REVUE DE STATISTIQUE APPLIQUÉE

G. MORLAT

Modèle pour des chroniques économiques mensuelles

Revue de statistique appliquée, tome 11, nº 2 (1963), p. 5-20
<http://www.numdam.org/item?id=RSA_1963__11_2_5_0>

© Société française de statistique, 1963, 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/

MODÈLE POUR DES CHRONIQUES ÉCONOMIQUES MENSUELLES (1)

G. MORLAT

(Électricité de France — Institut de Statistique de l'Université de Paris)

INTRODUCTION ET RESUME -

L'étude de chroniques mensuelles - en vue de la correction des mouvements saisonniers ou de la prévision - a donné lieu à de très nombreux travaux, et les modèles probabilistes, plus ou moins complètement explicités, qui justifient les méthodes d'analyse proposées, sont assez variés. Mais ils se ramènent le plus souvent à l'addition de diverses composantes : tendance, coefficients saisonniers, aléas résiduels. (Bibl. [4], [6]).Les deux premières composantes sont des fonctions certaines du temps, auxquelles il faut assigner une forme précise ; les aléas résiduels sont supposés généralement indépendants et de même loi. Bien que le choix des diverses fonctions qui interviennent soit arbitraire, les praticiens savent combien ce modèle manque souvent de souplesse.

Nous décrirons dans cette note un modèle simple, que nous pouvons caractériser brièvement en disant qu'il s'agit d'un processus à accroissements markoviens, à lois de probabilité de passage saisonnières. Un tel modèle qui généralise quelque peu le modèle sous-jacent à des pratiques courantes de "désaisonnalisation", se montre parfaitement adéquat à la représentation de chroniques telles que :

- *consommations mensuelles d'énergie,*
- *indices de production industrielle ou de production nationale,*
- *etc.*

Il rend effectivement des services, depuis un certain nombre d'années dans le traitement des consommations d'électricité en France (mensuelles ou hebdomadaires).

L'évolution de certaines chroniques mensuelles peut être représentée par une chaine de Markov Laplacienne, à paramètres périodiques. Cette chaine comportera donc dans le cas le plus général 36 parametres (moyennes, variances et covariances pour chaque mois). Pratiquement, on peut avoir intérêt à supposer les coefficients de corrélation égaux (modèle à 25 paramètres) voire même les variances égales (modèle à 14 paramètres).

---

(1) Note présentée à la Conférence Commune : Institut de Statistique Mathématique, The Institute of Management Sciences et Société d'Econométrie, DUBLIN, Septembre 1962. -

Pour des raisons dont l'interprétation économique est assez évidente, ce sont les accroissements successifs des chroniques de production qui sont bien représentés par de tels modèles.

L'estimation des paramètres ne soulève pas de grandes difficultés pratiques ; on peut très commodément prendre en compte et tester à l'aide des techniques statistiques classiques, des modifications éventuelles de certains paramètres au cours du temps (ce qui élimine la difficulté liée aux "ruptures de tendance" dans certains modèles classiques).

Lorsqu'on connait une chronique jusqu'à l'époque t, la loi de probabilité de sa valeur en t + h peut être obtenue très facilement. Cette loi peut servir utilement de base à des prévisions à court terme.

On désire souvent associer à une chronique donnée, une chronique fictive, "ressemblant" le plus possible à la première, mais dépourvue de mouvement saisonnier. C'est le problème de la "désaisonnalisation". Avec les modèles les plus simples admis habituellement, la solution consiste à retrancher (ou à diviser par) des coefficients saisonniers. Dans le cadre du modèle étudié ici, la solution est moins simple, mais on calcule sans peine une "chronique non saisonnière associée".

Nous donnerons "in fine" quelques indications concernant l'emploi d'un modèle à accroissements markoviens laplaciens pour l'étude des consommations électriques mensuelles françaises.

EQUATIONS D'UN MODELE A ACCROISSEMENTS MARKOVIENS LAPLACIENS -

Soit y_t une chronique (t exprimé en mois) représentant une grandeur économique de base : consommation d'énergie, production d'une industrie lourde, indice de production nationale, etc. dans un pays en expansion. Grosso modo, les possibilités d'accroissement, au cours d'une période de temps donnée, sont proportionnelles au niveau de production déjà atteint. Pour avoir de bonnes chances d'utiliser correctement un processus stationnaire, il faut penser que nous devons prendre en considération l'accroissement relatif de y , de préférence à toute autre grandeur liée à cette chronique. Des raisons habituelles de commodité technique conduisent à choisir la quantité :

$$\log \frac{y_t}{y_{t-1}}$$

pour caractériser l'accroissement au cours de l'unité de temps.

