



Environnement
Canada

Environment
Canada

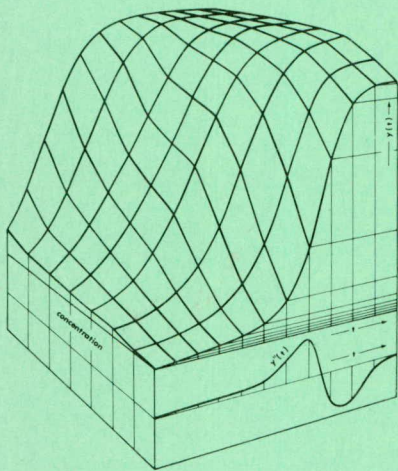
Service
des Forêts

Forestry
Service

22

LOG4, UN PROGRAMME APL D'ANALYSE DES DÉRIVÉES D'UNE FONCTION LOGISTIQUE MODIFIÉE

par Michel Boudoux



CENTRE DE RECHERCHES FORESTIÈRES DES LAURENTIDES
RÉGION DE QUÉBEC, QUÉBEC
RAPPORT D'INFORMATION Q-F-X-29

JUILLET 1972

LOG4, UN PROGRAMME APL D'ANALYSE DES DERIVEES D'UNE
FONCTION LOGISTIQUE MODIFIEE

par Michel Boudoux

CENTRE DE RECHERCHES FORESTIERES
DES LAURENTIDES
REGION DE QUEBEC, QUEBEC
RAPPORT D'INFORMATION Q-F-X-29

JUILLET 1972

TABLE DES MATIÈRES

	Page
RÉSUMÉ.....	4
ABSTRACT.....	5
INTRODUCTION.....	6
DISCUSSION DES DIFFÉRENTES ÉQUATIONS.....	8
DISCUSSION DU PROGRAMME	
a) rôle des sous-programmes.....	15
b) ordinogramme et formats de sortie.....	16
ANALYSE D'UN CAS CONCRET	
a) processus expérimental.....	20
b) interprétation des résultats.....	22
CONCLUSIONS.....	24
FIGURES	
Figure 1: résultats expérimentaux.....	26
Figure 2: valeurs en %.....	27
Figure 3: valeurs des logits.....	28
Figure 4a: fonctions primitives.....	29
Figure 4b: dérivées premières.....	29
Figure 4c: dérivées secondes.....	29
APPENDICES	
Appendice A: ordinogramme.....	30
Appendice B: "listing" des programmes.....	34

	Page
Appendice C: tableaux.....	42
Tableau 2: entrée des données.....	43
Tableau 3: résultats en % de $K(1)$ et $K(2)$	44
Tableau 4: analyse de régression.....	45
Tableau 5: logits observés et "prédits".....	46
Tableau 6: valeurs de $y(t)$, $y'(t)$ et $y''(t)$	47
Tableau 7: analyse de régression.....	48
Tableau 8: logits observés et "prédits".....	49
Tableau 9: valeurs de $y(t)$, $y'(t)$ et $y''(t)$	50
Tableau 10: comparaison des deux fonctions primitives.....	51
Tableau 11: comparaison des deux dérivées premières.....	52
Tableau 12: comparaison des deux dérivées secondes.....	53
RÉFÉRENCES.....	54

RÉSUMÉ

Le programme LOG4 permet l'ajustement d'une fonction sigmoïdale modifiée à des résultats expérimentaux.

Outre l'évaluation des paramètres de cette fonction, LOG4 donne également les valeurs des dérivées premières et secondes.

Un sous-programme fournit la comparaison entre les différentes courbes ajustées à deux ensembles de résultats expérimentaux. Cette comparaison est basée sur l'analyse des dérivées. Après avoir exposé en détail le fonctionnement du programme, on en fera une application à un cas concret.

ABSTRACT

The computer program LOG4 permits the adjustment of a modified sigmoid function to experimental results.

Besides the evaluation of the parameters of this function, LOG4 also gives derived first and second values.

A sub-program gives a comparison between different curves adjusted to two groups of experimental results. This comparison is based on the analysis of derived values. The function of the program is described in some detail, and then applied to a specific case.

INTRODUCTION

Dans de nombreux champs de la biologie expérimentale, les résultats obtenus adoptent souvent l'allure d'une courbe sigmoïdale, courbe dont peut rendre compte l'équation logistique. Néanmoins, la très grande rigidité de celle-ci sur le plan mathématique ne permet pas toujours un ajustement satisfaisant et peut mener, dans certains cas, à une perte importante d'information, sinon à des aberrations.

Le programme LOG4, que nous exposons ici, a pour but de pallier à cet inconvénient, et ceci à deux niveaux. D'abord en modifiant l'équation de la fonction logistique afin de l'assouplir, la rendant ainsi applicable à un plus grand nombre de cas, ensuite en introduisant le calcul des dérivées premières et secondes de la nouvelle équation, avec l'intérêt que cela présente dans l'analyse de la courbe, et partant dans l'interprétation des résultats expérimentaux.

Ce programme a été écrit en APL/360 (Pakin, 1968), étant donné la puissance de ce langage sur le plan mathématique. Néanmoins, LOG4 peut aisément être traduit en un langage plus universel, tel le FORTRAN, afin de pouvoir être employé par un grand nombre d'utilisateurs (une traduction en FORTRAN IV est actuellement à l'étude et sera disponible sous peu).

Le langage APL a ceci de particulier qu'il permet l'introduction des données et le déroulement des différentes phases du programme en mode conversationnel (temps partagé). En effet, les opérations

d'entrées/sorties se font à partir d'une console-terminale (du type IBM/2741) reliée à l'ordinateur au moyen d'une ligne téléphonique. En ne requérant pas de connaissances préalables en informatique, cette façon de procéder rend le programme accessible à un maximum de chercheurs.

En ce qui a trait à sa structure propre, le programme a été conçu de manière modulaire, c'est-à-dire que le bloc principal, LOG4 fait appel à un certain nombre de sous-programmes, en l'occurrence cinq. Ceci permet d'apporter au fur et à mesure des besoins des modifications ou des additions importantes, sans avoir pour autant à réécrire tout le programme.

Sous sa forme actuelle, LOG4 permet éventuellement l'introduction simultanée de deux ensembles de valeurs expérimentales et exécute la comparaison des deux séries de courbes résultantes.

Disons pour terminer que le calcul des paramètres de la fonction primitive est basée sur la méthode des moindres carrés, après transformation des données initiales en "logits". Cette technique a été étudiée récemment par Wilson et Douglas (1970). Ces auteurs suggéraient de remédier à la rigidité de la fonction logistique classique en introduisant dans les calculs un coefficient de pondération. Pour notre part, nous pensons qu'il est préférable d'ajuster une équation polynomiale aux valeurs observées des logits.

DISCUSSION DES DIFFERENTES EQUATIONS

L'équation logistique classique, encore appelée courbe sigmoïdale, fonction de Pearl-Reed ou courbe autocatalytique (Reed et Berkson, 1929) peut s'écrire, t étant la variable indépendante, de la manière suivante

$$y(t) = \frac{K}{1 + e^{-(rt + C)}} \quad (1)$$

ou encore, (Patten, 1971) sous sa forme différentielle:

$$\frac{\partial}{\partial t} y(t) = \{r - Cy(t)\}y(t) \quad (2)$$

ce qui signifie que la dérivée en un point t_i de la courbe est fonction du produit de la valeur de la courbe (en ce point t_i) multipliée par une expression contenant un taux de variation constant, r , et une fonction C de la courbe en ce point.

Si nous reprenons l'équation (1), nous constatons que $y(t)$ peut être défini pour chaque valeur de t , en autant que l'on connaisse les valeurs de K , r et C . Il existe plusieurs approches permettant de déterminer les estimateurs de ces paramètres (équation 2). Parmi les plus intéressantes citons: les techniques d'itération (Marquardt, 1963 et 1966), le calcul des estimateurs à vraisemblance maximum "maximum

likelihood estimators" (Antle et al., 1971; Harter et Moore, 1967; Gupta et Gnanadesikan, 1966; Berkson, 1957) et le développement en séries de Taylor (Glass, 1967). De leur côté, Berkson et Hoodges (1961) emploient la théorie des minimax (Méthode de Chebyshev). Il nous faut signaler aussi la méthode des gradients (Spang, 1962). Toutes ces méthodes ont leurs avantages respectifs, mais elles ont aussi en commun d'exiger des développements mathématiques qui sortent généralement du cadre habituel des préoccupations du biologiste.

Aussi avons-nous choisi d'employer une méthode moins élaborée peut-être mais offrant néanmoins, outre une grande simplicité, toute la précision requise.

Si nous posons provisoirement que dans l'équation (1):

$$rt + C = \phi(t) \quad (3)$$

$y(t)$ aura les propriétés suivantes (Reed et Berkson, 1929):

- a) il existe deux asymptotes horizontales à cette courbe et l'asymptote horizontale supérieure a pour valeur K ,
- b) il existe un point d'inflexion équidistant des deux asymptotes,
- c) la courbe sera symétrique autour de ce point d'inflexion.

On verra plus loin que dans bien des cas, ces trois propriétés - et surtout la troisième - rendent rédhitoire l'emploi de la fonction logistique classique.

On sait que la transformée logarithmique de l'équation (1) est:

$$\text{Ln} \left\{ \frac{y(t)}{K - y(t)} \right\} = \phi(t) \quad (4)$$

Il a été convenu que la quantité de gauche de l'expression (4) serait désignée sous le nom de "logit". On trouvera à la figure 3 une représentation graphique de ces logits pour un exemple donné. Un simple examen de ces points nous indique qu'ils ne peuvent être qu'imparfaitement approximés par une droite comme ce serait le cas si les points expérimentaux relevaient strictement de la fonction logistique.

