

REVUE DE STATISTIQUE APPLIQUÉE

C. BORGARD

Traitement d'un grand nombre de séries chronologiques par des modèles. Moyenne mobile avec fonction de transfert. Application à des séries de consommation de gaz

Revue de statistique appliquée, tome 37, n° 3 (1989), p. 5-17

http://www.numdam.org/item?id=RSA_1989__37_3_5_0

© Société française de statistique, 1989, 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/>

TRAITEMENT D'UN GRAND NOMBRE DE SÉRIES CHRONOLOGIQUES PAR DES MODÈLES MOYENNE MOBILE AVEC FONCTION DE TRANSFERT APPLICATION A DES SÉRIES DE CONSOMMATION DE GAZ

C. BORGARD

GAZ DE FRANCE, 23 Rue Philibert Delorme - 75017 PARIS

1. Généralités

1.1. Introduction

On s'intéresse, ici, aux séries caractérisées par une saisonnalité très marquée. De telles séries se rencontrent fréquemment dans l'étude de phénomènes à caractère socio-économique (liés aux cycles de l'activité économique), ou proches des sciences de la Nature (liés au caractère saisonnier du climat).

Le développement de modèles saisonniers utilisant les opérateurs de différenciation a été préconisé par BOX-JENKINS (voir [1], chapitre 9) pour étudier ce type de séries. Récemment, le champ d'application de ces modèles a été étendu aux consommations d'énergie : consommations de gaz ou d'électricité [6], aux installations d'équipements collectifs [3], aux ventes de produits de grande consommation, à certaines séries concernant la main d'oeuvre ou l'emploi [5], etc...

Dans l'exposé qui suit on se placera dans le cas de données mensuelles; la périodicité de la série est alors $S = 12$. Les formules présentées pourront se transposer facilement si l'on traite de données journalières ou de données horaires en remplaçant $S = 12$ respectivement par $S = 7$ et $S = 24$.

Notons une telle série N_t

Sous la forme canonique, le modèle multiplicatif saisonnier s'écrit :

$$\nabla \nabla_s N_t = \theta_1(B) \cdot \theta_2(B^S) \epsilon_t$$

où θ_1 et θ_2 sont des polynômes respectivement en B et B^S dont les racines sont toutes extérieures au cercle unité.

En pratique, on retient généralement l'expression :

$$\nabla \nabla_s N_t = (1 - \alpha_1 B)(1 - \alpha_S B^S) \epsilon_t$$

qui s'écrit ici :

$$\nabla \nabla_{12} N_t = (1 - \alpha_1 B)(1 - \alpha_{12} B^{12}) \epsilon_t \quad (1)$$

où ϵ_t est un bruit blanc

B, B^{12} sont les opérateurs de retard d'ordre 1 et d'ordre 12.

∇, ∇_{12} sont les opérateurs de différenciation à l'ordre 1 et à l'ordre 12

$$|\alpha_1| < 1 \text{ et } |\alpha_{12}| < 1$$

Ce modèle a été appliqué aux consommations mensuelles de gaz. Ces consommations sont fortement influencées par des variables d'environnement, principalement les variables climatiques. On a donc complété le modèle multiplicatif saisonnier par l'adjonction d'une fonction de transfert.

Le système décrit ici est exploité de façon industrielle sur une centaine de séries depuis plus de sept ans. La présentation de ce système est avant tout descriptive puisqu'elle rend compte d'une expérience; elle s'adresse donc d'abord aux praticiens.

On remarque toutefois que l'exemple traité ici n'est qu'un cas particulier d'un problème beaucoup plus général qui pourrait s'énoncer : comment une chronique saisonnière peut elle être expliquée par d'autres chroniques qui présentent la même saisonnalité ?

On retrace d'abord le cheminement qui a conduit à retenir les modèles ARIMA à fonction de transfert (paragraphe 2), puis on décrit le système actuellement en exploitation (paragraphe 3). On examine ensuite les apports concrets de la modélisation en processus du point de vue du contrôle de la fiabilité des données, de la stabilité des modèles (paragraphe 4) et enfin de la qualité statistique des estimations (paragraphe 5).