C'est cette quantité que nous proposons de représenter par un processus de Markov, à lois laplaciennes, dont les paramètres auront une évolution périodique annuelle. Bien qu'on puisse penser à priori que

doivent intervenir des liaisons assez "longues", quelques expériences sur des chroniques énergétiques ou des indices de production nous ont montré qu'un processus markovien du premier ordre convient parfaitement bien.

Notre modèle sera donc défini par les relations suivantes dont la forme "développée" se prête aux raisonnements qui suivront et en explicite bien le mécanisme :

$$\begin{cases} (1\ a) & \log y_t = x_t \\ (1\ b) & x_t = x_{t-1} + z_t \\ (1\ c) & z_t = m_{[t]} + w_t \sigma_{[t]} \ ; \ ([t] = t \bmod. 12) \\ (1\ d) & E(w_t) = 0, \ \text{var}\ w_t = 1 \\ (1\ e) & w_t = \rho_{[t]}\ w_{t-1} + \sqrt{1 - \rho^2_{[t]}}\ u_t \end{cases} \qquad (1)$$

u_t suite de variables aléatoires laplaciennes réduites indépendantes (a, 1, i, r).

Cette écriture met en évidence, de la façon la plus claire, la structure du modèle. A partir d'une suite d'aléas laplaciens réduits indépendants, on imagine la construction d'une chaîne simple de Markov à autocorrélation périodique (1 e - 1 d), de moyenne et de dispersion périodique (1 c) et ces grandeurs ainsi enchaînées constituent les accroissements successifs $\log \frac{y_t}{y_{t-1}}$ de la chronique étudiée (1 b - 1 a). Cette structure semble la plus simple qu'on puisse imaginer, en conformité avec les arguments esquissés au début de ce paragraphe.

Sous un autre aspect, les relations (1) définissent y_t comme un processus de Markov d'ordre 2 particulier, son logarithme, x_t, étant l'intégrale d'un processus d'ordre 1. (Markov-Laplace).

Indépendamment des valeurs initiales de la chronique, les relations (1) définissent un modèle à 36 paramètres. Pour certaines applications, on pourra se borner à 25 paramètres en posant $\rho_j = \rho$, voire même à 14 paramètres en adjoignant encore l'hypothèse supplémentaire $\sigma_j = \sigma$. C'est ce que nous ferons par exemple pour la consommation d'électricité en France, en raison de la limitation des données d'observation ; mais un modèle à 25 paramètres constituerait un progrès certain. On peut remarquer dès maintenant que le modèle à 14 paramètres donne une représentation formellement identique à un modèle classique de décomposition saisonnière, les écarts $m_j - \overline{m}$ pouvant être interprétés comme des coefficients saisonniers au sens habituel.

L'ESTIMATION DES PARAMETRES -

L'estimation des paramètres du modèle à accroissements markoviens ne présente guère de difficulté particulière. A partir de la différenciation de la chronique étudiée, on établira une table de BUYS BALLOT qui se prête au calcul des moyennes et variances par colonnes, ainsi que des covariances entre deux colonnes successives. Ces dernières permettent d'estimer correctement les coefficients d'autocorrélation ρ_j, pourvu que l'on puisse considérer que le produit :

$$\prod_{j=1}^{12} \rho_j$$

est négligeable devant chaque ρ_j (et ce sera souvent le cas en pratique) ; on peut alors traiter le problème de l'estimation d'un ρ_j comme celui de l'estimation d'un coefficient de corrélation ordinaire, à partir de couples indépendants d'observations. Si l'on a des raisons de soupçonner quelque changement de structure pour la chronique étudiée au cours des années d'observations disponibles, la vérification se présente comme un test d'homogénéité classique pour les paramètres en cause (généralement moyennes ou variances) ; on trouvera un exemple dans le dernier paragraphe, à propos des consommations électriques.

La question des "ruptures de tendance" ou autres phénomènes similaires, qui cause de graves ennuis et oblige à des opérations fort arbitraires avec d'autres modèles, apparaît donc ici comme un problème fort simple et classique.

o

o o

Il arrivera fréquemment qu'on ait à étudier, en même temps qu'une chronique mensuelle, la chronique annuelle correspondante. Il s'agira en général de répondre à des problèmes différents, mais une certaine cohérence entre les modèles utilisés peut sembler souhaitable. La chronique annuelle correspondant au modèle que nous étudions ici peut s'écrire (a désignant l'année) :

$$Y_a = y_{12a+1} + y_{12a+2} + \dots + y_{12(a+1)}$$

(car une chronique annuelle est souvent définie par addition des valeurs mensuelles : consommation d'énergie, indices de production, etc.).