Or, de tels cas sont on ne peut plus fréquents en biologie. En effet, quels sont les résultats expérimentaux qui peuvent être traduits rigoureusement par une courbe répondant aux propriétés de la courbe logistique?

En réalité cette fonction est beaucoup trop rigide et diverses tentatives ont déjà été faites pour la modifier (Wilson et Douglas, 1971). Ces auteurs suggéraient l'emploi d'un coefficient de pondération et il est bien évident que leur méthode présente un réel intérêt dans bien des cas. Néanmoins, nous pensons qu'il existe une autre manière d'améliorer l'ajustement aux points expérimentaux, c'est de considérer que les logits ne sont pas situés *a priori* sur une droite, mais plutôt sur une courbe polynomiale. De nombreux essais numériques nous ont permis d'établir que le meilleur ajustement était fourni par une

équation polynomiale du quatrième degré en t . En effet, parmi les très nombreux tests que nous avons fait subir au programme LOG4, employant simultanément des équations polynomiales du deuxième, troisième et quatrième degré, il est arrivé dans certains cas, assez rares il est vrai, que la courbe polynomiale du deuxième degré (parabole) présentait un minimum à l'intérieur des valeurs considérées i.e. les valeurs des logits calculés suivant l'équation (4). Ceci était attribuable à un degré d'ajustement insuffisant ou plus exactement à un manque de souplesse de la fonction. Or, une telle situation est à proscrire complètement, parce que dans ce cas, l'équation (6) n'était plus asymptotique à sa limite inférieure parallèlement à l'axe des abscisses, mais au contraire se relevait pour t tendant vers zéro, ce qui est aberrant. Cette situation s'est produite beaucoup moins fréquemment lorsqu'on employait une courbe du troisième degré, mais néanmoins, pour des raisons de sécurité nous avons décidé d'employer une équation polynomiale du quatrième degré qui a toute la flexibilité voulue pour s'adapter à des valeurs expérimentales qui seraient très peu ou pas croissantes aux limites inférieures de t .

Nous avons ainsi modifié l'équation (3) qui devient (voir figure 3):

$$\phi(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 \quad (5)$$

Les coefficients β_i de l'expression (5) peuvent aisément être estimés

par la méthode des moindres carrés en se servant de la relation (4) grâce à laquelle les logits sont directement calculés à partir des valeurs expérimentales.

Certes, l'équation (5) n'a pas la même résonance sur le plan biologique que l'équation (3), en l'occurrence:

$$rt + C = \phi(t) \quad (3)$$

Celle-ci constitue à proprement parler un "modèle mathématique" d'une réalité physique.

Si on la replace dans son champ d'application (équation 2), on constate que les paramètres r et C s'interprètent en termes strictement biologiques. En effet, ... the first form is $dN/dt = (r - CN)N$,^(*) which displays the growth rate (the terms in parentheses) as a decreasing function of population size. The constant C represents reduction of the growth rate per unit of N . The product CN is the total reduction, often termed "environmental resistance"... (Patten, 1971).

Or, rien de tel n'apparaît dans l'équation (5), qui n'est d'ailleurs qu'un ajustement mathématique d'une courbe à des points expérimentaux, ou plus exactement à des valeurs calculées. Mais en élaborant ce programme, notre propos n'était pas tant de présenter un nouveau modèle ou une amélioration à une méthode d'estimation des paramètres, que de mettre dans les mains du biologiste un outil mathématique

(*) N ayant ici la même signification que $y(t)$ dans l'équation (1).

à la fois simple et valable, lui permettant soit d'étudier une courbe de croissance, soit encore de comparer quantitativement plusieurs courbes entr'elles. Plutôt que d'élaborer sur la signification des paramètres r et C , notre démarche a été de chercher à obtenir avec autant d'exactitude que possible une fonction (équation 6, infra) rendant compte du phénomène étudié et ensuite d'étudier cette fonction par le biais de ses dérivées premières et secondes.

Au lieu de considérer comme l'a fait Patten (Patten, 1971), et avant lui de nombreux auteurs, que le terme " $r - CN$ " représente un *taux de croissance*, nous prendrons les dérivées premières et secondes, soient $y'(t)$ et $y''(t)$, et nous montrerons qu'elles constituent en réalité une *vitesse de croissance* et une *accélération* et que le point d'inflexion qui existe dans toute fonction sigmoïde est un moment où l'on passe d'un mouvement accéléré à un mouvement décéléré. Nous reviendrons plus loin sur ces points.

Les différentes équations seront donc les suivantes:

a) équation primitive

$$y(t) = \frac{K}{1 + e^{-\phi(t)}} \quad (6)$$

où K peut être estimé par la valeur maximum de $y(t)$ et où $\phi(t)$ est donné par la relation (5).

b) dérivée première

$$\frac{\partial}{\partial t} y(t) = y'(t) = y(t) \left\{ \frac{\theta(t)}{1 + e^{-\phi(t)}} \right\} e^{-\phi(t)} \quad (7)$$

$$\text{où: } \theta(t) = \frac{\partial}{\partial t} \phi(t) = \beta_1 + 2\beta_2 t + 3\beta_3 t^2 + 4\beta_4 t^3 \quad (8)$$

c) dérivée seconde

$$\frac{\partial^2}{\partial t^2} y(t) = y''(t) = y'(t) \left\{ \frac{\rho(t)}{\theta(t)} - \theta(t) + 2 \frac{\theta(t) e^{-\phi(t)}}{1 + e^{-\phi(t)}} \right\} \quad (9)$$

$$\text{où: } \rho(t) = \frac{\partial}{\partial t} \theta(t) = 2\beta_2 + 6\beta_3 t + 12\beta_4 t^2 \quad (10)$$

DISCUSSION DU PROGRAMME

Le programme LOG4 a été bâti en utilisant au maximum les possibilités offertes par les sous-programmes. Ceux-ci, au nombre de cinq, sont: INFOR, PCENT, REG*, INV* et RES*. On trouvera à l'Appendice B (page 34) la liste des instructions de ces différents programmes.

a) rôle des sous-programmes

Le sous-programme INFOR a pour but de permettre l'entrée des données expérimentales à partir d'une console-terminale IBM/2741. En outre, INFOR pose également à l'utilisateur une série de questions, des réponses desquelles dépendra la suite du traitement. La première question posée concerne le nombre d'ensembles de données qui devront être considérés. En effet, LOG4 peut exécuter le traitement d'un groupe de données soit Y(1), mais il peut également traiter consécutivement deux ensembles de résultats expérimentaux et comparer ensuite les résultats obtenus. Dans ce dernier cas, les deux ensembles seront désignés respectivement par Y(1) et Y(2). Cette information étant fournie à l'ordinateur, INFOR demandera dans un ordre croissant les valeurs observées. D'abord on entrera les valeurs de la variable indépendante t, ensuite les différentes valeurs de Y(1) et, s'il y a lieu, celles de Y(2).

* Ces sous-programmes appartiennent aux "librairies" du système APL/360 du Centre de Traitement de l'Information (Université Laval, Québec, P.Q.).

Après l'impression d'un premier tableau (cfr infra: formats de sortie), INFOR demandera si l'on désire que LOG4 traite les données telles que fournies, ou s'il est nécessaire d'exécuter une transformation de ces données en pourcentage de leur valeur maximum K . Un des principaux intérêts de cette transformation est le suivant: il arrive fréquemment que l'on désire comparer l'allure (par exemple le déroulement dans le temps) de deux phénomènes $Y(1)$ et $Y(2)$, mais que ceux-ci n'ont pas atteint nécessairement la même valeur finale: $K(1) \neq K(2)$. Dans ces cas, une transformation en pourcentage de $K(1)$ et $K(2)$ permet de contourner cet inconvénient. Si la transformation en pourcentage est demandée, LOG4 fera alors appel au sous-programme PCENT. Le calcul des coefficients de l'expression $\phi(t)$, (équation 5), est exécuté par les sous-programmes REG (régression polynomiale) et INV (inversion de matrice). Les résidus de l'équation 4 sont donnés par le sous-programme RES.

b) ordinogramme et formats de sortie

Ce paragraphe a pour but d'explicitier la logique interne du programme. Dans la discussion qui va suivre, on se référera à l'ordinogramme exposé à l'Appendice A. On désignera par $\{Y(1)\}$ l'ensemble des valeurs Y_i observées dans le premier ensemble, tandis que $\{\hat{Y}(1)\}$, $\{Y'(1)\}$ et $\{Y''(1)\}$ représenteront respectivement: l'ensemble des valeurs Y_i prédites (i.e. calculées depuis les relations 5 et 6), l'ensemble des valeurs de la dérivée première et l'ensemble des valeurs de la dérivée seconde par rapport à t . Les différentes adresses qui

apparaîtront dans le texte (1 ... 25) réfèrent à l'ordinogramme. Comme nous l'avons vu plus haut, LOG4 fonctionne en mode conversationnel: les entrées et les sorties (impression des résultats) prennent au maximum onze pages.

Au départ 1, le sous-programme INFOR demande 2 les valeurs de A, N, {t} et {Y(1)}, c'est-à-dire:

A = nombre de courbes à définir:

A = 1 signifie une courbe
A = 2 signifie deux courbes

N = nombre de points par courbe

{t} = ensemble des valeurs de la variable indépendante, tel que $\rho\{t\} \equiv N$, où $\rho\{t\}$ signifie la dimension de l'ensemble {t}.

{Y(1)} = ensemble des valeurs des ordonnées des points appartenant au premier ensemble. Chaque point de la courbe est donc défini par un couple {t; Y(1)}, et chaque couple définit un point et un seul.

Si A = 1, il y aura impression 3 d'un premier tableau (tableau 1), en l'occurrence un tableau de N lignes par deux colonnes, respectivement {t} et {Y(1)}.

