C A H I E R D U L A M S A D E

Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision

(Université Paris-Dauphine)

Equipe de Recherche Associée au C.N.R.S. N° 656

UNE APPROCHE SIMPLIFIEE DE LA METHODE DE BOX & JENKINS POUR L'ANALYSE ET LA PREVISION DES SERIES TEMPORELLES UNIDIMENSIONNELLES

J. ABADIE
D. TRAVERS

N° 26-1980 Janvier 1980

J. ABADIE et D. TRAVERS

UNE APPROCHE SIMPLIFIEE DE LA METHODE DE BOX ET JENKINS POUR L'ANALYSE ET LA PREVISION DES SERIES TEMPORELLES UNIDIMENSIONNELLES

Résumé

L'analyse des séries temporelles unidimensionnelles par la méthode de BOX & JENKINS comporte trois étapes : identification, estimation, validation.

La phase initiale d'identification est souvent difficile dans la pratique. Nous proposons ici une approche interactive simplifiée, reposant sur des manipulation algébriques simples (factorisation de polynômes, recherche de racines communes ou voisines de deux polynômes, etc.) Cela permet de placer la méthode à la portée de l'utilisateur moyen.

J. ABADIE and D. TRAVERS

A SIMPLIFIED APPROACH OF THE BOX-JENKINS METHOD FOR UNIDIMENSIONAL TIME-SERIES ANALYSIS AND FORECASTING

Abstract

Time Series Analysis by the BOX-JENKINS method involves three stages : identification, estimation,validation.

The initial identification stage is often difficult in practice. We propose in this paper a simplified interactive approach based on simple algebraic manipulations (factorization and roots of polynomes,etc).This enables the method to be handled by the average user.

PLAN

1. - PRESENTATION GENERALE

L'analyse des séries temporelles unidimensionnelles par la méthode de Box & Jenkins comporte traditionnellement trois étapes :

- Identification de la forme de modèle la mieux adaptée à la série étudiée;

- Estimation des paramètres du modèle;

- Validation : le modèle retenu convient-il bien ? Sinon pourquoi ? et comment l'améliorer.

Sous cette forme, la méthode est déjà d'un emploi très général et son intérêt n'est plus à démontrer. Pourtant, la phase initiale d'identification se révèle, en pratique, être une étape assez délicate à franchir. Elle repose sur l'analyse des fonctions d'autocorrélation et d'autocorrélation partielle de la série et il arrive malheureusement assez souvent que les corrélogrammes observés soient relativement peu typiques. Une grande expérience est alors nécessaire pour reconnaître le modèle théorique auquel ils s'apparentent, surtout dans le cas saisonnier.

Nous proposons ici une approche simplifiée qui doit permettre de placer la méthode à la portée de l'utilisateur moyen, tel qu'on peut le rencontrer dans le contexte de la gestion des entreprises, et non plus seulement du statisticien spécialisé. Dans cette optique, l'analyse des fonctions d'autocorrélation et d'autocorrélation partielle est remplacée par des manipulations algébriques simples sur les opérateurs du modèle (factorisation de polynômes, recherche de racines communes ou voisines, etc.).

Cette approche a été rendue possible par la mise au point d'une nouvelle méthode d'estimation (cf. § 3) qui utilise une méthode moderne d'optimisation et dans laquelle les premiers résidus sont tout simplement considérés comme des paramètres supplémentaires à estimer simultanément avec les paramètres proprement dits du modèle. Contrairement à la méthode habituelle d'estimation (estimation avec "prévisions à rebours", cf. § 3.2.2), cette méthode ne fait pas appel à la réversibilité du

modèle et peut donc s'appliquer même à des modèles très mal identifiés, notamment en ce qui concerne le degré de différenciation.

Le schéma général de la méthode de Box & Jenkins se trouve alors profondément modifié puisque l'identification n'est plus un préalable nécessaire à l'estimation. Bien au contraire, c'est l'estimation successive de différents modèles qui peut permettre de préciser peu à peu le modèle à retenir en définitive (cf. § 4). Les deux étapes identification et estimation ne constituent plus ainsi qu'une étape complexe unique que nous appelons "choix du modèle".

Cette étape "choix du modèle" rend la méthode beaucoup plus accessible et souple. Elle conduit progressivement à l'identification et à l'estimation du modèle ARIMA ou SARIMA le mieux adapté. Quant à l'analyse, si délicate, des autocorrélations et autocorrélations partielles, elle est maintenant réservée à l'étape de la validation. Elle est alors beaucoup plus facile à réaliser puisqu'il ne s'agit plus que de vérifier que la série des résidus possède bien les caractéristiques d'un bruit blanc (cf. § 5).

