

# REVUE DE STATISTIQUE APPLIQUÉE

JACQUES GIROD

## **Ajustement de type logarithmique et prévision des diagrammes de charge**

*Revue de statistique appliquée*, tome 23, n° 1 (1975), p. 83-95

[http://www.numdam.org/item?id=RSA\\_1975\\_\\_23\\_1\\_83\\_0](http://www.numdam.org/item?id=RSA_1975__23_1_83_0)

© Société française de statistique, 1975, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## AJUSTEMENT DE TYPE LOGARITHMIQUE ET PRÉVISION DES DIAGRAMMES DE CHARGE <sup>(1)</sup>

Jacques GIROD

*Pour une étude proposée par EURATOM "Incidence à long terme de l'énergie nucléaire sur le choix des moyens de production d'électricité de la Communauté Economique Européenne"\* , on a dû effectuer une prévision des diagrammes de charge jusqu'à l'an 2000. Pour un terme aussi lointain, les techniques statistiques habituelles se sont avérées inopérantes et ont été remplacées par des tests de cohérence de quelques hypothèses posées a priori. Pour un horizon moins éloigné, 1980 ou 1985, il a été en revanche possible de proposer des méthodes dont les fondements sont mieux assurés. Cet article est un approfondissement théorique de l'une d'entre elles.*

Confronté avec l'étude de séries de deux variables, pour des objectifs d'explication ou de prévision, on commence généralement par éprouver les modèles d'ajustement les plus simples, linéaires et logarithmiques. En cas d'échec, on poursuit avec des modèles plus compliqués dont la signification et la justification, contrairement aux précédents, ne sont pas toujours assurées. Dans le cas présent, pour une raison qui apparaîtra dans la suite, on passera à ce qu'on peut appeler le premier degré de complexité d'un ajustement semi-logarithmique : ajouter une constante à la variable du logarithme, soit l'équation :

$$y = a \cdot \text{Log}(x + b) + c \quad (\text{logarithme népérien}) \quad (1)$$

Dans une première partie on proposera une méthode d'estimation des paramètres, l'introduction de b interdisant le recours aux moindres carrés. On montrera ensuite que l'équation (1) peut s'interpréter à l'aide d'une relation particulière sur la pente de la série. On appliquera enfin les résultats à la prévision des diagrammes de charge de la consommation d'électricité.

-----

(1) Article remis le 22/11/73

\* Modèle énergétiques – Structure de la production d'électricité dans le Marché Commun ; I.R.E.P., Grenoble 1971.

## 1 – METHODE D'ESTIMATION DES PARAMETRES

A partir de l'équation d'ajustement (1), on considère le modèle statistique:

$$y = a \cdot \text{Log}(x + b) + c + e_x \quad (\text{M1})$$

où les résidus  $e_x$  sont supposés aléatoires, gaussiens, de moyennes nulles, sans corrélation entre eux et de lois indépendantes de  $x$ .

Les estimations des paramètres de M1 sont obtenues par une méthode classique d'itérations rendant maximum la vraisemblance de ces paramètres [1]. Le point particulier est cependant le choix des valeurs initiales des itérations. Les propriétés de l'équation d'ajustement et de sa fonction inverse permettent de n'en estimer au départ qu'une seule, en l'occurrence  $a$ .

### 1.1. Paramètres du modèle M1.

Le logarithme de la vraisemblance s'écrit, à une constante près :

$$K = \sum_{i=1}^N (y_i - a \cdot \text{Log}(x_i + b) - c)^2 \quad (2)$$

( $N$  est le nombre d'observations  $x_i, y_i$ )

L'estimateur du maximum de vraisemblance (le vecteur  $v$  de composantes  $\hat{a}, \hat{b}, \hat{c}$ ) est solution de l'équation (ou plutôt du système d'équations non linéaires) :

$$L(\hat{v}) = 0$$

où  $L(v)$  est le vecteur des dérivées premières de  $K$  au point  $v$ .

Si  $v_0$  est une valeur proche de  $\hat{v}$ , on peut écrire :

$$L(v) = L(v_0) - M(v_0) \cdot (\hat{v} - v_0) + \epsilon(\hat{v} - v_0).$$

Sous certaines conditions, on peut trouver un vecteur  $v_1$ , compris entre  $v_0$  et  $\hat{v}$  au sens d'une norme préalablement définie, et satisfaisant l'égalité :

$$v_1 = v_0 + [M(v_0)]^{-1} \cdot L(v_0)$$

et le nombre d'itérations dépend de la rapidité de convergence, c'est-à-dire de la décroissance plus ou moins rapide de la norme  $\|v_{n+1} - v_n\|$  vers 0.