2. Une expérience de modélisation

2.1 Vers les modèles ARIMA, Pourquoi ?

La variable étudiée C_t représente des consommations d'énergie en vue d'assurer en grande partie des besoins de chauffage : cette variable dépend fortement de la variable naturelle θ_t température observée au temps t . On remarque que les variables θ_t et C_t sont auto-corrélées (inertie thermique pour la première et inertie des comportements des consommateurs pour la seconde) et, de plus, que la liaison entre ces deux variables subit des déformations au cours du temps.

Des études préliminaires s'appuyant sur l'Analyse de Données et la Ridge-régression ont permis de sélectionner quelles autres variables climatiques peuvent influencer sur les consommations; finalement on a retenu en plus de la température observée à la station météorologique le rayonnement solaire global R_t évalué par la Météorologie Nationale.

Les premiers modèles cherchant à relier C_t aux variables θ_t et R_t étaient des modèles régressifs qui ne pouvaient fonctionner que sur des courtes périodes (l'année au maximum) et on ne parvenait pas à se débarrasser totalement des auto-corrélations.

On a donc cherché à établir des modèles à fonction de transfert dont l'intérêt apparait de la façon suivante :

Lorsque l'on a deux variables $X(t)$ et $Y(t)$, centrées, stationnaires, et reliées entre elles par une liaison linéaire de la forme :

$$Y(t) = V_0 X(t) + V_1 X(t-1) + \dots + V_p X(t-p) = V(B)X(t)$$

et que l'on souhaite estimer les coefficients V_0, V_1, \dots, V_p , on peut être tenté de calculer le coefficient de la régression de $Y(t)$ en $X(t)$. En désignant par $\rho(h)$ le coefficient d'auto-corrélation d'ordre h de $X(t)$, on a :

$$\frac{E[Y(t) \cdot X(t)]}{E[X(t)^2]} = V_0 + V_1 \rho(1) + \dots + V_h \rho(h) + \dots + V_p \rho(p)$$

La régression de $Y(t)$ en $X(t)$ ne permet donc pas d'estimer directement les V_h . Considérons, par contre, la variable $Z(t)$ construite à partir de $Y(t)$ en lui appliquant l'opérateur $\psi^{-1}(B)$ tel que $X(t) = \psi(B) \cdot a_t$ où a_t désigne un bruit blanc.

Alors : $Z(t) = \psi^{-1}(B) \cdot V(B) \cdot X(t) = V(B)a_t$ et $E[Z(t) \cdot a(t-h)] = E[V(B)a(t)a(t-h)]$

d'où :

$$V_h = \frac{E[Z(t) \cdot a(t-h)]}{E[a(t)]^2}$$

On obtient les coefficients V_h non nuls de la liaison linéaire en calculant les corrélations croisées entre $Z(t)$ et $a(t-h)$. En pratique, pour s'assurer de la significativité des V_h , on pourra être amené à étudier la corrélation croisée entre les bruits blancs de $X(t)$ et $Y(t)$.

2.2. Formulation du modèle à fonction de transfert

Pour stationnariser les variables C_t et R_t on a appliqué la transformation logarithmique et pour chacune des trois variables les opérateurs de double différenciation. On obtient le modèle :

$$\nabla \nabla_{12} \text{Log} C_t = W_0 \nabla \nabla_{12} \theta_t + W'_0 \nabla \nabla_{12} \text{Log} R_t + (1 - \alpha_1 B)(1 - \alpha_{12} B^{12}) \epsilon_t$$

ce qui correspond avant différenciation, compte tenu des conditions initiales et de la relation (1) à :

$$\text{Log} C_t = W_0 \theta_c + W'_0 \text{Log} R_t + N_t \tag{2}$$

Globalement ce modèle est satisfaisant, car les critères traditionnels retenus pour valider les conditions d'application de la méthode "BOX-JENKINS" sont vérifiés (le résidu ϵ_t est bien un bruit blanc, les estimations des coefficients $W_0, W'_0, \alpha_1, \alpha_{12}$ sont peu corrélées). Mais, pour chaque mois, on est amené à introduire

des paramètres spécifiques. On définit pour cela 12 variables température et 12 variables rayonnement solaire de la façon suivante :

$$\theta_{mt} = \theta_t \text{ si } t \text{ correspond au mois d'indice } m \text{ (} m = 1, 2, \dots, 12 \text{)}$$

$$= 0 \text{ sinon}$$

$$R_{mt} = R_t \text{ si } t \text{ correspond au mois d'indice } m$$

$$= 0 \text{ sinon}$$

En regroupant les mois consécutifs pour lesquels les paramètres W_o et W'_o ne sont pas significativement différents on dégage l'existence de quatre périodes homogènes, appelées saisons :