Si A = 2, le sous-programme INFOR demandera 4 de lire {Y(2)} et imprimera 5 le tableau 2, soit N lignes par trois colonnes qui sont en l'occurrence {t}, {Y(1)} et {Y(2)}. Ensuite le sous-programme INFOR lira la valeur de B, c'est-à-dire que si B = 0, il n'y aura aucune transformation exécutée sur {Y}, par contre, si B = 1, le sous-programme PCENT transformera {Y(1)} et éventuellement {Y(2)} en pourcentage de K(1) et K(2) respectivement.

Si $B = 0$, 6, le programme continuera en séquence 10 et si $B = 1$, 7.
 le sous-programme PCENT exécutera la transformation 8 et imprimera 9
 le tableau 3, de N lignes par 2 ou 3 colonnes: $\{t\}$, $\{Y(1)\}$ et $\{Y(2)\}$,
 selon le cas. Dans ce tableau $\{Y(1)\}$ et $\{Y(2)\}$ seront exprimés en
 pourcentage de $K(1)$ et $K(2)$. A ce stade 10, commence le calcul de la
 première courbe, soit $\hat{Y}(1)$.

Pour pouvoir estimer les différents coefficients $\hat{\beta}_i$ de l'é-
 quation polynomiale $\phi(t)$, équation (5), on transformera 11 d'abord les
 valeurs Y_i en valeurs Λ_i , Λ_i étant le logit correspondant à un Y_i ob-
 servé; puis le sous-programme REG calculera 12 les différents paramè-
 tres de la régression polynomiale et imprimera 13 le tableau 4: 8 li-
 gnes par 5 colonnes, c'est-à-dire:

5	$\hat{\beta}_0$	0	0	0
1	$\hat{\beta}_1$	écart-type de $\hat{\beta}_1$	t_1	0
2	$\hat{\beta}_2$	écart-type de $\hat{\beta}_2$	t_2	0
3	$\hat{\beta}_3$	écart-type de $\hat{\beta}_3$	t_3	0
4	$\hat{\beta}_4$	écart-type de $\hat{\beta}_4$	t_4	0
0	d.l. (régression)	S.C.	C.M.	F
0	d.l. (erreur)	S.C.	C.M.	0
0	d.l. (total)	S.C.	écart-type de l'estimée	R carré

Dans ce tableau, les $\hat{\beta}_i$ représentent les coefficients du polynôme $\phi(t)$,
 les valeurs t_i sont celles des variables "t" de Student calculées pour

les différents $\hat{\beta}_i$.

Les lignes 6, 7 et 8 donnent l'analyse de variance de la régression: d.l. représente les degrés de liberté, S.C. les sommes des carrés, C.M. les carrés moyens, F la variable "F" de Snedecor et R carré le coefficient de détermination.

Ensuite, le sous-programme RES calculera 14 les résidus de l'équation $\phi(t)$ et imprimera 15 le tableau 5 ayant N-1 lignes et 4 colonnes respectivement $\{t\}$, $\{\Lambda_i\}$, $\{\hat{\phi}(t)\}$ et $\{\epsilon\}$. La colonne $\{\epsilon\}$ représentent les résidus des Λ_i .

Ensuite, le programme LOG4 calculera 16 les différentes valeurs des équations $\hat{y}(t)$, $y'(t)$ et $y''(t)$ et les imprimera 19 au tableau 6. Ce tableau aura N lignes par 6 colonnes, respectivement $\{t\}$, $\{Y(1)\}$, $\{\hat{Y}(1)\}$, $\{\epsilon Y\}$, $\{Y'(1)\}$ et $\{Y''(1)\}$. Une fois ce tableau imprimé, si A = 1, le programme s'arrêtera 20.

Dans le cas où A = 2, le programme retournera à l'adresse 11 après avoir remplacé 21 les valeurs expérimentales Y(1) par $\{Y(2)\}$. Le programme continuera alors selon les mêmes séquences que celles décrites ci-dessus et il y aura impression des tableaux 7, 8 et 9, analogues aux tableaux 4, 5 et 6. Après cela, LOG4 comparera les valeurs différentes des ensembles $\{1\}$ et $\{2\}$. Tout d'abord, un tableau 10 permettra de faire la comparaison entre les deux fonctions primitives estimées. Ce tableau comporte N lignes et 4 colonnes, respectivement $\{t\}$, $\{\hat{Y}(1)\}$, $\{\hat{Y}(2)\}$ et $\{\epsilon\hat{Y}\}$. Cette dernière colonne représentant les différences observées (résidus) entre les deux fonctions primitives. Venant ensuite, le tableau 11 représente également des comparaisons, mais entre les dérivées

premières cette fois. Les 4 colonnes représentent respectivement $\{t\}$, $\{Y'(1)\}$, $\{Y'(2)\}$ et $\{\epsilon Y'\}$.

Quant au dernier tableau, le tableau 12, il se lit comme suit: $\{t\}$, $\{Y''(1)\}$, $\{Y''(2)\}$ et $\{\epsilon Y''\}$. Il nous donne donc une comparaison entre les dérivées secondes des deux fonctions primitives. Ce tableau était imprimé, le programme LOG4 s'arrête 25.

Lorsque deux ensembles ont été analysés ($A = 2$), les tableaux 10, 11 et 12 permettent une comparaison globale qui peut se représenter graphiquement avec beaucoup de précision (figure 4). Néanmoins, il n'est pas forcément nécessaire de recourir à de telles constructions graphiques: un examen des quatrièmes colonnes de ces trois tableaux fournit suffisamment d'information sur la manière dont les deux ensembles évoluent l'un vis à vis de l'autre.

ANALYSE D'UN CAS CONCRET

a) processus expérimental

Afin d'illustrer l'intérêt que peut présenter ce programme sur le plan biologique, nous avons choisi de l'appliquer à un ensemble de données tirées d'une expérience visant à déterminer l'influence de différentes solutions nutritives sur la croissance de semis de résineux (*Picea mariana* (Mill.) B.S.P.).

Parmi celles-ci, deux solutions, A et B, furent choisies, d'abord parce que chimiquement très différentes entr'elles et ensuite parce qu'ayant

induit à la fin de l'expérience une même hauteur moyenne. En gros, le protocole expérimental fut le suivant: deux groupes de cinquante semis âgés de 9 semaines et ayant sensiblement à cet âge la même hauteur moyenne avec un même écart-type, reçurent, l'un une dose de la solution A, l'autre une dose de la solution B. Ces deux groupes furent mesurés (mm) chaque semaine et leur hauteur moyenne portée sur un graphique (figure 1). L'expérience se termina lorsque les semis furent âgés de trente-quatre semaines. Ces mêmes résultats, mais exprimés en % de leur valeur maximum K sont reportés à la figure 2.

Les tableaux 2 et 3 (Appendice C) donnent les valeurs exactes de ces résultats expérimentaux. Au tableau 2 on trouvera les valeurs réelles des hauteurs à différentes époques (par intervalle d'une semaine) tandis qu'au tableau 3 ces hauteurs seront converties en % de la hauteur atteinte à la fin de l'expérience, soit trente-quatre semaines. Les tableaux 4 et 7 donnent les différents paramètres des équations de régressions $\phi(t)$ ainsi que l'analyse de variance de ces régressions: dans les deux cas, le coefficient de détermination R^2 est hautement significatif (seuil 99.9%). Ceci explique le très bon ajustement observé aux tableaux 5 et 8.

Ces tableaux donnent pour les différentes valeurs de "t" (variable indépendante) la valeur des logits calculés et des logits prédits au moyen de l'équation de régression (5). On retrouve d'ailleurs ces valeurs à la figure 3. La dernière colonne des tableaux 5 et 8 est constituée des résidus, c'est-à-dire des différences entre les Λ_i et les $\hat{\Lambda}_i$. Etant donné que la courbe polynomiale

$$\hat{\Lambda}_i = \phi(t) \quad (4')$$

fournit un excellent ajustement, la somme des carrés des résidus est très faible, soit 0.378 et 0.008 pour les deux ensembles.

Les tableaux 6 et 9 donnent pour les deux ensembles respectivement les hauteurs observées, les hauteurs calculées, les résidus des hauteurs, les valeurs de la dérivée première et les valeurs de la dérivée seconde. Dans ce cas également, la somme des carrés des résidus est très faible ce qui confirme la valeur de l'ajustement.

b) analyse numérique des résultats

1° Dans le cas d'une seule courbe: considérons tout d'abord sur la figure 4a, la courbe $\hat{Y}(2)$. Il s'agit de semis ayant reçu une dose de la solution B. La croissance assez lente au départ s'accroît aux environs de la 22^{ième} semaine et atteint son niveau de saturation vers la 32^{ième} semaine.

Tout ceci peut d'ailleurs être réduit approximativement d'un simple examen des figures 1 et 2. Néanmoins, si nous désirons pousser plus loin notre investigation, il sera nécessaire de faire appel au calcul des dérivées. En effet, $\hat{Y}(t)$ représente la hauteur moyenne atteinte par le semis au temps t , tandis que $\hat{Y}(t + \Delta t)$ représente sa hauteur après un temps Δt .

L'expression:

$$\frac{\hat{Y}(t + \Delta t) - \hat{Y}(t)}{\Delta t} \quad (11)$$

représente donc l'accroissement en hauteur de ces semis durant l'in-

tervalle Δt , et, d'une façon plus générale, $\frac{\partial}{\partial t} y(t)$ peut être considéré comme une *vitesse de croissance*. En effet, elle est exprimée en mm par unité de temps (LT^{-1}), qui sont les unités physiques d'une vitesse. La courbe $Y'(2)$ est donc bien la courbe de *vitesse de croissance* moyenne du groupe B. Cette courbe part - théoriquement - de zéro et croit jusqu'à un point b' où la vitesse est maximum. La détermination précise de ce point maximum peut se faire en annulant la dérivée seconde correspondante, $Y''(2)$, soit $t = b''$ (figure 4c).