Si les y_t obéissent au modèle défini par les relations (1), les Y_a doivent théoriquement obéir à des lois de probabilité fort complexes et peu maniables. En pratique, nous avons cependant admis que les Y_a peuvent eux aussi être représentés par un modèle à accroissements markoviens laplaciens. L'examen des statistiques rend cette hypothèse fort plausible pour les chroniques que nous avons étudiées ; ce modèle annuel est bien entendu dépourvu de toute périodicité, et l'autocorrélation entre $\log Y_a$ et $\log Y_{a+1}$ est souvent pratiquement nulle. (Bibl. [5]).

La cohérence entre les paramètres d'accroissements moyens des modèles mensuel et annuel est bien facile à vérifier.

Un point plus délicat concerne la cohérence entre la variance du modèle annuel, d'une part, les variances et covariances du modèle mensuel, d'autre part. Cela constitue d'ailleurs en quelque sorte un test d'adéquation du modèle mensuel Markov-Laplace, et mérite quelques développements.

Soit Σ le paramètre de dispersion du modèle annuel :

$$\Sigma^2 = \text{var} \log \frac{Y_a}{Y_{a-1}}$$

et σ le paramètre de dispersion (supposé constant) du modèle mensuel ; on supposera aussi ρ constant.

On peut montrer, en utilisant un développement limité - ce qui implique que les accroissements mensuels et annuels, de la chronique étudiée, soient assez petits - qu'on a approximativement :

$$\Sigma^2 = \text{var} \log \frac{Y_a}{Y_{a-1}} \simeq \text{var}\,(x_{12a+1} + x_{12a+2} + \ldots + x_{12(a+1)})$$

La cohérence des modèles mensuel et annuel implique une certaine relation entre Σ^2 d'une part, ρ et σ d'autre part.

(Nous supposons, au moins comme une approximation, que l'autocorrélation de la chronique annuelle est nulle). Nous allons établir cette relation.

Supposons plus généralement une périodicité de longueur k dans le modèle défini par les formules (1) que nous reproduisons partiellement ci-dessous :

$$x_t = x_{t-1} + z_t$$

$$E(z_t) = m_{[t]} \qquad [t] = t \text{ modulo } k$$

$$\text{var } z_t = \sigma^2$$

$$\text{cov}\,(z_t\, z_{t+1}) = \rho\,\sigma^2$$

Nous voulons calculer :

$$\Sigma^2 = \text{var}\,(x_1 + x_2 + \ldots + x_k)$$

$$= \text{var}\,[k\, z_1 + (k-1)\, z_2 + \ldots + z_k]$$

$$\begin{aligned}\frac{\Sigma^2}{\sigma^2} = \sum_{j=1}^{k} j^2 &+ 2\rho\,[k\,(k-1) + (k-1)\,(k-2) + \ldots] \\ &+ 2\rho^2[k\,(k-2) + (k-1)\,(k-3) + \ldots] \\ &+ 2\rho^3[k\,(k-3) + \ldots \qquad\qquad] \\ &+ \ldots\ldots\ldots\ldots \\ &+ 2\,\rho^{k-1} k\end{aligned}$$

Chacune des sommes entre crochets étant poursuivie jusqu'à ce qu'un terme s'annule.

La ligne de rang λ comporte ainsi une somme de la forme :

$$\begin{aligned} & k \quad (k-\lambda) \\ + & (k-1)\ (k-\lambda-1) \\ + & (k-2)\ (k-\lambda-2) \\ + & \ldots\ldots\ldots\ldots \\ + & (1+\lambda)\end{aligned}$$

Posant $1 + \lambda = u$, $k - \lambda = v$, et écrivant cette somme dans l'ordre inverse, il vient :

$$u + 2(u+1) + 3(u+2) + \ldots + v(u+v-1)$$

$$= u(1+2+\ldots+v) + \sum_{1}^{v} j(j-1)$$

$$= u.\ \frac{v(v+1)}{3} + \frac{(v-1)\,v(v+1)}{3} = v(v+1)\left[\frac{u}{2} + \frac{v-1}{3}\right]$$

ou encore, en revenant aux notations initiales :

$$(k-\lambda)(k-\lambda+1)\left(\frac{1+\lambda}{2} + \frac{k-\lambda-1}{3}\right)$$