Les mois de Novembre à Mars constituent	: l'hiver
Les mois d'Avril et Mai constituent	: le printemps
Les mois de Juin à Août constituent	: l'été
Les mois de Septembre et Octobre constituent	: l'automne

3. Exploitation

La mise au point et les essais des modèles se sont déroulés entre 1977 et 1980; puis un système d'exploitation a été mis en place dans le courant de l'année 1981. Ce système d'exploitation comprend une base de données et une base de modèles.

3.1. Alimentation de la base de données

Les consommations mensuelles sont collectées auprès d'environ 2000 points de mesures. A l'aide de tables de correspondance on aboutit après agrégation à un ensemble réduit de 100 unités de traitement. A chaque unité de traitement est associée une station météorologique de référence choisie dans un ensemble de 36 stations possibles.

3.2. Fonctions de la base de modèles

A chaque unité de traitement correspond un modèle. La base de modèles est constituée des paramètres spécifiques de chacun des 100 modèles.

Ces modèles sont utilisés pour assurer en tout ou partie les opérations suivantes :

- contrôle de qualité des données,
- maintenance de modèles associés,
- estimation des consommations sous des références climatiques données,
- estimation des prévisions et de leurs intervalles de confiance.

4. Amélioration de la qualité numérique

On examine maintenant quels sont les principaux apports de la modélisation en processus du point de vue de la qualité des estimations statistiques.

4.1. Contrôle des données

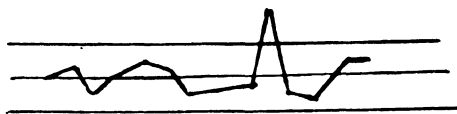
Les données mensuelles provenant d'environ 2000 points source font préalablement l'objet de bilans simplifiés consistant à vérifier, au niveau de grandes régions, dans une limite de précision tolérable que les équations de conservation des flux sont bien respectées. Ce premier filtrage reste assez imprécis; il est suivi d'opérations de contrôle de qualité des données au niveau de chaque unité de traitement. Le fondement de ces opérations est le suivant :

On a vu que chaque série présentait une forte inertie et se trouvait donc pourvue d'une assez bonne mémoire. Chaque observation ne doit donc pas trop s'écarter de sa position naturelle et s'il y a des accidents de trajectoire ils sont alors facilement détectables. Pour cela, on peut par exemple rechercher les valeurs de la série qui s'écartent de leur valeur attendue d'un nombre fixé d'écarts-types.

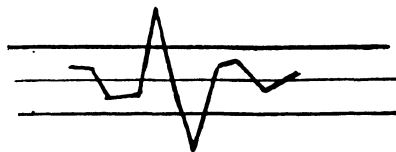
Les résidus anormaux demandent à être analysés au fil du temps; les principales causes d'écarts importants sont les suivantes :

- a) Des erreurs de données sur la variable à expliquer provenant soit d'erreur de relevé des appareils de mesure, de transcription ou de transmission. Nous incluons dans cette rubrique le cas où les relèves ne se font pas aux dates prescrites par le calendrier (la valeur mensuelle ne recouvre pas exactement le mois calendaire : si, par exemple, elle déborde le mois t de n jours, la valeur en $t + 1n$ inclura pas ces n jours) et aussi le cas des données manquantes, heureusement fort rare lorsque l'on travaille sur des séries mensuelles.
- b) Cas d'une erreur de donnée sur les variables explicatives. C'est le cas où l'on a une erreur sur l'une des variables climatiques transmises. Ce cas est très rare.
- c) Il y a un décalage dans les saisons; par exemple l'une des saisons se prolonge un peu plus tard que d'habitude.
- d) Il y a un changement dans le rythme de développement de la variable à étudier, par exemple à la suite d'une action commerciale plus soutenue.
- e) Il y a un changement conjoncturel de comportement de la clientèle.