Appliquée au modèle M1 cette procédure conduit à :

$$L(v) = -2 \times \begin{bmatrix} \sum (y - a \cdot \text{Log}(x + b) - c) \cdot \text{Log}(x + b) \\ \sum (y - a \cdot \text{Log}(x + b) - c) \cdot \frac{a}{x + b} \\ \sum (y - a \cdot \text{Log}(x + b) - c) \end{bmatrix}$$

$$M(v) = 2 \times \begin{bmatrix} \Sigma (\text{Log}(x+b))^2 & \Sigma a \cdot \frac{\text{Log}(x+b)}{x+b} & \Sigma \text{Log}(x+b) \\ \Sigma a \cdot \frac{\text{Log}(x+b)}{x+b} & \Sigma \frac{a^2}{(x+b)^2} & \Sigma \frac{a}{x+b} \\ \Sigma \text{Log}(x+b) & \Sigma \frac{a}{x+b} & N \end{bmatrix}$$

A partir de  $M(v)$ , il est aisé de calculer la matrice asymptotique des variances et covariances de  $\hat{a}$ ,  $\hat{b}$ ,  $\hat{c}$ , et mieux, celle de ces trois paramètres complétée par la variance des écarts  $e_x$ , soit  $\sigma^2$ , estimé par :

$$s^2 = \frac{1}{N-3} \Sigma (y - \hat{a} \cdot \text{Log}(x + \hat{b}) - \hat{c})^2$$

Pour ce faire, il faut ajouter au nouveau vecteur  $L'(v)$  une composante supplémentaire :

$$\frac{\partial K}{\partial \sigma^2}$$

On vérifie facilement que la nouvelle matrice d'information s'écrit :

$$M'(v) = \begin{bmatrix} \frac{N}{2 \sigma^4} & 0 \\ 0 & \frac{1}{\sigma^2} \cdot M \end{bmatrix}$$

et la matrice asymptotique des variances – covariances

$$V = \begin{bmatrix} \frac{2 \sigma^4}{N} & 0 \\ 0 & \sigma^2 \cdot M^{-1} \end{bmatrix}$$

On notera en particulier que [2] :

- $s^2$  est sans corrélation avec  $\hat{a}$ ,  $\hat{b}$ ,  $\hat{c}$  ;
- les variances et covariances des paramètres sont proportionnels à  $s^2$  et indépendantes de  $c$ .

En ce qui concerne la fin des itérations, on peut se fixer l'une des deux règles suivantes :

- convenir d'arrêter le processus lorsque deux vecteurs successifs vérifient :

$$\| v_{n+1} - v_n \| \leq \epsilon_1$$

$\epsilon_1$  étant une valeur fixée au départ :

– convenir d'arrêter le processus quand le vecteur des dérivées prend des valeurs proches de zéro :

$$\|L(v_n)\| \leq \epsilon_2$$

En pratique, on constate qu'il suffit généralement de trois ou quatre itérations pour vérifier l'une et/ou l'autre de ces conditions.

## 1.2 Choix de la solution initiale – Le modèle auxiliaire.

Le choix de la solution initiale consiste à trouver un vecteur  $v_0$  tel que le vecteur des dérivées premières au point  $v_0$  soit aussi proche que possible de zéro. Montrons que la détermination d'un tel vecteur peut être limitée à l'estimation préalable d'une seule de ses composantes,  $a_0$  :

La fonction inverse de l'équation d'ajustement (1) s'écrit :

$$x = A + B \cdot R^y \quad (3)$$

où

$$A = -b, B = e^{-\frac{c}{a}}, R = e^{\frac{1}{a}} \quad (4)$$

Ce type d'ajustement, proche des fonctions logistiques et de Gompertz, a été fréquemment étudié [3, 4, 5]. W.L. Stevens a trouvé une méthode de résolution ne requérant qu'une seule valeur initiale pour R. Dans son principe, cette méthode est identique à celle adoptée pour le modèle précédent : procédure itérative cherchant à maximiser la fonction de vraisemblance.