Donc :

$$\frac{\Sigma^2}{\sigma^2} = \frac{k(k+1)(2k+1)}{6} + 2\rho(k-1)\,k\left(1 + \frac{k-2}{3}\right)$$

$$+ 2\rho^2(k-2)(k-1)\left(\frac{3}{2} + \frac{k-3}{2}\right) + 2\rho^3(k-3)(k-2)\left(\frac{4}{2} + \frac{k-4}{3}\right)$$

$$+ 2\rho^4(k-4)(k-3)\left(\frac{5}{2} + \frac{k-5}{3}\right) + \ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$+ 2\rho^{k-1}k$$

Dans le cas particulier $k = 12$, on trouve :

$$\frac{\Sigma^2}{\sigma^2} = \phi(\rho) = 650 + 1144\rho + 990\rho^2 + 840\rho^3 + 696\rho^4 + 560\rho^5 + 434\rho^6 +$$

$$+ 320\rho^7 + 220\rho^8 + 136\rho^9 + 70\rho^{10} + 24\rho^{11}$$

Nous donnons ci-dessous une table de $\phi(\rho)$:

ρ	$\phi(\rho)$	$\sqrt{\phi(\rho)}$
0	650	25,50
- 0,10	544,72 ...	23,34
- 0,20	455,04 ...	21,33
- 0,30	377,76 ...	19,44
- 0,40	310,49 ...	17,62
- 0,50	251,57 ...	15,86
- 0,60	199,17 ...	14,11
- 0,70	152,57 ...	12,35
- 0,80	110,66 ...	10,52
- 0,90	72,46 ...	8,51
- 1,00	36	6

FORMULES POUR LA PREVISION A COURT TERME -

Bornons-nous, pour simplifier les calculs, au cas d'un modèle à 14 paramètres, et reprenons les formules (1) qui définissent notre modèle :

$$x_t = x_{t-1} + z_t$$

$$z_t = m_{[t]} + w_t \rho$$

$$w_t = \rho w_{t-1} + \sqrt{1-\rho^2}\, u_t$$

u_t : aléas laplaciens réduits indépendants (a. l. i. r.).

La troisième formule exprime que les z_t forment une chaîne de Markov Laplacienne, de corrélation ρ au premier ordre. L'équation de régression du premier ordre s'écrit évidemment :

$$E(w_{t+1}/t) = \rho w_t$$

et l'équation de régression d'ordre h s'écrit :

$$E(w_{t+h}/t) = \rho^h w_t$$

On en déduit les espérances mathématiques, à l'époque t - c'est-à-dire lorsqu'on connaît la chronique jusqu'à x_t - des quantités : z_{t+1}, $z_{t+2}, \ldots, z_{t+h}$, qui s'écrivent :

$$E(z_{t+1}) = m_{[t+1]} + \rho\,(z_t - m_{[t]})$$

$$E(z_{t+2}) = m_{[t+2]} + \rho^2(z_t - m_{[t]})$$

.....

.....

$$E(z_{t+h}) = m_{[t+h]} + \rho^h(z_t - m_{[t]})$$

et on en déduira encore, par addition de toutes ces quantités à x_t, l'espérance en t de x_{t+h}, soit :

$$E(x_{t+h}/t) = x_t + E \sum_{j=1}^{h} z_{t+j}$$

$$= x_t + m_{[t+1]} + m_{[t+2]} + \ldots \quad + m_{[t+h]} + (\rho + \rho^2 + \ldots + \rho^h)\,(z_t - m_{[t]})$$

Remplaçant z_t par $x_t - x_{t-1}$ et réorganisant un peu l'expression précédente, on obtient le résultat :

$$\begin{aligned} E(x_{t+h}/t) = {} & m_{[t]} + m_{[t+1]} + \ldots + m_{[t+h]} \\ & + (1 + \rho + \rho^2 + \ldots + \rho^h)\,(x_t - m_{[t]}) \\ & + (\rho + \rho^2 + \ldots + \rho^h)\, x_{t-1} \end{aligned}$$

ou
$$E(x_{t+h}/t) = m_{[t]} + m_{[t+1]} + \dots + m_{[t+h]} + \frac{1-\rho^{h+1}}{1-\rho}(x_t - m_{[t]}) + \rho\,\frac{1-\rho^h}{1-\rho}\,x_{t-1}$$