Les cas a) et b) sont dits "accidentels". La série de résidus décroche brusquement et retrouve généralement assez vite son profil habituel :



cas a)
erreur de transmission ou cas b)



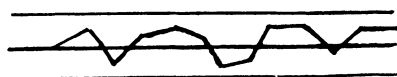
cas a)
"problème d'horloge"

Le cas c) est détecté rapidement grâce à l'exploitation de modèles "pilote" travaillant sur des données journalières. Les données du jour j sont connues en $j + 2$. L'expérience montre qu'il faut attendre cinq jours pour confirmer une date de changement de saison.

Le cas d) se résorbe après que l'on ait réactualisé les paramètres du modèle, généralement en fin de saison :

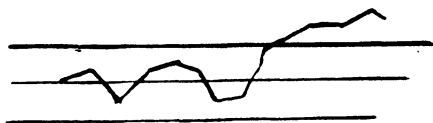


cas c) ou d)
avant réactualisation

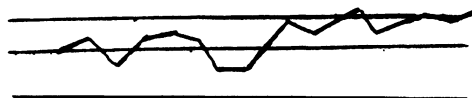


cas c) ou d)
après réactualisation

Le cas e) est plus complexe. Malgré une réactualisation des paramètres les écarts résiduels ne retrouvent pas une distribution proche du hasard. On observe des écarts de même signe pendant plusieurs mois consécutifs pouvant recouvrir même plusieurs saisons :



cas e)
avant réactualisation



cas e)
après réactualisation

Si le cas e) n'apparaît que sur une série isolée, il s'agit d'un phénomène transitoire lié à l'apparition ou au contraire à la perte d'un client qui à lui seul représente une part importante de la consommation totale.

Si la grande majorité des séries se trouve ainsi affectée il s'agit alors d'un changement conjoncturel du comportement de la clientèle. La structure des consommations est soudainement modifiée mais le modèle ne peut pas répartir la variation de consommation entre une part relative aux conditions climatiques et une part qui ne l'est pas. Ce phénomène conjoncturel peut être traité par l'introduction de variables d'intervention.

4.2. Réactualisation des paramètres des modèles

4.2.1. Pourquoi réactualiser les paramètres ?

Les paramètres des modèles évoluent au cours des années car les besoins énergétiques s'adaptent à des situations économiques différentes : matériels de production de chaleur plus performants, développement de l'isolation des locaux, mise en place de systèmes de régulation thermique, ... Cette évolution se fait progressivement selon un rythme saisonnier car les travaux sont généralement réalisés en dehors des périodes de chauffage.

4.2.2. Quand faut-il réactualiser ?

Il y a des moments privilégiés pour réactualiser les paramètres : ce sont les fins de saison. La prise en compte des données de toute une saison peut apporter des indications sur l'évolution de certains paramètres. Ces informations sont d'abord qualitatives : confirmation ou non des tendances éventuelles constatées dans le passé. Elles conduisent généralement à des prévisions mieux adaptées.

4.2.3. Qualité des ajustements successifs

Les premiers réajustements des modèles se sont accompagnés d'une diminution sensible des variances résiduelles et après deux années d'exploitation on a observé une stabilisation.

4.2.4. Evolution du paramètre W_0 de la fonction de transfert entre 1982 et 1985

Selon le sens de l'évolution de W_0 , les séries se répartissent comme suit :

SAISON	SENS DE L'EVOLUTION DE W_0		
	Décroissance	Stabilité	Croissance
HIVER	10	65	25
PRINTEMPS	70	25	5
ETE	40	40	20
AUTOMNE	40	40	20

Le paramètre de printemps a subi une forte décroissance et a rejoint la valeur moyenne du paramètre d'automne. Ainsi, on peut admettre qu'une variation de un degré de la température entraîne maintenant la même variation relative de

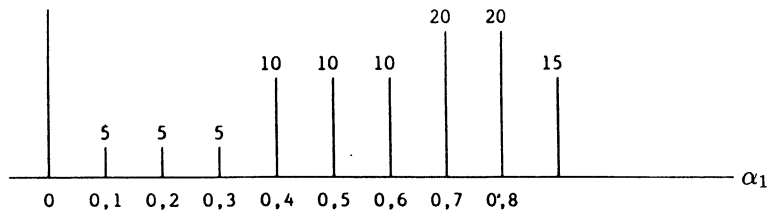
consommation au printemps ou en automne. L'habitude qui consistait à laisser le chauffage fonctionner au printemps plus tard qu'il n'était besoin se serait donc considérablement modifiée au cours de ces années. Depuis 1985, on note une très grande stabilité des paramètres.