Soit donc le modèle statistique :

$$x = A + B \cdot R^y + \eta_y$$

où nous supposons que  $\eta_y$  jouit dans un premier temps des mêmes propriétés que  $e_x$ .

Les estimateurs A, B et R sont solutions de l'équation  $J(\hat{w}) = 0$ , où J(w) est le vecteur des dérivées premières de la vraisemblance :

$$H = \Sigma (x - A - B \cdot R^y)^2$$

Si I est la matrice d'information, à partir d'une estimation initiale  $w_0$ , on peut parvenir de proche en proche au vecteur  $w_{n+1}$  :

$$w_{n+1} = w_n + [I(w_n)]^{-1} \cdot J(w_n)$$

Dans le cas présent, il se trouve que, à la première itération par exemple, les solutions  $A_1$  et  $B_1$  ne dépendent pas des valeurs initiales  $A_0$  et  $B_0$  mais seulement de  $R_0$ .

On a :

$$J(W) = 2x \begin{bmatrix} A \cdot N + B \cdot \sum R^y - X \\ A \cdot \sum R^y + B \cdot \sum R^{2y} - X_2 \\ A \cdot B \cdot \sum y R^{y-1} + B^2 \cdot \sum y R^{2y-1} - B \cdot X_3 \end{bmatrix}$$

avec  $X_1 = \sum x, X_2 = \sum x R^y, X_3 = \sum x y R^{y-1}$

$$I(W) = 2x \begin{bmatrix} N & \sum R^y & B \cdot \sum y R^{y-1} \\ \sum R^y & \sum R^{2y} & B \cdot \sum y R^{2y-1} \\ B \cdot \sum y R^{y-1} & B \cdot \sum y R^{2y-1} & B^2 \cdot \sum y^2 R^{2y-2} \end{bmatrix}$$

On le voit,  $I(w)$ , ne dépend que de  $B$  et  $R$ . Son inverse peut s'écrire :

$$I^{-1}(w) = \begin{bmatrix} F_1 & F_2 & F_3/B \\ F_2 & F_4 & F_5/B \\ F_3/B & F_5/B & F_6/B^2 \end{bmatrix}$$

où les  $F_1 \dots F_6$  ne dépendent maintenant que de  $R$ .

En désignant par  $\delta w_o$  le vecteur accroissement, on montre facilement qu'on obtient les estimations.

$$A_1 = A_o + \delta A_o = F_1 \cdot X_1 + F_2 \cdot X_2 + F_3 \cdot X_3$$

$$B_1 = B_o + \delta B_o = F_2 \cdot X_1 + F_4 \cdot X_2 + F_5 \cdot X_3$$

$$B_1 \cdot \delta R_o = F_3 \cdot X_1 + F_5 \cdot X_2 + F_6 \cdot X_3$$

Ainsi, à partir de  $R_o$  et des  $N$  couples  $(x, y)$  on calcule les  $F_i$ , les  $X_i$  et delà,  $A_1, B_1$  et  $B_1 \delta R_o$ , pour parvenir à une nouvelle valeur  $R_1$ . Théoriquement, rien n'empêche de poursuivre les itérations sur la base du triplet  $(A_1, B_1, R_1)$  ; la précision de l'ajustement peut ainsi être améliorée. Il faut remarquer cependant que, sauf cas particulier, le gain ne justifie guère la poursuite du calcul. Les estimations de  $M_2$  sont des paramètres auxiliaires que la procédure examinée pour  $M_1$  devra améliorer. Surtout, il faut noter que toutes les hypothèses nécessaires pour les calculs précédents ne sont pas strictement vérifiées ou du moins ne peuvent-elles l'être à la fois pour  $M_1$  et  $M_2$  ; on pense principalement aux conséquences de l'échange effectué entre variable expliquée et variable explicative.

Le passage du modèle auxiliaire aux paramètres initiaux du modèle principal est assuré par les relations :

$$\begin{aligned}
 a_1 &= \frac{1}{\text{Log } R_1} \\
 b_1 &= -A_1 \\
 c_1 &= -\frac{\text{Log } B_1}{\text{Log } R_1}
 \end{aligned}
 \tag{5}$$

### 1.3 – Estimation préalable de $a_0$

Pour plusieurs raisons dont quelques unes rejoignent les fondements incertains des hypothèses relatives à  $M_2$ , il est préférable d'estimer au préalable  $a_0$  plutôt que  $R_0$ . L'égalité  $R_0 = e^{1/a_0}$  permet ensuite de passer d'un paramètre à l'autre.