Pour évaluer la précision d'une prévision à h mois d'échéance, on calculera la variance conditionnelle de x_{t+h}, connaissant la chronique jusqu'en x_t, variance que nous noterons :

$$\mathrm{var}\,(x_{t+h}/t)$$

or :
$$x_{t+h} = x_t + z_{t+1} + z_{t+2} + \dots + z_{t+h}$$

Appelons $\Sigma^2\sigma^2$ la variance de $z_{t+1} + \dots + z_{t+h}$, et $\Lambda\,\sigma^2$ la covariance de cette somme avec z_t. (R le coefficient de corrélation correspondant). Nous aurons :

$$\mathrm{var}\,(x_{t+h}/t) = \Sigma^2\sigma^2\,(1 - R^2) = \Sigma^2\sigma^2\left(1 - \frac{\Lambda^2}{\Sigma^2}\right) = (\Sigma^2 - \Lambda^2)\,\sigma^2$$

Or, on calcule sans peine Λ et Σ :

$$\Lambda = \frac{\mathrm{cov}\,(z_t,\ z_{t+1} + \dots + z_{t+h})}{\sigma^2} = \rho + \rho^2 + \dots + \rho^h = \rho\,\frac{1-\rho^h}{1-\rho}$$

et :

$$\Sigma^2 = \frac{\mathrm{var}\,(z_{t+1} + \dots + z_{t+h})}{\sigma^2} = h + 2\,(h-1)\,\rho + 2\,(h-2)\,\rho^2 + \dots + 2\,\rho^{h-1}$$

$$= h + 2\,[(h-1)\,\rho + (h-2)\,\rho^2 + \dots + \rho^{h-1}]$$

La somme entre crochets peut être décomposée sous la forme :

$$\begin{array}{lllll}
\rho^{h-1} + \rho^{h-2} + \rho^{h-3} + \dots & & & & + \rho \\
\quad + \rho^{h-2} + \rho^{h-3} + \dots & & & & + \rho \\
\qquad\quad + \rho^{h-3} + \dots & & & & + \rho \\
\qquad\qquad\quad + \dots & & & & \\
\qquad\qquad\qquad\qquad & + \rho^2 & + \rho & & \\
\qquad\qquad\qquad\qquad & & + \rho & &
\end{array}$$

et on voit qu'elle s'écrit :

$$\frac{\rho}{1-\rho}\,(1-\rho) + (1-\rho^2) + (1-\rho^3) + \dots + (1-\rho^{h-1}) = \frac{\rho}{1-\rho}\left[h-1 - \rho\,\frac{1-\rho^{h-1}}{1-\rho}\right]$$

d'où on tire finalement :

$$\mathrm{var}\,(x_{t+h}/t) = \sigma^2\left[h + 2\,(h-1)\,\frac{\rho}{1-\rho} - 2\left(\frac{\rho}{1-\rho}\right)^2 \frac{1-\rho^{h-1}}{1-\rho} - \left(\frac{\rho}{1-\rho}\right)^2 (1-\rho^h)^2\right]$$

A l'époque t, la loi de probabilité conditionnelle de y_{t+h} est une loi logarithmo-normale dont les paramètres sont fixés par les formules que nous venons d'établir. Sur le plan pratique, il semble que ces formules devraient rendre quelques services pour des prévisions à court terme (de l'ordre de 1 à 12 mois par exemple).

o

o o

PROCESSUS NON SAISONNIER ASSOCIE -

On sait tout l'intérêt qu'attachent les conjoncturistes à l'établissement de chroniques dépourvues de mouvement saisonnier. Avec le modèle d'accroissements de Markov-Laplace décrit dans cette note, nous pouvons associer, à un processus x_t de paramètres :

$$(m_1, m_2, \ldots, m_{12}, \sigma_1, \sigma_2, \ldots, \sigma_{12}, \sigma)$$

(pour nous borner, par raison de simplicité, au modèle à 25 paramètres) un autre processus ξ_t, non saisonnier, de paramètres :

$$(m, \sigma, \rho)$$

tel qu'une réalisation de ξ ressemble d'aussi près qu'il est possible (mouvement saisonnier mis à part) à la réalisation "correspondante" de x.

Nous conviendrons de prendre, pour achever de définir ξ :

$$m = \frac{\Sigma m_j}{12} \quad , \quad \sigma^2 = \frac{\Sigma \sigma_j^2}{12}$$

et d'associer à une réalisation de x, la réalisation de ξ résultant des mêmes a. l. i. r.

Il est bien clair que le calcul des ξ observés peut être effectué directement à partir des x, ou mieux encore, on exprimera la chronique $\eta = e^{\xi}$ en fonction de y, par les formules suivantes :

$$\eta_t = \eta_{t-1} \exp. \left\{ m - m_{[t]} \cdot \frac{\sigma}{\sigma_{[t]}} \right\} \cdot \left(\frac{y_t}{y_{t-1}} \right)^{\frac{\sigma}{\sigma_{[t]}}}$$

(L'établissement de cette "formule de désaisonnalisation" ne présente aucune difficulté : il suffit d'écrire explicitement les relations (1) pour les deux modèles, et d'éliminer les a. l. i. r. supposés identiques).