A gauche et à droite de ce point b' , il existe sur la courbe $Y'(2)$ deux points d'inflexion a' et c' . Ces deux points correspondent respectivement à un maximum (a'') et à un minimum (c'') de la dérivée seconde. Cette dérivée seconde, $Y''(2)$, a comme unités LT^{-2} , qui sont physiquement les unités d'une accélération, en l'occurrence une *accélération de la croissance*. Cette accélération $Y''(2)$ sera tantôt positive tantôt négative suivant que la *vitesse de croissance* $Y'(2)$, augmente ou diminue. Cette accélération sera nulle au point b'' ($t \cong 25$ semaines) et on remarquera que $t_b = t_{b'} = t_{b''}$.

En étudiant, non plus une courbe $\hat{Y}(t)$, mais bien l'ensemble des trois courbes (i.e. fonction primitive, dérivée première et seconde), il est possible de beaucoup mieux interpréter le phénomène observé et de définir en termes de "t" (variable indépendante) ses points remarquables et ses tendances. Certes cette interprétation ne peut en aucune manière se substituer à l'interprétation strictement biologique qui devra être faite, mais néanmoins, elle lui offre un support plus solide.

Par exemple, dans le cas présent, il serait possible d'affirmer que l'action de la solution B a été la plus marquée aux environs de la vingt-deuxième semaine (t_a), ce qui correspond à une accélération maximum de croissance ($t_a = t_{a''}$).

Si comme c'est le cas ici, il y a deux ensembles de points à envisager, on peut recommencer la même interprétation pour l'autre ensemble, *mutatis mutandis*. D'autre part, l'étude de ces fonctions dérivées nous permet d'exécuter alors une comparaison complète des deux ensembles.

2° Dans le cas de deux courbes: lorsqu'il s'agit de comparer plusieurs courbes de croissance, la méthode graphique est loin d'offrir toute la précision voulue! Dans ces cas, on pourra se référer aux tableaux 10, 11 et 12 (Appendice C) dont les dernières colonnes donnent les valeurs exactes des différences entre les fonctions primitives, les dérivées premières et les dérivées secondes.

CONCLUSIONS

La fonction logistique "modifiée" que nous avons discutée ici offre une très grande souplesse dans l'ajustement aux points expérimentaux. Le programme LOG4 est d'un emploi excessivement simple et a été conçu de telle sorte qu'il puisse être à la portée d'un nombre maximum d'utilisateurs. Certes dans quelques cas il n'offrira pas le même degré de précision qu'on pourrait attendre d'un programme basé

par exemple sur un processus d'itération⁽¹⁾, néanmoins ses autres qualités, vitesse d'exécution et facilité d'usage nous semblent justifier son emploi dans la majorité des situations.

Il apparaît clairement en outre que seul le calcul des dérivées permet une analyse rationnelle de la fonction primitive et offre un outil extrêmement précieux dans l'interprétation des résultats.

De plus, le fait que le programme offre la possibilité de comparer deux ensembles de points introduits simultanément permet un gain de temps parfois considérable dans certains travaux.

(1) En effet, il apparaît que pour être applicable, le programme LOG4 doit rencontrer les deux hypothèses de départ suivantes: tout d'abord, il faut sssumer que la distribution des résultats expérimentaux est bien de nature sigmoïdale *lato sensu*, cela d'ailleurs va de soi, ensuite il est très important de vérifier si la valeur K peut être considérée en première approximation comme une estimation valable de l'asymptote horizontale supérieure; si tel n'était absolument pas le cas, il conviendrait alors de considérer les résultats obtenus par LOG4 avec la plus grande attention!

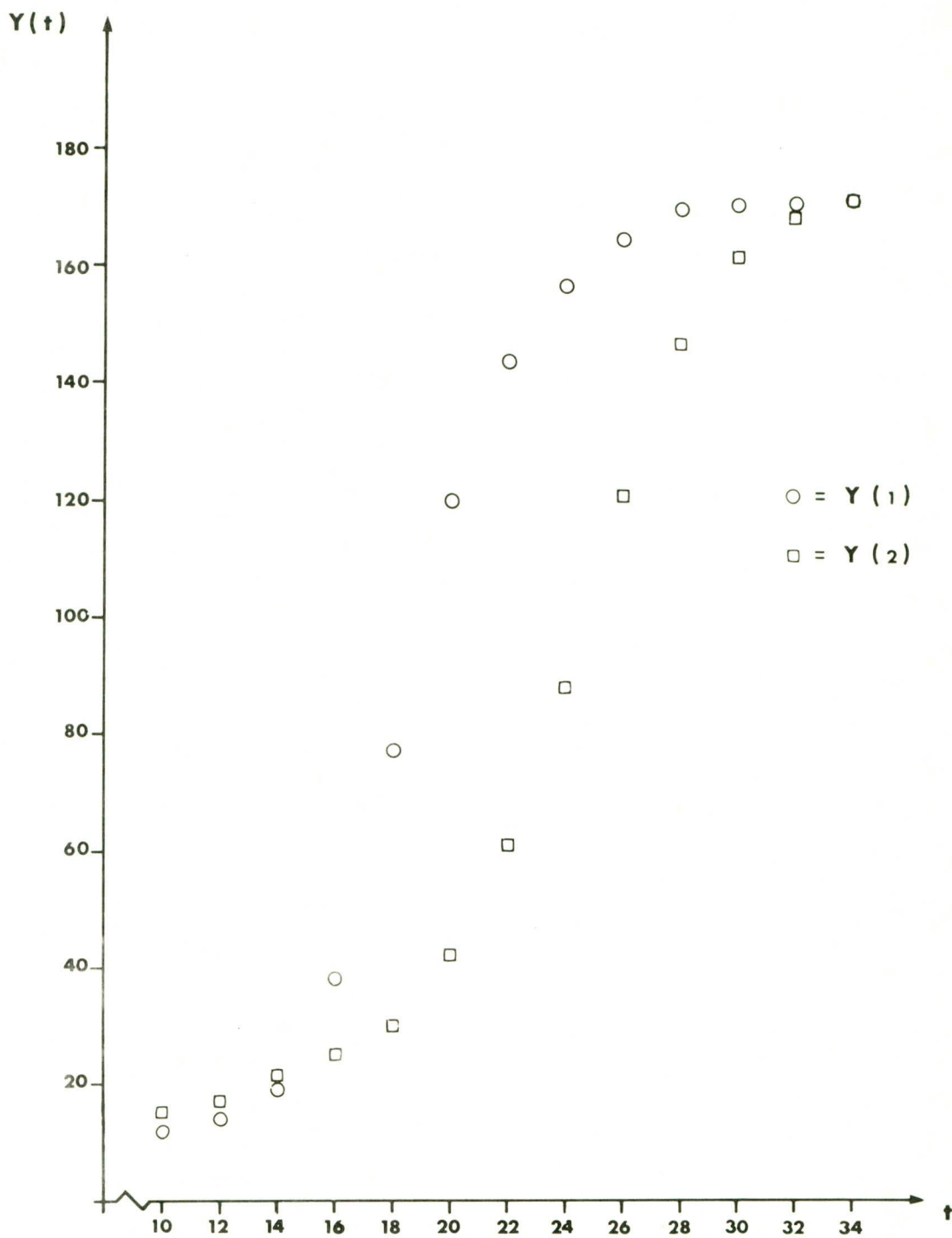


Fig. 1. Valeurs expérimentales $\{Y(1)\}$ et $\{Y(2)\}$

En abscisse, t représente des semaines.

Chaque point représente la hauteur moyenne

d'un groupe de cinquante semis.