Pour  $a_0$ , plusieurs méthodes sont possibles :

- développement limité de l'équation (3) et ajustement parabolique,
- élimination de  $b$  et  $c$  à partir des équations (1) et (3),
- utilisation de la pente de l'équation (1)

Pour les séries utilisées lors des calculs, aucune de ces méthodes n'a donné des résultats satisfaisants. On leur a préféré les deux procédures suivantes :

- 1)  $a_0$  est déduit de l'estimation de  $M_1$  par la méthode des moindres carrés avec  $b = 0$ .
- 2)  $a_0$  est déterminé à partir des différences premières de l'équation (1) :

$$\Delta(y) = y_{x-1} - y_x$$

Si  $x + b$  est suffisamment grand,  $\Delta(y)$  vérifie :

$$\Delta(y+n) = \frac{a \Delta(y)}{a - n \Delta(y)}$$

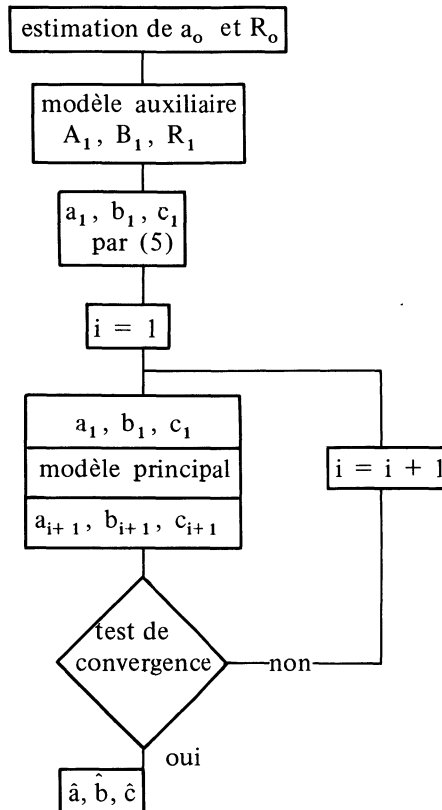
et

$$a = \frac{n \Delta(y+n) \Delta(y)}{\Delta(y+n) - \Delta(y)}$$

En pratique, on choisit la valeur  $a_0$  correspondant à  $n = N - 1$  ou  $n = N - 2$ .

La double estimation a pour seul but d'amorcer la convergence des itérations pour au moins une des valeurs de ce paramètre. On sait que pour qu'elle ait lieu, il faut au départ se situer assez près de la vraie valeur  $\hat{v}$ . Sur certaines séries plus ou moins régulières la seconde estimation peut conduire à des valeurs de  $a_0$  aberrantes (par exemple, négatives avec une tendance croissante). Dans ce cas, il faut partir de la première, généralement moins précise quand l'hypothèse d'ajustement est vérifiée, mais jamais aberrante puisque dépendante de toutes les observations.

L'organigramme ci-dessous résume la procédure d'estimation des paramètres  $a$ ,  $b$  et  $c$ .



## 2 - APPLICATIONS AUX DIAGRAMMES DE CHARGE

L'exemple d'application concerne les diagrammes de charge, plus précisément les diagrammes de la consommation d'électricité (pertes incluses) de la France pendant les mois de décembre. Les coordonnées des diagrammes sont respectivement les 24 heures de la journée et les puissances appelées en ces différentes heures par l'ensemble des utilisateurs. Les données statistiques de base que nous considérons sont les relevés effectués par E.D.F. tous les troisièmes mercredis de chaque mois [6].

On les notera  $P_{hix}$ , les indices se rapportant à la puissance appelée à l'heure  $h$  du mois  $i$  de l'année  $x$ .

Pour l'entreprise fournissant l'énergie électrique, les diagrammes de charge sont d'une grande importance. Ils sont le complément indispensable des prévisions à court, moyen et long terme de la consommation en ce sens qu'ils la répartissent entre les différentes heures des jours de l'année.