On pourra observer que dans le cas particulier d'un modèle à 14 paramètres, le processus non saisonnier associé est formellement identique au résultat de la technique classique, fondée sur l'hypothèse de coefficients saisonniers additifs, et mettant en jeu le filtre de BUYS BALLOT, à condition d'appliquer ce modèle classique aux accroissements de la chronique (Bibl. [2]).

o

o o

APPLICATION A LA CHRONIQUE DES CONSOMMATIONS MENSUELLES D'ELECTRICITE EN FRANCE -

Les tableaux ci-annexés donnent quelques résultats essentiels. Bornons-nous ici aux commentaires les plus nécessaires.

TABLEAU 1 - Les données de base sont les consommations moyennes des jours ouvrables, établies et utilisées par nos Collègues de la Direction des Etudes et Recherches (Bibl. [3]).

TABLEAU 2 - Il s'agit bien entendu des logarithmes naturels. Les moyennes indiquées au bas du tableau sont les valeurs empiriques, sauf pour les mois d'août et septembre, pour lesquels les z_t ne sont manifestement pas stationnaires (influence de l'allongement des congés payés, intervenu en 1956). Une étude graphique a conduit à retenir pour m_j les chiffres indiqués, respectivement avant et après 1956.

TABLEAU 3 - Nous donnons ici, outre les m relatifs à la période 1956-1960, les valeurs empiriques des écarts-types et coefficients d'autocorrélation du premier ordre.

Pour les écarts-types, il faut noter que les valeurs trouvées pour février et mars sont gravement affectées par le mois de février 1956 (froid exceptionnel).

Avec un échantillon d'une dizaine d'années, les variations de σ_j ne sauraient être regardées comme significatives. Nous retiendrons une valeur constante :

$$\sigma = .012$$

En ce qui concerne les coefficients de corrélation, les valeurs trouvées pour août, septembre, octobre, ne sauraient être retenues (rôle de la modification de la durée des congés) ; considérant que diverses influences perturbatrices détériorent la corrélation, nous admettrons au vu des valeurs expérimentales :

$$\rho = -.50$$

C'est d'ailleurs très précisément la valeur que l'on obtient si l'on calcule, comme il est dit dans les pages précédentes, la valeur de ρ qui explique la dispersion des accroissements des consommations annuelles.

Si nos présomptions sont fondées, nous obtiendrons des valeurs absolues plus élevées pour l'autocorrélation empirique, lorsque nous pourrons disposer, au départ, des valeurs de consommations corrigées des pertes et de l'influence de la température.

TABLEAU 4 - Le tableau 4 donne les valeurs du processus non saisonnier associé γ_t.

L'intérêt des résultats - au reste provisoires - tient au fait que la technique de calcul est justifiée par un modèle précis ; d'autre part, sa mise en oeuvre pourra être grandement améliorée par la prise en compte de consommations corrigées des effets de la température, des pertes de transports, etc. sans compter les raffinements dont peut facilement bénéficier le modèle lui-même.

Tableau 1

C_t = consommation moyenne des jours ouvrables, par mois, pour l'ensemble de la FRANCE (sans corrections) de 1950 à 1960
en millions de KWh

Années	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Septembre	Octobre	Novembre	Décembre
1950	96,7	94,5	84,3	92,8	93,2	92,8	89,6	83,2	94,7	99,9	105,9	111,7
1951	109,7	108,5	108,5	106,4	107,4	105,4	101,5	93,8	105,9	111,4	116,0	120,5
1952	122,7	119,1	115,1	114,1	111,1	110,8	107,0	97,1	112,9	116,5	124,9	129,1
1953	127,2	120,7	115,4	115,3	114,9	115,9	110,1	91,1	114,0	120,8	129,5	121,6
1954	132,3	131,3	125,4	126,7	127,1	125,6	119,0	108,1	124,7	130,3	136,1	140,1
1955	142,8	141,6	142,9	136,8	137,0	134,7	130,7	118,8	134,9	143,3	152,4	153,6
1956	155,3	161,5	151,3	149,9	149,1	149,3	142,8	126,3	149,6	158,0	170,0	170,1
1957	173,1	166,3	162,9	163,0	161,0	160,3	152,5	135,4	162,3	169,1	181,6	186,8
1958	185,8	182,3	184,7	180,6	175,2	174,4	165,8	139,0	172,2	182,3	192,2	195,1
1959	195,1	188,9	184,6	185,5	182,4	181,9	173,0	145,6	182,1	192,4	204,8	210,6
1960	212,3	211,1	206,3	202,0	198,1	198,4	191,6	166,1	206,8	219,7	231,4	238,1