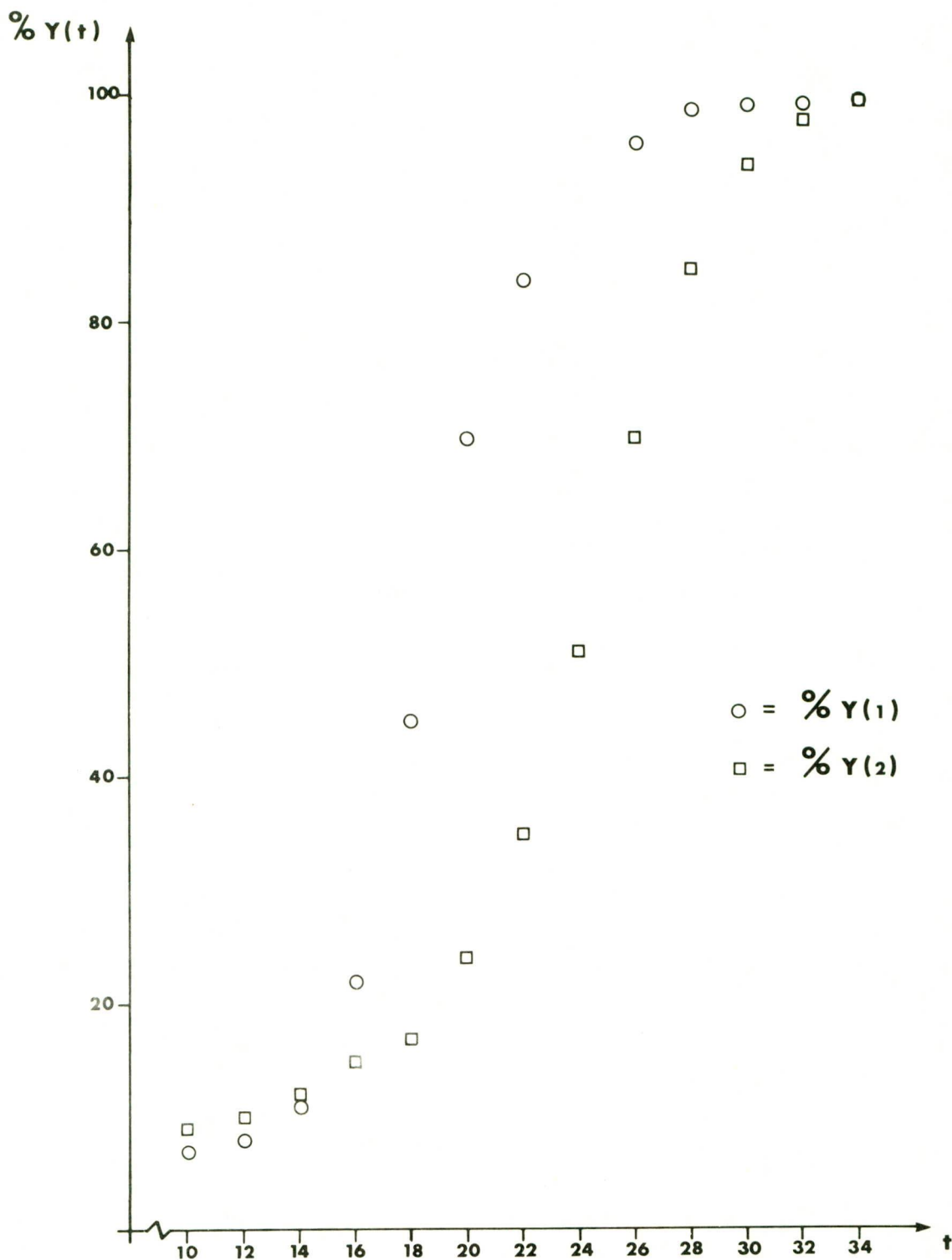


Fig. 2. Mêmes valeurs qu'à la figure 1, mais exprimées en % de $K(1)$ et $K(2)$, qui dans ce cas-ci sont égaux (les valeurs expérimentales ont été arrondies à l'unité).

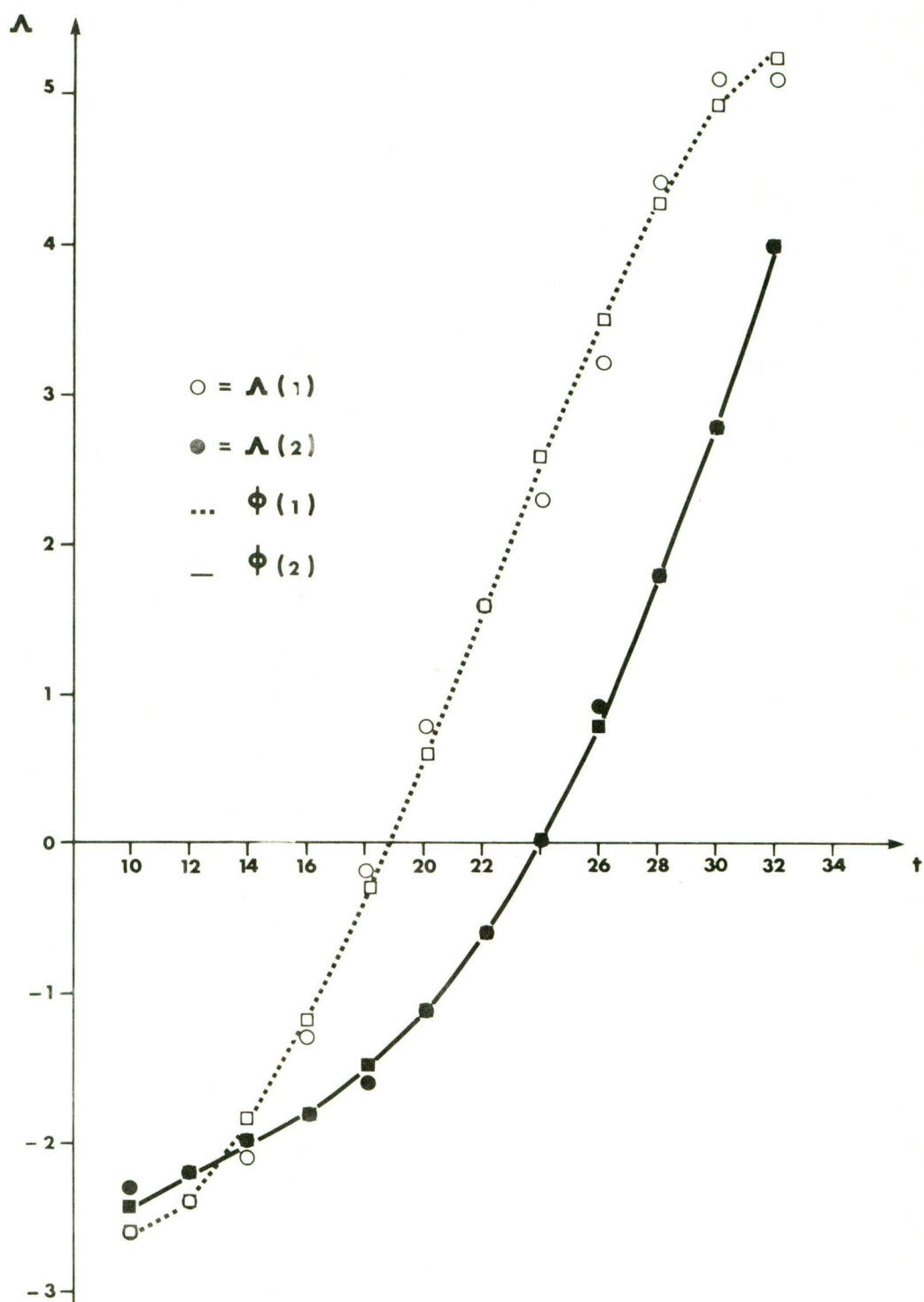


Fig. 3. Valeurs des logits λ_i calculés depuis les $\{Y(1)\}$ et $\{Y(2)\}$.