Leur déformation au cours du temps traduit le développement des usages électriques et également les modifications, spontanées et plus souvent provoquées, dans le comportement des utilisateurs. Sur un horizon de 5 à 10 ans, ils servent de base au choix des investissements pour la production d'électricité.



## 2.1 – Préviation des diagrammes de charge.

De très nombreuses méthodes de préviation des diagrammes de charge ont été élaborées [7]. Parmi elles, l'une consiste à remplacer ces diagrammes par des diagrammes dits réduits obtenus en divisant chaque puissance  $P_{hix}$  par la puissance moyenne du jour considéré [8]. On peut ainsi percevoir l'évolution et la déformation des diagrammes indépendamment de l'accroissement annuel de la consommation d'électricité et des fluctuations ou écarts accidentels (température, ensoleillement. . .)

On définit les puissances réduites par :

$$\pi_{hix} = \frac{P_{hix}}{P_{.ix}} \quad (6)$$

$$P_{.ix} = \frac{1}{24} \sum_{h=1}^{24} P_{hix}$$

On vérifie immédiatement que

$$\sum_{h=1}^{24} \pi_{hix} = 24 \text{ pour tout } i \text{ et tout } x. \quad (7)$$

Sous réserve de prévoir les puissances moyennes  $P_{.ix}$  la préviation des diagrammes réduits, c'est-à-dire à la préviation des  $\pi_{hix}$ .

L'étude [8] déjà citée a mis en évidence une propriété intéressante des  $\pi$ , à savoir leur tendance linéaire, croissante ou décroissante, en fonction du temps  $x$ . Sur cette base, on peut proposer le modèle statistique :

$$\pi_{hx} = a_h \cdot x + b_h + e_{hx}$$

(l'indice  $i$  sera dorénavant sous-entendu)

Les projections correspondantes ne donnent des résultats acceptables qu'à court terme, au maximum pour une échéance de trois ou quatre ans. Elles supposent, par le biais du coefficient  $a_h$  constant dans le temps, une stabilité de l'environnement économique et énergétique et, de surcroît, une croissance ou une décroissance illimitée des puissances réduites. Cette hypothèse est en contradiction aussi bien avec les conditions d'offre à court terme du producteur qu'avec les interventions constantes des institutions publiques chargées de modifier la forme du diagramme notamment à l'aide de la tarification.

L'objectif poursuivi en proposant un autre type d'ajustement est d'étendre le domaine de validité des projections tout en s'assurant qu'à court terme elles sont au moins aussi satisfaisantes qu'avec le modèle linéaire <sup>(1)</sup>.

Cette extension du champ d'application passe par une justification théorique de l'évolution des puissances réduites et non plus seulement par la constatation d'une liaison statistique avec le temps. On peut chercher un tel fondement en considérant la pente du nuage d'observations.

-----  
(1) Il faut entendre par "satisfaisantes" un coefficient de corrélation pour  $M1$  au moins égal à celui obtenu pour le modèle linéaire.

On remplace la pente  $a_h$  du modèle linéaire par la fonction  $p(x)$  vérifiant la relation de récurrence :

$$p(x+1) = \frac{a \cdot p(x)}{a + p(x)} \quad (8)$$

Cette relation traduit bien, en première approximation, le phénomène que l'on veut représenter : décroissance de  $p(x)$  vers zéro quand  $x$  croît. En outre, elle est formellement proche de l'équation  $P' = \frac{K \cdot P}{K+P}$  ( $K \geq 0$ ), équation classique de la théorie des servomécanismes donnant les corrections successives apportées à une valeur initiale de  $P$ . Cette référence aux asservissements tente de prendre en compte l'influence des réactions tarifaires destinées à corriger, d'une façon plus ou moins continue et avant d'atteindre des situations aberrantes, les déformations trop importantes du diagramme de charge.

Par un double changement de variables :

$$\begin{aligned} P(x) &= f(x) - a \\ f(x) &= \frac{g(x+1)}{g(x)} \end{aligned}$$

on déduit la solution de (8) :

$$P(x) = \frac{a}{x + b}$$

La fonction intégrale est :

$$\pi(x) = a \cdot \text{Log}(x + b) + c$$

dont nous savons estimer les trois paramètres  $a$ ,  $b$ ,  $c$  à partir des observations  $(\pi, x)$

En vue d'assurer la définition du logarithme, on supposera  $b \geq 0$  ( $x$  est lui-même positif).