Tableau 2

$$\text{Log } \frac{C_t}{C_{t-1}} = x_t - x_{t-1} = z_t$$

Années	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Septembre	Octobre	Novembre	Décembre	
1951	018	-011	0	-020	009	-018	-038	-079	121	051	040	037	
1952	018	-029	-034	-009	-021	-008	-035	-098	150	031	070	033	
1953	-015	-053	-045	-001	-003	009	-051	-189	224	058	070	001	
1954	021	-008	-046	011	033	-012	-053	-097	143	044	043	030	
1955	019	-008	009	-044	002	-017	-029	-097	128	060	061	008	
1956	011	038	-065	-010	-005	001	-045	-122	169	054	073	001	
1957	017	-040	-020	0	-012	-004	-050	-119	181	041	071	028	
1958	-005	-019	013	-022	-030	-005	-050	-177	213	057	053	015	
1959	0	-031	-023	005	-017	-003	-049	-173	224	054	063	028	
1960	008	-006	-023	-021	-020	002	-035	-143	219	061	052	029	
m_j	005	-017	-023	-011	-010	-006	-004	-100 -150	150 200	051	060	021	076

Tableau 4

Valeurs de γ_t - Processus non saisonnier associé à C_t -
Base 100 : Décembre 1950

Années	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Septembre	Octobre	Novembre	Décembre
1951	98,3	99,5	102,5	102,1	104,7	104,1	105,3	108,2	105,8	106,4	105,0	107,4
1952	109,5	108,8	108,4	109,2	108,7	109,1	110,8	111,8	112,4	111,0	112,9	115,0
1953	113,4	110,0	108,3	110,2	111,6	113,9	113,8	104,7	113,3	114,8	116,6	115,0
1954	117,5	119,3	117,4	120,8	123,1	123,1	122,6	123,8	123,7	123,6	122,3	124,1
1955	126,6	128,5	133,5	130,1	132,4	131,7	134,5	135,8	133,5	135,5	136,6	135,7
1956	137,8	145,9	140,8	141,8	143,4	145,2	146,0	150,9	147,2	148,6	151,6	149,5
1957	152,2	149,6	151,1	153,8	154,4	155,6	155,6	161,6	159,6	159,0	161,8	163,9
1959	170,7	169,3	170,5	174,4	174,2	175,8	176,0	173,0	178,2	179,9	181,7	184,1
1960	185,8	188,9	190,2	189,5	188,7	191,4	194,2	196,8	201,7	204,9	204,5	207,3

Tableau 3

Paramètres du modèle markovien

	m_j	σ_j	ρ_j
Janvier	005	014	-·34
Février	-017	(023)	·27
Mars	-023	(023)	-·24
Avril	-011	015	-·66
Mai	-010	012	-·02
Juin	-006	008	-·42
Juillet	-044	008	-·42
Août	-150	-	(·51)
Septembre	200	-	(-·95)
Octobre	051	010	(·41)
Novembre	060	012	-·18
Décembre	021	013	-·50

On retiendra : σ = ·012 ρ = -50

QUELQUES CONCLUSIONS -

Nous avons déjà indiqué quelles sont les chroniques économiques pour lesquelles le modèle à accroissements markoviens laplaciens doit à priori bien convenir. L'application aux consommations mensuelles d'électricité en France confirme que cette présomption était fondée.

Au cours de la dernière année scolaire, nous avons eu l'occasion de décrire devant des étudiants de troisième cycle de la Faculté des Sciences, en statistique mathématique et économétrie, le modèle qui fait l'objet de ce mémoire. Nous les avons invités à se livrer à quelques applications numériques à diverses chroniques économiques, et leurs travaux ont quelque peu enrichi notre expérience.

Nos étudiants ont pu constater, par exemple, que le modèle à accroissements markoviens laplaciens représente assez bien l'évolution des indices de production industrielle, mais l'estimation du coefficient d'autocorrélation présente quelque difficulté, en raison des limites de précision dans la définition des indices. On peut noter au passage que ce coefficient d'autocorrélation peut fournir une mesure fort indirecte, mais fort significative néanmoins, de la précision réelle des chroniques analysées ; nous en tirons ainsi une présomption extrêmement encourageante, concernant les consommations mensuelles d'électricité.