4.2.5. Valeurs des paramètres de la fonction bruit $(1 - \alpha_1 B)(1 - \alpha_{12} B^{12})\epsilon_t$

La fonction bruit est caractérisée par les deux paramètres α_1 et α_{12}

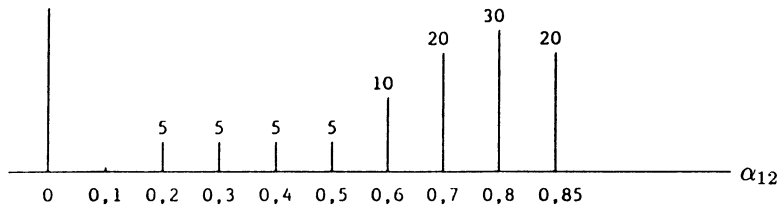
. Distribution du paramètre α_1

Nombre
de séries en %



. Distribution du paramètre α_{12}

Nombre
de séries en %



Les paramètres α_1 et α_{12} caractérisent l'allure de la trajectoire de chaque série. Il est intéressant de suivre, en pratique, l'évolution de ces paramètres. On a observé que les courbes de répartition données ci-dessus évoluent peu, avec toutefois une légère translation vers les valeurs croissantes pour α_1 , vers les valeurs décroissantes pour α_{12} .

5. Connaissance des réalisations

5.1. Estimation de la consommation d'une région pour un climat de référence donné

Pour un mois donné t , on cherche la consommation qui se serait produite si l'on avait observé un climat de référence défini par des valeurs $\bar{\theta}_t$ et \bar{R}_t . On utilise le modèle (2) avec les estimations de W_o et W'_o pour calculer la consommation correspondante \bar{C}_t .

C_t, θ_t, R_t et N_t sont reliés par

$$\text{Log}C_t = W_o\theta_t + W'_o\text{Log}R_t + N_t \quad (2)$$

En pratique, on construit des estimateurs sans biais pour W_o, W'_o, α_1 et α_{12} et à l'instant t on dispose de leurs valeurs numériques

$$\bar{W}_o, \bar{W}'_o, \bar{\alpha}_1 \text{ et } \bar{\alpha}_{12}$$

on pose

$$\bar{N}_t = \text{Log}C_t - \bar{W}_o\theta_t - \bar{W}'_o\text{Log}R_t \quad (2')$$

et on a alors comme estimation de $\text{Log}\bar{C}_t$ la valeur $\widetilde{\text{Log}}\bar{C}_t$ donnée par :

$$\widetilde{\text{Log}}\bar{C}_t = \bar{W}_o\bar{\theta}_t + \bar{W}'_o\text{Log}\bar{R}_t + \bar{N}_t \quad (2'')$$

La vraie valeur de \bar{C} est en fait donnée par la relation (2) qui s'écrit

$$\text{Log}\bar{C}_t = W_o\bar{\theta}_t + W'_o\text{Log}\bar{R}_t + N_t \quad (2''')$$

Les formules (2), (2'), (2'') et (2''') permettent de calculer l'erreur sur l'estimation de $\text{Log}\bar{C}_t$. On a :

$$\begin{aligned} \widetilde{\text{Log}}\bar{C}_t - \text{Log}\bar{C}_t &= (\bar{W}_o - W_o)(\bar{\theta}_t - \theta_t) + (\bar{W}'_o - W'_o)\text{Log}\frac{\bar{R}_t}{R_t} \\ \Delta\text{Log}\bar{C}_t &= \frac{\Delta\bar{C}_t}{\bar{C}_t} = (\bar{\theta}_t - \theta_t) \Delta W_o + \text{Log}\frac{\bar{R}_t}{R_t} \Delta W'_o = Z_t \end{aligned}$$

Z_t est une variable aléatoire de moyenne nulle. Sa variance devrait dépendre du mois étudié. Mais en pratique on remarque que la variance de ΔW_o dépend de la saison. Elle est faible en hiver, là où la dispersion des variables climatiques est importante. On admet donc, en première approximation, que la variance σ^2 de Z_t est constante, quel que soit le mois considéré : $\sigma^2 \approx 10^{-4}$.