Les estimations numériques des paramètres étant connues, on choisit comme prévision de  $\pi$  l'espérance mathématique de la variable estimée pour différentes valeurs de  $x$ . Quelques rectifications, en général mineures, sont nécessaires pour vérifier l'égalité (7).

## 2.2 – Résultats numériques

Parmi l'ensemble des diagrammes de charge mensuels, on a retenu celui de décembre, particulièrement significatif parce que les puissances appelées y atteignent, on le suppose du moins, leur maximum annuel. Quant aux années considérées, les effets de la nouvelle tarification établie en 1957 ne commençant à se faire sentir d'une façon significative qu'à partir de 1960, on ne retiendra que les années postérieures à cette date.

On est ainsi conduit à traiter un ensemble de 24 séries horaires, chacune comportant 12 observations (1961–1972). Les valeurs en sont les  $\pi_{hx}$  déduits des  $P_{hx}$  de décembre par la relation (6).

Les séries révèlent pour la plupart des heures de la journée une tendance nette, croissante ou décroissante selon les cas, accompagnée d'une diminution progressive de la pente. Pour les heures de la nuit et la majorité des heures de la matinée, les observations sont peu dispersées autour de la tendance. Elles le sont nettement plus pour les heures de l'après-midi, avec parfois des valeurs aberrantes. Enfin, pour certaines heures, 6h, 20h, 21h et 22h, il ne se dégage aucune tendance nette. Elles ont de ce fait été exclues des calculs et, pour la prévision, ont fait l'objet d'un traitement particulier (agrégation et estimation).

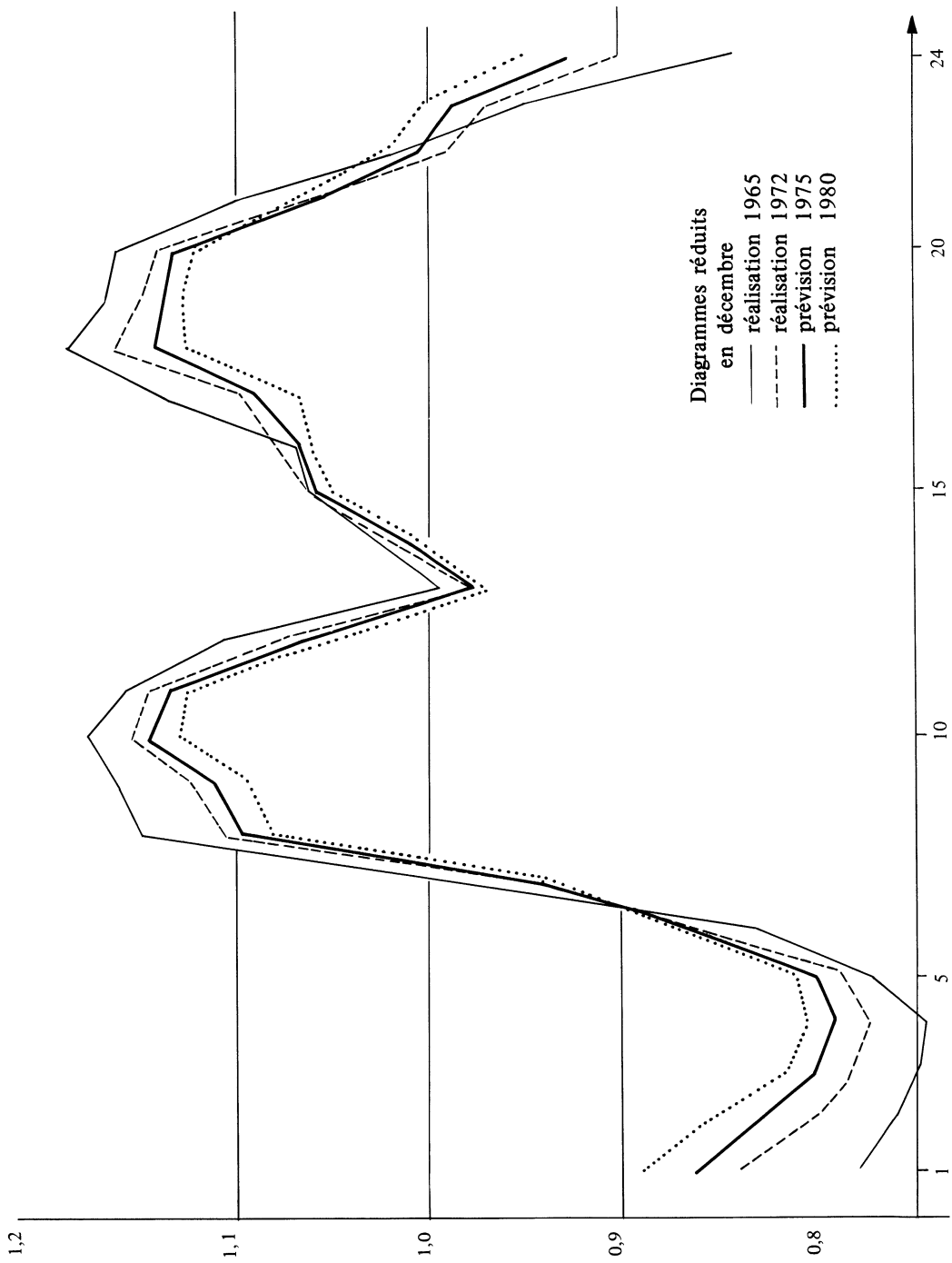
Les valeurs des paramètres  $\hat{a}_h$ ,  $\hat{b}_h$ ,  $\hat{c}_h$  des 20 séries analysées sont données dans le tableau ci-après. On y trouve également :