Par ailleurs, nos étudiants ont pu aussi constater que notre modèle s'applique fort mal aux chroniques de production des industries légères, telle l'industrie textile. Cet échec ne nous a causé aucune peine, mais la recherche de modèles plus satisfaisants pour ces chroniques-là, nous semble proposer un problème encore ouvert.

D'autre part, nous avons indiqué que le modèle justifie une méthode d'analyse employée en France pour des consommations hebdomadaires d'électricité. C'est un argument positif, certes, mais nous sommes persuadé qu'à cette échelle de temps, des liaisons un peu plus complexes doivent se faire jour, et le raffinement du modèle dans cette direction mérite aussi quelques efforts supplémentaires.

NOTE

Cet article a été rédigé au cours de l'été 1962, en vue d'une Communication à DUBLIN, et nous indiquions à plusieurs reprises que les résultats dont il est fait état (pour ce qui concerne l'application à l'analyse de la consommation d'électricité) devaient être considérés comme provisoires. Nous avons obtenu, depuis, des chiffres de consommation électrique corrigée de l'influence des pertes de transport et des effets de la température, et si l'exploitation de ces nouvelles chroniques "raffinées" n'a pu être complètement menée à bien - car cela demande du temps et des moyens de calcul - il est dès maintenant possible d'en tirer quelques conclusions assez importantes pour que nous en fassions ici brièvement état. L'autocorrélation entre les accroissements logarithmiques mensuels de consommations s'avère différente de ce qu'elle était avec les chiffres bruts. mais contrairement à notre présomption initiale, il semble qu'elle se rapproche de zéro ! Cela doit être regardé comme un test positif de l'adéquation du modèle markovien. si l'on considère que :

1/ lorsqu'on applique un modèle markovien à une chronique résultant en réalité d'un processus classique :

$$x_t = f(t) + \varepsilon_t$$

alors on doit théoriquement trouver un coefficient d'autocorrélation précisément égal à $-\frac{1}{2}$ *entre les différences successives* $x_t - x_{t-1}$ *(pourvu que la tendance* $f(t)$ *ne s'écarte pas sensiblement d'une fonction linéaire).*

2/ l'influence des perturbations dûes aux pertes de transport et aux effets de température peut se représenter, grosso-modo, par l'introduction d'écarts aléatoires ε_t*, s'ajoutant à la chronique* $x_t = \log C_t$*, et ne se cumulant pas. Dans la mesure où la variance de ces écarts supplémentaires est appréciable devant celle des accroissements relatifs* Z_t*, leur introduction doit donc rapprocher le coefficient d'autocorrélation* ρ *de* $\frac{1}{2}$.

3/ La disparition, ou à tout le moins l'atténuation importante, de cette autocorrélation négative, lorsqu'on a éliminé l'effet des pertes et de la température, montrerait donc que le modèle markovien (avec accroissements pratiquement indépendants) s'applique bien aux consommations "purifiées" ce qui est résultat bien plus fort que ce que nous avions d'abord présumé.

En tout état de cause, nous devons souligner une fois de plus que le problème est loin d'être clos ; comme souvent dans ce domaine, l'accession à des données d'observation plus fines, permet seule de mieux discerner l'adéquation des modèles. Si des modèles markoviens nous satisfont mieux dans le cas de la consommation électrique - domaine privilégié sans doute - nous ne doutons guère que leur emploi doive s'avérer fécond pour une foule d'autres chroniques économiques.

Les perfectionnements qu'il faut leur apporter apparaîtront avec le développement des expériences pratiques, dont un certain nombre sont en cours.

BIBLIOGRAPHIE

[1] A. T. BHARUCHA-REID - Elements of the theory of Markov Processes and their applications. Mc GRAW HILL. 1960.

[2] G. Th. GUILBAUD - L'étude statistique des oscillations économiques. Cahiers du Séminaire d'Econométrie. N°1. 1951.

[3] D. JUNG - Evolution et prévision de la consommation d'énergie électrique en France. Journal de la Société de Statistique de Paris. Avril ; Mai ; Juin 1956.

[4] J. MERAUD et A. TYMEN - Les variations saisonnières de l'activité économique. Revue de Statistique appliquée. 1960 N°4.

[5] G. MORLAT - Modèle probabiliste pour la prévision de la demande d'électricité. Bulletin de l'Institut International de Statistique. 33° Session. Paris. 1961.

[6] J. SHISKIN - Decomposition of economic time series : Business fluctuations are broken down into seasonal, cyclical and irregular factors by computers. Science 19.12.58.

[7] H. WOLD - Demand analysis. Wiley. 1953.