Lors de la construction du modèle on s'est assuré que \bar{W}_o et \bar{W}'_o étaient indépendantes de N_t et suivaient des lois normales. Z_t est donc une variable aléatoire normale indépendante de N_t .

Lorsque $\bar{\theta}_t$ et \bar{R}_t sont donnés, la relation (2''') permet d'écrire $\text{Log} \bar{C}_t = \text{constante} + N_t$

Les variables \bar{C}_t et Z_t sont donc indépendantes et l'on a $E(\Delta \bar{C}_t) = 0$

$\text{Var}(\Delta \bar{C}_t) = E(\bar{C}_t^2) \text{Var}(Z_t)$ dont l'ordre de grandeur est $\bar{C}_t^2 \sigma^2$.

L'erreur $\Delta \bar{C}_t$ faite sur l'estimation de \bar{C}_t suit donc approximativement une loi normale de moyenne nulle et d'écart-type $\bar{C}_t \sigma$.

$$\Delta \bar{C}_t \in \text{L.N.}(0, \bar{C}_t \sigma)$$

5.2. Erreur sur la somme de consommations

Considérons n consommations C_1, C_2, \dots, C_n et désignons par $\bar{C}_1, \bar{C}_2, \dots, \bar{C}_n$ les valeurs correspondant au climat choisi pour référence.

Formons :

$$C = \sum_i C_i$$

$$\bar{C} = \sum_i \bar{C}_i$$

$$\Delta \bar{C} = \sum_i \Delta \bar{C}_i$$

On rencontre ce problème lorsque pour un mois donné on cherche la consommation totale de N régions ou lorsque pour une région donnée on cherche la consommation cumulée sur n mois.

$\Delta \bar{C}$ suit approximativement une loi normale dont la moyenne est nulle et dont la variance s'écrit $\sum_i E(\bar{C}_i^2) \sigma^2 = E(\sum_i \bar{C}_i^2) \sigma^2$ quantité approximativement égale à $\sum_i \bar{C}_i^2 \sigma^2$ qui peut toujours être majorée par $\bar{C}^2 \sigma^2$.

On a donc, en première approximation :

$$\frac{\Delta \bar{C}}{\bar{C}} \in \text{L.N.}(0, \sigma') \text{ avec } \sigma' < \sigma$$

Remarque :

La dispersion des \bar{C}_i provient surtout des consommations pour le chauffage qui sont très faibles en dehors de la période hivernale. Ainsi, en dehors de l'hiver, si on cumule les consommations sur N régions, on admettra que $\bar{C}_i \# \frac{\bar{C}}{N}$ et donc $\text{Var}(\bar{\Delta C}) \# \sum C_i^{-2} \sigma^2 \# \bar{C}^2 \frac{\sigma^2}{N}$

Sur cette période on prendra pour approximation de la loi de $\frac{\Delta \bar{C}}{\bar{C}}$ une loi

normale de moyenne nulle et d'écart type $\frac{\sigma}{\sqrt{N}}$

$$\frac{\Delta \bar{C}}{\bar{C}} \epsilon L.N(O, \frac{\sigma}{\sqrt{N}})$$

5.3 Application : Estimation de la consommation totale de N régions cumulée sur n mois

Désignons par \bar{C}_i^j la consommation du mois i dans la région j et par \bar{C}_i^N la consommation totale des N régions le mois i

$$\bar{C} = \sum_{j=1}^N \sum_{i=1}^n \bar{C}_i^j = \sum_{i=1}^n \bar{C}_i^N$$

On calcule l'erreur sur la consommation \bar{C} lorsque l'on connaît l'erreur sur la consommation totale des N régions pour un mois d'hiver, notée a . En pratique, a est inférieur à 1 % .