- le coefficient de corrélation  $R^2_{LI}$  avec l'hypothèse d'un modèle linéaire,
- le coefficient de corrélation  $R^2$  avec l'hypothèse M1,
- les valeurs empiriques et théoriques des  $\pi_h$  pour l'année 1972,
- les valeurs prévues pour 1975 et 1980.

h	Modèle (M1)				Mod.Linéaire $R^2_{LI}$	Réalisation 1972	Valeurs Théoriques $\pi_h$		
	$\hat{a}_h$	$\hat{b}_h$	$\hat{c}_h$	$R^2$			1972	1975	1980
1	0,1494	8,75	0,3917	0,986	0,973	0,841	0,841	0,863	0,891
2	0,1260	6,87	0,4453	0,947	0,928	0,807	0,833	0,833	0,858
3	0,0804	4,87	0,5626	0,973	0,947	0,785	0,790	0,802	0,820
4	0,0635	3,35	0,6095	0,933	0,889	0,776	0,783	0,793	0,807
5	0,0365	2,16	0,6964	0,825	0,760	0,788	0,793	0,799	0,811
6	–	–	–	–	0,287	0,854	0,850	0,857	0,865
7	–0,0223	0	1,0031	0,531	0,376	0,950	0,948	0,944	0,936
8	–0,0526	0,35	1,2393	0,905	0,832	1,104	1,107	1,095	1,080
9	–0,1083	8,09	1,4467	0,957	0,944	1,121	1,122	1,107	1,090
10	–0,0814	9,82	1,4039	0,710	0,698	1,152	1,153	1,145	1,125
11	–0,0342	0,89	1,2251	0,676	0,599	1,143	1,138	1,133	1,121
12	–0,0473	5,07	1,1967	0,453	0,434	1,077	1,063	1,058	1,044
13	–0,0940	18,27	1,2930	0,682	0,679	0,977	0,981	0,977	0,969
14	–0,0278	4,39	1,0932	0,636	0,604	1,021	1,016	1,013	1,005
15	–0,0327	5,89	1,1584	0,398	0,377	1,060	1,062	1,059	1,050
16	–0,0332	5,56	1,1693	0,394	0,370	1,077	1,072	1,066	1,059
17	–0,1426	19,30	1,5871	0,775	0,771	1,094	1,096	1,084	1,065
18	–0,0818	8,21	1,3954	0,822	0,805	1,161	1,149	1,140	1,125
19	–0,0119	0	1,1703	0,448	0,349	1,148	1,138	1,135	1,125
20	–	–	–	–	0,529	1,139		1,130*	1,120*
21	–	–	–	–	0,350	1,064	} 3,195*	1,056*	1,060*
22	–	–	–	–	0,004	0,990		1,005*	1,020*
23	0,0571	0	0,8326	0,771	0,689	0,969		0,975	0,986
24	0,1182	4,23	0,6532	0,911	0,881	0,902	0,908	0,920	0,950
TOT.						24,000	24,000	24,000	24,000

\* Estimations

Le graphique suivant superpose quatre diagrammes de charge pour deux années passées, 1965 et 1972 et pour deux années futures, 1975 et 1980.



