sigma\hat{\theta}_2$*).*

This test gives an answer to the null hypothesis (i.e. the difference between $\hat{\theta}_{j1}$ *and* $\hat{\theta}_{j2}$ *can be explain, or not, by random effects such as noises on measures, sampling process...).*

However it is desirable to specify some practical difficulties in the use and involved conclusions of such procedures. Firstly, it is well known that in non linear cases there is often not only one minimum for the least squares function, so to avoid erroneous conclusions we have applied systematically to each *set of data to be compared the minimization procedure by choosing as starting values of parameters the previously found values obtained for* all *the other sets of data (for example, for total RNA we had 7 sets of data, then 42 compulation have been done, and always the same minimum has been found for each experiment).*

Secondly, the estimation of covariance matrix is not always very good often because the hypothesis on errors of measures are very simple and the power of the test have to be considered, in fact the second one has a bad power, which is illustrated when significant difference is observed for the vector θ *but not for the parameters taken individually. Thus statistical conclusions have to take in account such considerations. However, in spite of these problems, we think that it is better to follow such a way rather than to limit analysis on comparisons of absolute values of parameters which negect the variability of estimations (in our model the case of parameter b is a good example in that direction) a such approach permits, at least, to find for instance the best discriminant parameter. In any case when more complex hypothesis are suggested by experimental features then simulations methods can been used to approach parameters distribution.*

3. Results.

The results of computations are summarized in tables I and II.

REFERENCES.

1. Chassé, J. L., Legay, J. M. & Pavé, A. (1977) *Ann. Zool. Ecol. Anim.*, **9**, 425-441.
2. Lotka,, A. J. (1956) *Elements of mathematical biology*, Dover, New York.
3. Volterra, V. (1931) *Leçons sur la théorie mathématique de la lutte pour la vie*, Gauthier Villars, Paris.
4. Pavé, A. & Pagnotte, Y. (1977) *Comput. in Biol. Med.*, **7**, 301-310.
5. Kostitzin, V. A. (1937) *Biologie mathématique*, Armand Colin, Paris, 66-72.
6. Fournier, A. (1974) *Contribution à l'étude de l'adaptation fonctionnelle quantitative des tRNA à la biosynthèse protéique.* Thèse Doct. 3e cycle, Lyon 1.
7. Prudhomme, J. C. (1976) *Contribution à l'étude de la biosynthèse de la fibroine dans la glande séricigène de* Bombyx mori. Thèse Doct. ès-Sciences, Lyon 1.
8. Kurata, K. & Daillie, J. (1978) *Bull. Sericult Exp. Stat.*, 27 (under press).
9. Kurata, K., Takeshita, H., Shigematsu, H. & Sakate, S. (1974) *J. Sericult. Sci. Japan*, **43**, 296-303.
10. Tashiro, Y., Shimadzu, T. & Shiro, M. (1976) *Cell Struct. and Funct.*, **1**, 205-222.
11. Kirimura, J. (1962) *Bull. Sericult. Exp. Stat. Gov. Gen. Chosen*, **17**, 515.
12. Gillot, S. & Daillie, J. (1968) *C. R. Acad. Sci.*, **266**, 2295-2298.
13. Matsuura, S., Morimoto, T., Nagata, S. & Tashiro, Y. (1968) *J. Cell. Biol.*, **38**, 589-603.
14. Boisvieux, J. F. (1977) *Modélisation et commande des processus biologiques, aspects théoriques et mise en œuvre.* Thèse Doct. ès-Sciences, Paris 6.
15. Fletcher, R. (1970) *Comput. J.*, **13**, 317-322.
16. Harwell (1973) Harwell subrontine library. Theoretical Physics Div. U.K.A.E.E. Research Group.
17. Morrison, D. F (1967) *Multivariate Statistical Methods*, Mac Graw Hill, New York.
18. Pontriaguine, L. (1975) *Equations differentielles ordinaires*, Ed. MIR, Moscou.
19. Vila, J. P. To be published in « Dictionnaire des modèles », Contrat D.G.R.S.T.. Paris.
20. Fayard, J. M. (1976) *Etude biométrique des glandes séricigènes en rapport avec la production de soie chez* Bombyx mori L. Thèse Doct. 3e cycle. Lyon 1.