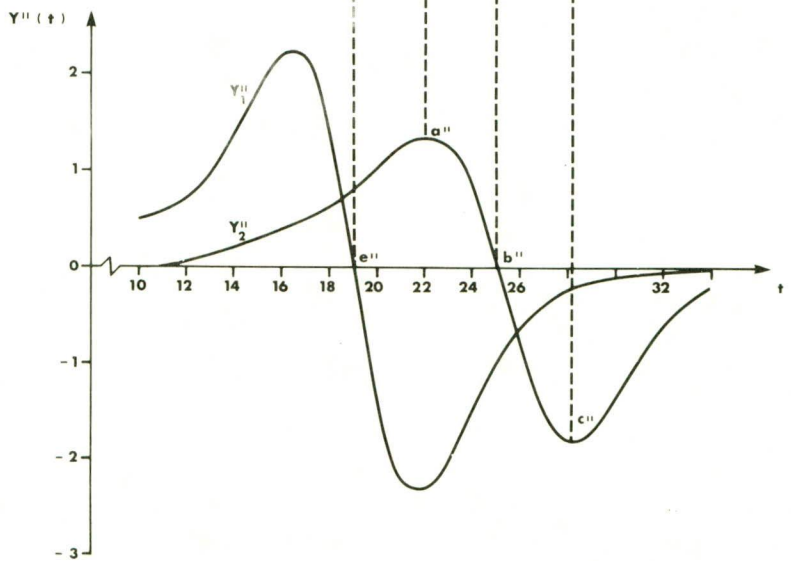
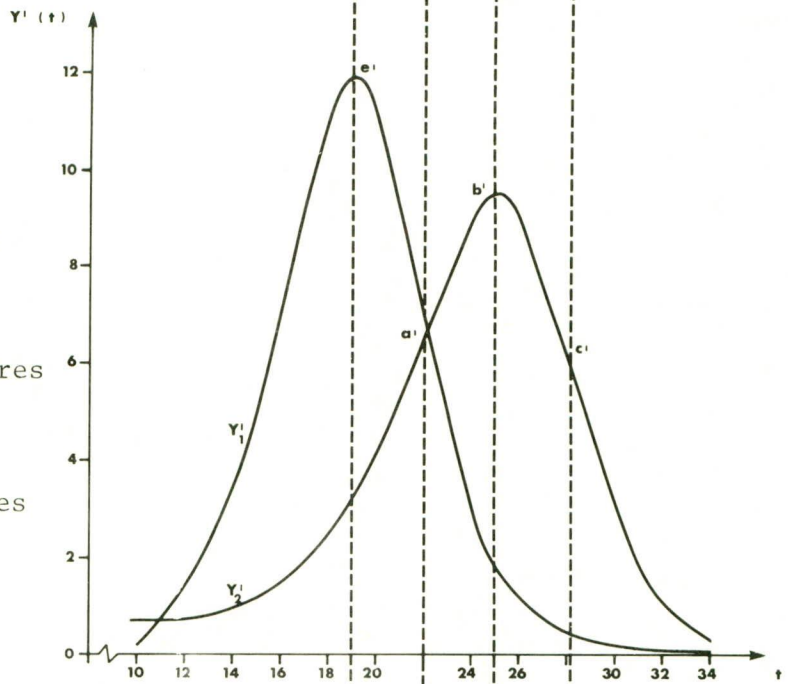
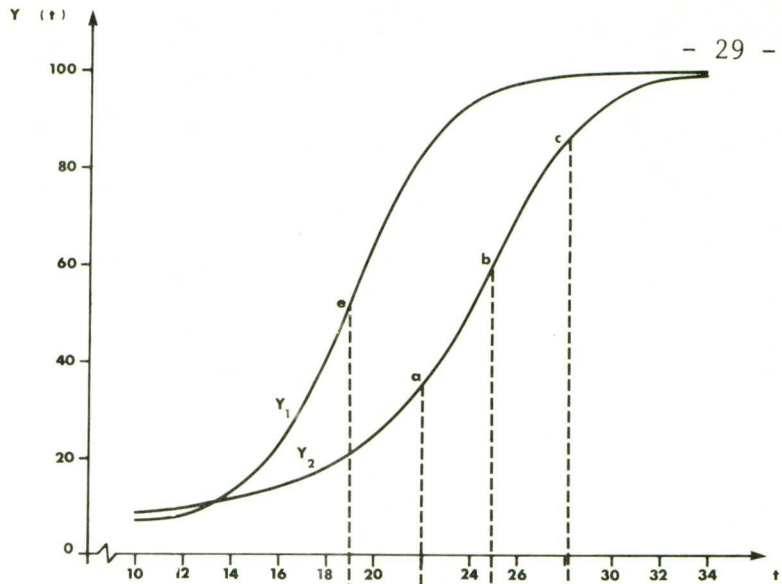
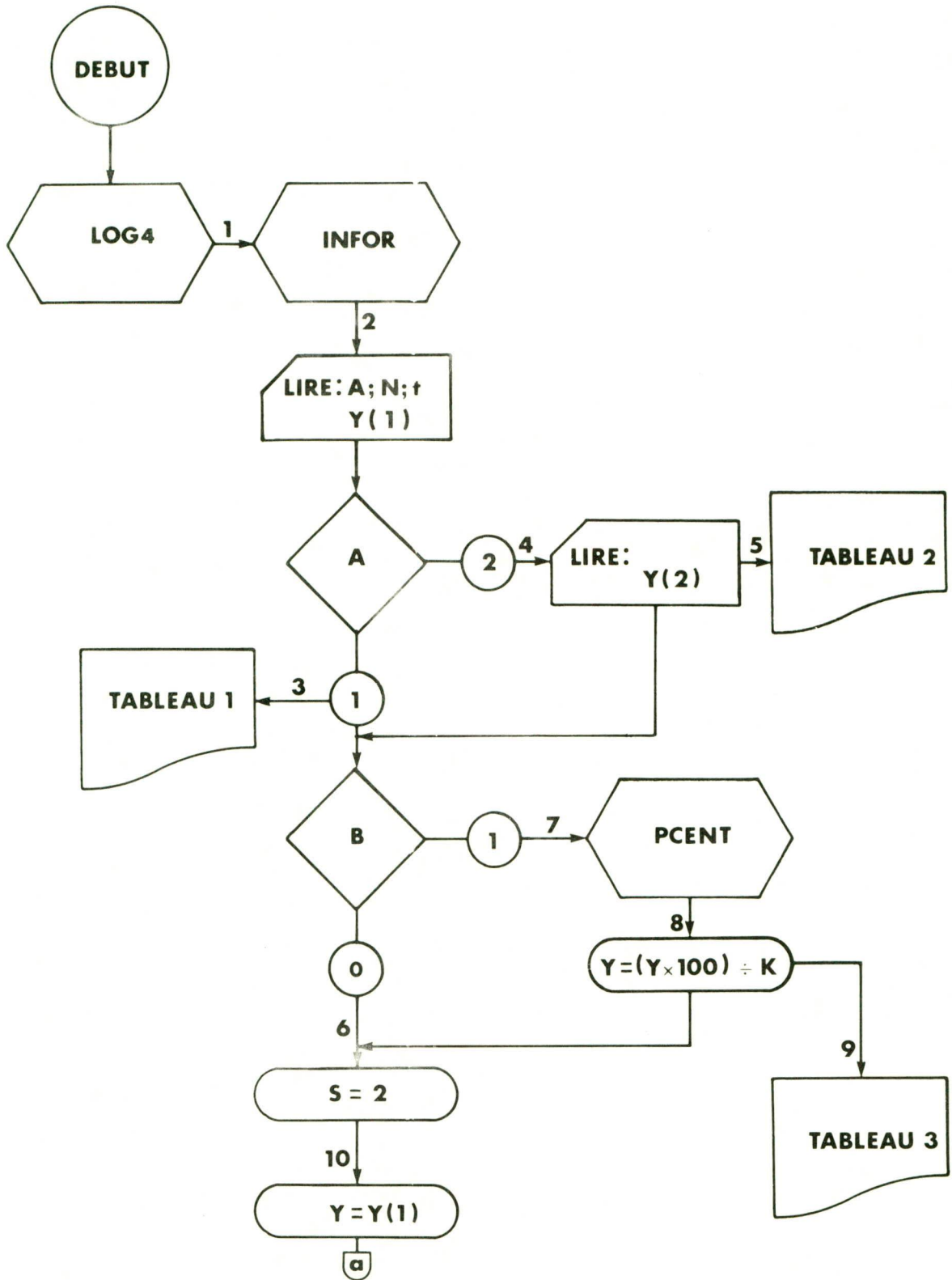
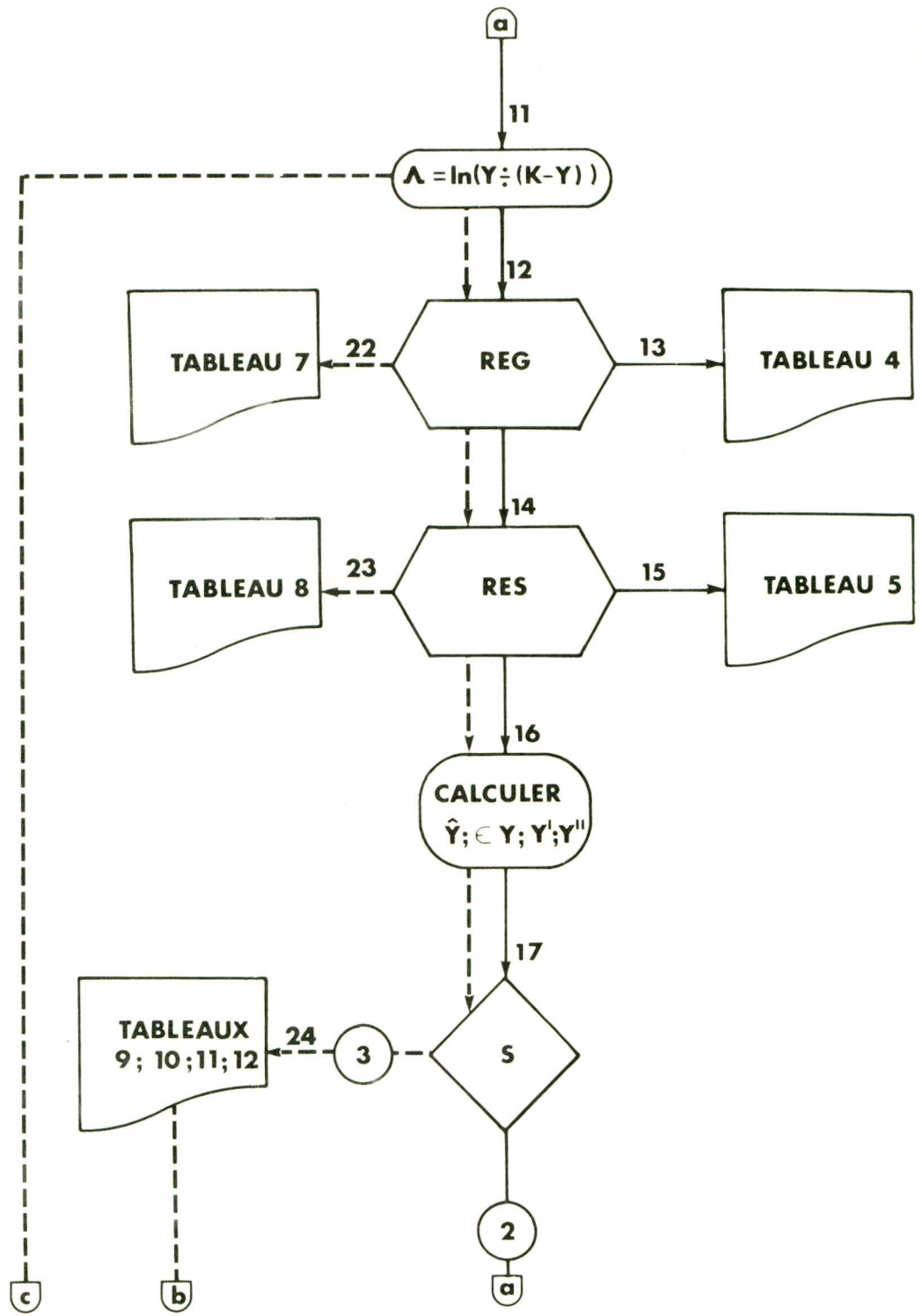
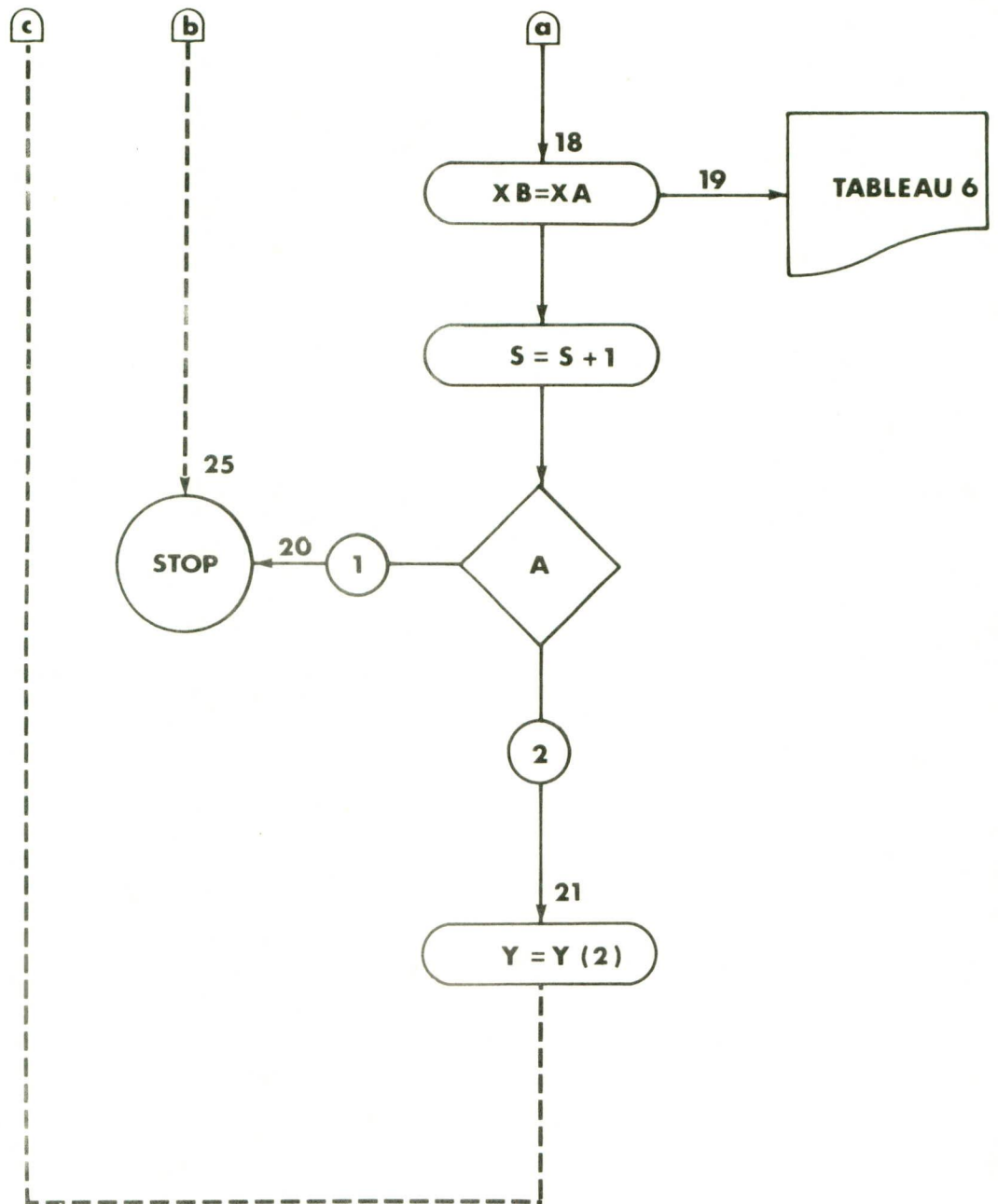