Application numérique : $N = 100$

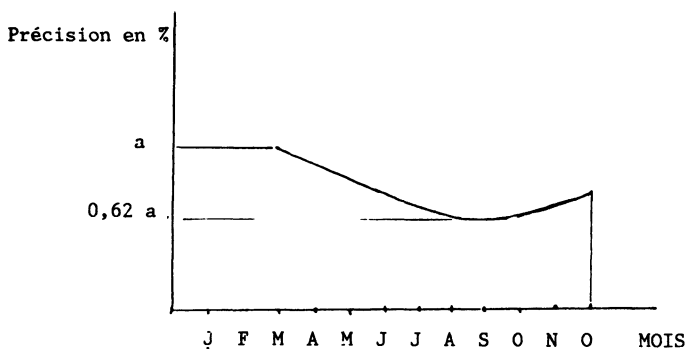
CONSOMMATIONS MENSUELLES

1	MOIS i	1	2	3	4	5	6	7	8	9	10	11	12
2	$\frac{\bar{C}_i^N}{\bar{C}}$.16	.14	.12	.08	.06	.03	.02	.02	.03	.07	.11	.16
3	$\frac{ \Delta \bar{C}_i^N }{\bar{C}_i^N}$	a	a	a	$a/10$	$a/10$	$a/10$	$a/10$	$a/10$	$a/10$	$a/10$	a	a
4	$\frac{ \Delta \bar{C}_i^N }{a\bar{C}}$.16	.14	.12	.008	.006	.003	.002	.002	.003	.007	.11	.16

CONSOMMATIONS CUMULEES

5	CUMUL JUSQU'À n	1	2	3	4	5	6	7	8	9	10	11	12
6	$\frac{\sum_{i=1}^n \bar{C}_i^N}{\bar{C}}$.16	.30	.42	.50	.56	.59	.61	.63	.66	.73	.84	1.
7	$\frac{\sum_{i=1}^n \Delta \bar{C}_i^N }{a\bar{C}}$.16	.30	.42	.428	.434	.437	.439	.441	.444	.451	.561	.721
8	$\frac{\sum_{i=1}^n \Delta \bar{C}_i^N }{\sum_{i=1}^n \bar{C}_i^N}$	α	α	α	.86 α	.78 α	.74 α	.72 α	.70 α	.67 α	.62 α	.67 α	.72 α

La précision de l'estimation de la consommation totale cumulée sur n mois varié en cours d'année comme l'indique le graphique suivant :



On remarque que, dans ces conditions, la précision est optimale en octobre et qu'elle varie peu autour de cette valeur optimale les six derniers mois de l'année.

6 Conclusion

Le système d'exploitation qui vient d'être décrit sert, en pratique, à suivre l'évolution de séries mensuelles en cours d'année. C'est donc bien d'abord un outil de gestion. Chaque mois il permet de contrôler les données reçues et de calculer les consommations passées et prévues associés à des conditions d'environnement prédéterminées.

Ces estimations sont connues avec leurs intervalles de confiance, ce qui permet d'assurer la cohérence entre les prévisions et les réalisations.

Mais les apports d'un tel système d'exploitation sont beaucoup plus riches si l'on mémorise les résultats obtenus pour chaque série. On peut, en effet, suivre les déformations des paramètres au cours des réestimations successives et en tirer des enseignements sur l'évolution de la structure de chaque série.

On peut, aussi, en recoupant les résultats des estimations de toutes les séries dégager les spécificités de certains groupes de séries caractéristiques de comportements régionaux très marqués.

Finalement, l'exploitation d'un tel système, conçu pour gérer le présent, informe aussi sur les tendances profondes des séries et fournit des indications sur leur évolution à long terme.

Bibliographie

- [1] BOX-JENKINS (1976) *Time Series Analysis - Forecasting and control Holden-day.*
- [2] G. JENKINS (1981) *Case Studies in time series analysis - GJP Publication.*
- [3] O. ANDERSON (1980) *Forecasting Public Utilities - North Holland.*
- [4] G. MAC LEOD (1983) *Box-Jenkins in practice - GJP Publication.*
- [5] R. MAC CLEARY (1980) *Applied time series analysis for social sciences – Sage Publications.*
- [6] C. BORGARD & G. GEORGE (1982) *Une nouvelle modélisation des émissions de gaz. Congrès ATG.*