### 2.3 – Conclusions.

On se limitera ici aux conclusions d'ordre purement statistique.

1) On remarque que pour toutes les régressions effectuées on a  $R^2 > R^2_{L1}$ , ce qui satisfait à un des objectifs du modèle M1.

De la comparaison, pour 1972, entre réalisation et valeur théorique on tire les renseignements suivants :

- écart maximum : 0,014 soit 1,3 % par rapport à la valeur observée ;
- écart-type de la différence (calculée pour les 24 heures) : 0,006.

2) Quant aux estimations des paramètres, on voit que les valeurs de b sont de l'ordre de quelques unités. Dans deux cas elles dépassent 10 et dans trois autres cas elles sont nulles, rendant ainsi équivalent le modèle testé avec le modèle semi-logarithmique classique.

Notons au passage que b peut s'interpréter comme un paramètre-retard. En effet, si la tendance fondamentale du phénomène correspond bien à une fonction logarithmique l'incidence des facteurs d'asservissement a pour résultat de donner comme mesure du phénomène à l'instant t la mesure du phénomène, supposé non influencé par ces facteurs, tel qu'il serait apparu à l'instant t + b.

En ce qui concerne encore b, on doit constater que la variance de l'estimateur est généralement élevée, contrairement à celles de  $\hat{a}$  et de  $\hat{c}$ .

Les matrices des variances-covariances montrent aussi que les trois estimateurs sont fortement corrélés entre eux. Ce sont là deux inconvénients de la méthode d'ajustement proposée.

3) Pour les prévisions de la période 1975–1980, il est difficile de juger de leur qualité en soi. Les enseignements statistiques en l'occurrence ne suffisent plus ; il faut les renforcer par des jugements de valeurs spécifiques au domaine d'application, ici l'électricité, et corroborer les résultats par d'autres méthodes.

Pour terminer on voudrait signaler quelques prolongements possibles dans l'utilisation du modèle M1. Bien que d'autres applications économiques puissent être envisagées, en particulier en prenant une autre grandeur que le temps comme variable explicative, on se limitera là encore aux diagrammes de charge.

1. Estimer les paramètres (a, b, c) sous la contrainte explicative  $\sum \pi_{hx} = 24$  pour tout x.

2. Chercher, pour une chronique donnée  $\pi_1$ , parmi l'ensemble des  $24 \times 12$  chroniques  $\pi_{hix}$  supposées ajustées par M1, la chronique équivalente  $\pi_2$  définie par :

$$\pi_2(x+k) = \pi_1(x) + \epsilon_x$$

ou

$$\pi_2(x+k) = h \cdot \pi_1(x) + 1 + \epsilon_x$$

La première série est décalée de k par rapport à la seconde et joue, pour la prévision de  $\pi_2$ , le rôle de série-anticipation;

## REFERENCES BIBLIOGRAPHIQUES

- [1] RAULT (C.) – “Etude économétrique de la possession d’un ensemble de biens durables de consommation”, *Annales de l’INSEE*, n° 1, mai 1969.
- [2] OLIVER (F.R.) – “Another generalisation of the logistic growth function”, *Econometrica*, vol. 37, n° 1, janvier 1969.
- [3] NELDER (J.A.) – “The fitting of a generalisation of logistic curve”, *Biometrics*, vol. 17, n° 1, 1961.
- [4] OLIVER (F.R.) – “Aspects of maximum likelihood estimation of logistic growth function”, *Journal of American Statistical Association*, vol. 61, n° 3, 1966.
- [5] STEVENS (W.L.) – “Asymptotic Regressions”, *Biometrics* ; septembre 1951.
- [6] E.D.F. – “Statistiques de la production et de la consommation” ; 1961. . . 1972.
- [7] Congrès de l’UNIPEDE. – n° 15 ; Cannes 1971
- [8] JUNG (D.) et PIOGER (G.) – “Etude de la modulation saisonnière des diagrammes de charge des jours ouvrables en France et de l’évolution de leur forme”, *l’Economie Electrique*, n° 41, 42, 43 ; 1965.