Fig. 4. De haut en bas,
les figures 4a, 4b et 4c.
Fig. 4a: courbes $\hat{Y}(1)$ et
 $\hat{Y}(2)$, en %.
Fig. 4b: dérivées premières
 $y'_1(t)$ et $y'_2(t)$.
Fig. 4c: dérivées secondes
 $y''_1(t)$ et $y''_2(t)$.







Appendice B

Liste d'instructions des sous-programmes INFOR,
PCENT, REG, RES, INV et du programme LOG4.

VINFOR[]V

▽ INFOR

```
[1] 'NOMBRE DE COURBES :'  
[2] A←[ ]  
[3] ' '  
[4] 'NOMBRE DE VALEURS OBSERVEES :'  
[5] N←[ ]  
[6] ' '  
[7] R←((N),3)ρ0  
[8] 'ENTRER LES VALEURS X :'  
[9] R[;1]←[ ]  
[10] ' '  
[11] 'ENTRER LES VALEURS Y :'  
[12] R[;2]←[ ]  
[13] ' '  
[14] →(A=1)/IA  
[15] 'ENTRER LES SECONDES VAL. Y :'  
[16] R[;3]←[ ]  
[17] ' '  
[18] 'I3,2F11.5' ΔFMT R  
[19] ' '  
[20] →IB  
[21] IA:'I3,F11.5' ΔFMT R  
[22] ' '  
[23] IB:'TRANSFORMATION EN o/o :'  
[24] B←[ ]
```

▽

∇PCENT[]∇

```
∇ PCENT
[1]  R[;2]←100×R[;2]÷R[(N);2]
[2]  →(A=1)/PA
[3]  R[;3]←100×R[;3]÷R[(N);3]
[4]  'I3,2F11.5' ΔFMT R
[5]  →PB
[6]  PA:'I3,F11.5' ΔFMT R
[7]  PB:' '
∇
```


VREG[]V

▽ T←V REG X;N;U;M;R;Q;S;B
[1] T←((3+ρV),5)ρ0
[2] T[1ρV;1]←V[ρV],⁻1+V
[3] Q←(QR)+.×R+U-(ρU)ρM←(+/[1] U←X[;V])÷N←(ρX)[1]
[4] T[3+ρV;3]←Q[ρV;ρV]
[5] Q←⁻1 0 +Q
[6] B←(-M+.×B,⁻1),B←,(S+INV 0⁻1 +Q)+.×Q[;ρV]
[7] T[1ρV;2]←B
[8] T[1+ρV;3]←(1+B)+.×Q[;ρV]
[9] T[(ρV)+13;2]←((ρV)-1),(N-ρV),N-1
[10] T[2+ρV;3]←-/T[(3+ρV),1+ρV;3]
[11] T[(ρV)+12;4]←T[(ρV)+12;3]÷T[(ρV)+12;2]
[12] T[1+1(ρV)-1;3]←(T[2+ρV;4]×+/S×(1(ρV)-1)○.=1(ρV)-1)*0.5
[13] T[1+1(ρV)-1;4]←T[1+1(ρV)-1;2]÷T[1+1(ρV)-1;3]
[14] T[3+ρV;4]←T[2+ρV;4]*0.5
[15] T[1+ρV;5]←T[1+ρV;4]÷T[2+ρV;4]
[16] T[3+ρV;5]←T[1+ρV;3]÷T[3+ρV;3]

▽

▽RES[]▽

▽ R←T RES X;V;O;E
[1] V←T[1+i⁻⁴+ρT[;1];1]
[2] R←(O←X[;T[1;1]])-E←T[1;2]+T[1+iρV;2]+.×QX[;V]
[3] R←Q(4,(ρX)[1])ρ(i(ρX)[1]),O,E,R
▽

VINV[[]]V

```
∇ RB←INV RA;RK;RS;RP;RI
[1] →((2=ρρRA)∧=/ρRA)ρ4
[2] 'NO INVERSE!'
[3] →~RB←1
[4] RK←L/ρRA
[5] RS←RK
[6] RP←ιRK
[7] RA←RA[;(ιRS),1]
[8] RA[;1+RS]←(ιRS)≤1
[9] RI←(|RA[ιRK;1])ι[|RA[ιRK;1]
[10] RP[1,RI]←RP[RI,1]
[11] RA[1,RI;ιRS]←RA[RI,1;ιRS]
[12] →(1E-30>|RA[1;1])ρ2
[13] RA[1;]←RA[1;]÷RA[1;1]
[14] RA←RA-((~(ιRS)≤1)×RA[;1])°.×RA[1;]
[15] RA←RA[1+RS|ιRS;(1+ιRS),1]
[16] RP←RP[1+RS|ιRS]
[17] →(0<RK←RK-1)/8
[18] RB←RA[;RPιιRS]
```

∇

∇LOG4[[]]∇

∇ LOG4

```

[1] PAGE←[]
[2] INFOR
[3] →(B=0)/QA
[4] PAGE←[]
[5] PCENT
[6] QA:PAGE←[]
[7] XA←((N),6)ρ0
[8] XA[;2]←R[;2]
[9] Y←((N),5)ρ0
[10] Y[;1]←XA[;1]←R[;1]
[11] Y[;2]←Y[;1]*2
[12] Y[;3]←Y[;1]*3
[13] Y[;4]←Y[;1]*4
[14] K+2
[15] I+1
[16] QB:Y[I;5]←R[I;K]÷R[N;K]-R[I;K]
[17] I+I+1
[18] →(I≤(N-1))/QB
[19] POLY←((N-1),5)ρ,Y
[20] V←15
[21] 'I3,4F11.5' ΔFMT T←V REG POLY
[22] ' '
[23] PAGE←[]
[24] G←T[;2]
[25] 'I3,3F11.5' ΔFMT U←T RES POLY
[26] PAGE←[]
[27] PHI←G[1]+(G[2]×Y[;1])+(G[3]×Y[;2])+(G[4]×Y[;3])+(G[5]×Y[;4])
[28] THETA←G[2]+(2×G[3]×Y[;1])+(3×G[4]×Y[;2])+(4×G[5]×Y[;3])
[29] RHO←(2×G[3])+(6×G[4]×Y[;1])+(12×G[5]×Y[;2])
[30] EX←1÷(*PHI)
[31] KEX←THETA×EX÷(1+EX)
[32] XA[;3]←XA[N;2]÷(1+EX)
[33] XA[;4]←XA[;2]-XA[;3]
[34] XA[;5]←XA[;3]×KEX
[35] XA[;6]←(XA[;5])×((2×KEX)+(RHO÷THETA)-(THETA))
[36] →(K=3)/QC
[37] XB←XA
[38] 'I3,5F11.5' ΔFMT XB
[39] PAGE←[]
[40] ' '
[41] QC:K←K+1
[42] →(K≠3)/QD
[43] →(A=1)/QE
[44] XA[;2]←R[;3]
[45] I←1

```

[46] →QB
[47] QD: 'I3,5F11.5' ΔFMT XA
[48] PAGE←
[49] 'I3,3F11.5' ΔFMT(XA[;1];XB[;3];XA[;3];XB[;3]-XA[;3])
[50] PAGE←
[51] 'I3,3F11.5' ΔFMT(XA[;1];XB[;5];XA[;5];XB[;5]-XA[;5])
[52] PAGE←
[53] 'I3,3F11.5' ΔFMT(XA[;1];XB[;6];XA[;6];XB[;6]-XA[;6])
[54] QE: ' '

▽

Appendice C

Analyse d'un cas concret: tableaux des résultats.

Tableaux 2 à 12: le tableau 1 n'existe pas étant donné qu'il y a deux ensembles de résultats.

TABLEAU DEUX

NOMBRE DE COURBES :

:
2

NOMBRE DE VALEURS OBSERVEES :

:
13

ENTRER LES VALEURS X :

:
10 12 14 16 18 20 22 24 26 28 30 32 34

ENTRER LES VALEURS Y :

:
12 14 19 38 77 120 144 157 165 170 171 171 172

ENTRER LES SECONDES VAL. Y :

:
15 17 21 25 30 42 61 88 121 147 162 169 172

10	12.00000	15.00000
12	14.00000	17.00000
14	19.00000	21.00000
16	38.00000	25.00000
18	77.00000	30.00000
20	120.00000	42.00000
22	144.00000	61.00000
24	157.00000	88.00000
26	165.00000	121.00000
28	170.00000	147.00000
30	171.00000	162.00000
32	171.00000	169.00000
34	172.00000	172.00000

TRANSFORMATION EN O/O :

:
1

TABLEAU TROIS

10	6.97674	8.72093
12	8.13953	9.88372
14	11.04651	12.20930
16	22.09302	14.53488
18	44.76744	17.44186
20	69.76744	24.41860
22	83.72093	35.46512
24	91.27907	51.16279
26	95.93023	70.34884
28	98.83721	85.46512
30	99.41860	94.18605
32	99.41860	98.25581
34	100.00000	100.00000

TABLEAU QUATRE

5	2.94637	0.00000	0.00000	0.00000
1	-1.30506	1.55370	-0.83997	0.00000
2	0.09138	0.12082	0.75630	0.00000
3	-0.00175	0.00399	-0.43833	0.00000
4	0.00001	0.00005	0.15700	0.00000
0	4.00000	93.02334	23.25584	430.49928
0	7.00000	0.37814	0.05402	0.00000
0	11.00000	93.40149	0.23242	0.99595

TABLEAU CINQ

1	-2.59027	-2.63935	0.04909
2	-2.42354	-2.42105	-0.00249
3	-2.08600	-1.92333	-0.16267
4	-1.26025	-1.21150	-0.04876
5	-0.21007	-0.34802	0.13795
6	0.83625	0.60750	0.22875
7	1.63761	1.59830	0.03931
8	2.34820	2.57048	-0.22229
9	3.16004	3.47301	-0.31297
10	4.44265	4.25769	0.18496
11	5.14166	4.87919	0.26248
12	5.14166	5.29502	-0.15336

TABLEAU SIX

10	6.97674	6.66483	0.31192	0.17450	0.54444
12	8.13953	8.15815	-0.01861	1.38290	0.73655
14	11.04651	12.74911	-1.70259	3.42291	1.37946
16	22.09302	22.94366	-0.85064	7.05261	2.19160
18	44.76744	41.38631	3.38114	11.14890	1.43945
20	69.76744	64.73694	5.03050	11.21297	-1.42331
22	83.72093	83.17800	0.54293	6.92780	-2.34201
24	91.27907	92.89375	-1.61468	3.12116	-1.38134
26	95.93023	96.99100	-1.06076	1.24239	-0.58321
28	98.83721	98.60426	0.23295	0.48885	-0.22502
30	99.41860	99.24542	0.17319	0.19677	-0.08947
32	99.41860	99.50085	-0.08225	0.07439	-0.04152
34	100.00000	99.57879	0.42121	0.00743	-0.02973

TABLEAU SEPT

5	-5.14703	0.00000	0.00000	0.00000
1	0.60960	0.23227	2.62455	0.00000
2	-0.04929	0.01806	-2.72860	0.00000
3	0.00181	0.00060	3.03676	0.00000
4	-0.00002	0.00001	-2.58421	0.00000
0	4.00000	49.00407	12.251021014	7.54229
0	7.00000	0.00845	0.00121	0.00000
0	11.00000	49.01252	0.03475	0.99983

TABLEAU HUIT

1	-2.34820	-2.35274	0.00455
2	-2.21021	-2.18102	-0.02919
3	-1.97276	-2.00940	0.03665
4	-1.77156	-1.79669	0.02513
5	-1.55463	-1.50869	-0.04594
6	-1.12986	-1.11825	-0.01162
7	-0.59866	-0.60522	0.00656
8	0.04652	0.04351	0.00301
9	0.86396	0.83403	0.02994
10	1.77156	1.76540	0.00616
11	2.78501	2.82967	-0.04466
12	4.03129	4.01187	0.01942

TABLEAU NEUF

10	8.72093	8.68480	0.03613	0.74259	-0.03723
12	9.88372	10.14681	-0.26309	0.74860	0.05007
14	12.20930	11.82192	0.38738	0.96884	0.17739
16	14.53488	14.22545	0.30943	1.49621	0.36271
18	17.44186	18.11327	-0.67141	2.48649	0.64788
20	24.41860	24.63364	-0.21504	4.16763	1.04631
22	35.46512	35.31506	0.15006	6.61637	1.34114
24	51.16279	51.08755	0.07525	8.98566	0.81914
26	70.34884	69.72056	0.62828	9.09553	-0.79910
28	85.46512	85.38847	0.07664	6.23663	-1.79011
30	94.18605	94.42585	-0.23980	2.96415	-1.32725
32	98.25581	98.22222	0.03359	1.07774	-0.59941
34	100.00000	99.49834	0.50166	0.32868	-0.20584

TABLEAU DIX

10	6.66483	8.68480	-2.01997
12	8.15815	10.14681	-1.98867
14	12.74911	11.82192	0.92718
16	22.94366	14.22545	8.71821
18	41.38631	18.11327	23.27303
20	64.73694	24.63364	40.10330
22	83.17800	35.31506	47.86295
24	92.89375	51.08755	41.80620
26	96.99100	69.72056	27.27044
28	98.60426	85.38847	13.21579
30	99.24542	94.42585	4.81957
32	99.50085	98.22222	1.27863
34	99.57879	99.49834	0.08045

TABLEAU ONZE

10	0.17450	0.74259	-0.56809
12	1.38290	0.74860	0.63430
14	3.42291	0.96884	2.45407
16	7.05261	1.49621	5.55640
18	11.14890	2.48649	8.66241
20	11.21297	4.16763	7.04534
22	6.92780	6.61637	0.31143
24	3.12116	8.98566	-5.86449
26	1.24239	9.09553	-7.85315
28	0.48885	6.23663	-5.74778
30	0.19677	2.96415	-2.76737
32	0.07439	1.07774	-1.00335
34	0.00743	0.32868	-0.32125

TABLEAU DOUZE

10	0.54444	-0.03723	0.58167
12	0.73655	0.05007	0.68647
14	1.37946	0.17739	1.20208
16	2.19160	0.36271	1.82889
18	1.43945	0.64788	0.79157
20	-1.42331	1.04631	-2.46963
22	-2.34201	1.34114	-3.68315
24	-1.38134	0.81914	-2.20048
26	-0.58321	-0.79910	0.21589
28	-0.22502	-1.79011	1.56508
30	-0.08947	-1.32725	1.23778
32	-0.04152	-0.59941	0.55789
34	-0.02973	-0.20584	0.17611

REFERENCES

- ANTLE, C., L. KLIMKO, W. HARKNESS, 1970. Confidence intervals for the parameters of the logistic distribution.
Biometrika 57 (2).
- BERKSON, J., 1957. Tables for the maximum likelihood estimate of the logistic function.
Biometrics 13, pp 28-34.
- BERKSON, J., J.L. HODGES, 1961. A minimax estimator for the logistic function.
Proceedings of the fourth Berkeley Symposium on mathematical statistics and probability, 4 pp 77-86.
University of California Press, Berkeley and Los Angeles.
- GLASS, N.R., 1967. A technique for fitting non-linear models to biological data.
Ecology 48 pp 1010-1013.
- GUPTA, S.S., M. GNANADESIKAN, 1966. Estimation of the parameters of the logistic distribution.
Biometrika 53 pp 565-570.
- HARTER, J.L., A.H. MOORE, 1967. Maximum-likelihood estimation from censored samples, of the parameters of a logistic distribution.
J. Am. Statis. Ass. 62 pp 675-683.
- MARQUARDT, D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters.
J. Soc. Industr. Appl. Math. II(2) pp 431-441.

- MARQUARDT, D.W., 1966. Least-squares estimation of nonlinear parameters IBM Share Library distribution, no. 309401, MLIN.
- PAKIN, S., 1961. APL/360 Reference manual. Science Research Associates Inc. Illinois.
- PATTEN, B.C., 1971. A primer for ecological modeling and simulation with analog and digital computers, in System analysis and simulation in ecology, édité par B.C. Patten. Academic Press, New York et Londres.
- REED, L.J., J. BERKSON, 1929. The application of the logistic function to experimental data. J. Phys. Chem. 33 pp 760-779.
- SPANG, H.A., 1962. A review of minimization techniques for nonlinear functions. J. Soc. Industr. Appl. Math. 4 pp 343-365.
- WILSON, A.L., A.W. DOUGLAS, 1969. A note on nonlinear curve fitting. Am. Statis. pp 37-38.