Enfin, nous n'avons pas voulu perdre de vue que l'objectif ultime de toute analyse de série temporelle reste la prévision. Nous avons donc examiné les conséquences du choix d'un modèle sur la prévision (cf. § 6).

Dans un souci de clarté et de brièveté, notre approche est illustrée par un nombre restreint d'exemples, mais on trouvera en annexe une revue rapide des modèles retenus pour un plus grand nombre de séries.

2. - RAPPELS RAPIDES SUR LES MODELES UTILISES

2.1. - Modèles ARIMA pour séries non-saisonnières

Pour représenter une série temporelle non-saisonnière $\{z_t\} = \{z_1, z_2 \ldots z_n\}$, on peut envisager un modèle ARIMA (p, d, q) de la forme :

$$\phi(B)\ w_t = \theta(B)\ a_t \qquad (2.1.1)$$

avec :

$$\phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 \ldots - \varphi_p B^p$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 \ldots - \theta_q B^q$$

$$w_t = \begin{cases} \tilde{z}_t = z_t - \bar{z} & \text{si } d = 0 \\ \nabla^d z_t & \text{si } d > 0 \end{cases} \qquad (\{w_t\} : \text{série stationnaire})$$

$$Bz_t = z_{t-1} \qquad (B : \text{opérateur de retard})$$

$$\nabla z_t = (1 - B) z_t = z_t - z_{t-1} \qquad (\nabla : \text{opérateur de différence})$$

$$a_t \sim \eta(0, \sigma_a^2) \qquad (\text{résidus aléatoires, normaux et indépendants})$$

En pratique, des valeurs modérées de p, d et q suffisent pour les séries souvent courtes (cf. § 3.1 pour la longueur minimum souhaitable) que l'on rencontre en économie et en gestion : $d \leq 2$; $p + q \leq 3$ ou parfois 4.

Le modèle ARIMA (2,d,2) s'écrit, par exemple :

$$(1 - \varphi_1 B - \varphi_2 B^2) w_t = (1 - \theta_1 B - \theta_2 B^2) a_t \qquad (2.1.2)$$

soit

$$w_t = \varphi_1 w_{t-1} + \varphi_2 w_{t-2} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \qquad (2.1.3)$$

Parfois, il peut être utile d'introduire une constante θ_o et le modèle (2.1.2) devient le modèle ARIMA (p,d,q) $+ c^{te}$:

$$\phi(B) w_t = \theta_o + \theta(B) a_t \qquad (2.1.4)$$

2.2. - Tranformation initiale

Dans certains cas, l'opérateur ∇ n'est pas suffisant pour transformer la série $\{z_t\}$ en une série $\{w_t\}$ stationnaire, et on peut alors utiliser au préalable une transformation non linéaire du type

logarithme ou fonction puissance* . Mais il ne faut pas oublier de corriger le biais introduit au stade de la prévision**.

2.3. - Stationnarité et inversibilité

L'utilisation des polynômes $\phi(B)$ et $\theta(B)$ permet d'écrire facilement le modèle (2.1.1) sous forme de filtre linéaire :

$$z_t = \Psi(B)\ a_t \qquad (2.3.1)$$

avec $$\Psi(B) = \frac{\theta(B)}{\phi(B)\ \nabla^d} = 1 - \psi_1 B - \psi_2 B^2 - \ldots \qquad (2.3.2)$$

ou sous forme inversée :

$$\pi(B)\ \nabla^d\ z_t = a_t \qquad (2.3.3)$$

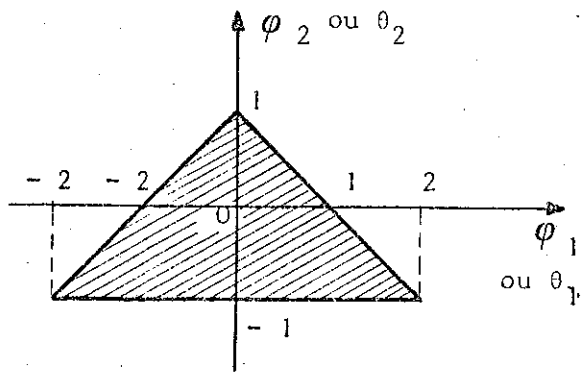
avec $$\pi(B) = \frac{\phi(B)}{\theta(B)} = 1 - \pi_1 B - \pi_2 B^2 - \ldots \qquad (2.3.4)$$

Le passage entre les formes (2.1.1) et (2.3.1) ou (2.3.3) du modèle ARIMA suppose que les paramètres φ_i et θ_i remplissent respectivement, les conditions de stationnarité et d'inversibilité, c'est-à-dire que les racines des polynômes $\phi(B)$ et $\theta(B)$ soient situées à l'extérieur du cercle unité dans le plan complexe.

Pour un modèle ARIMA (2,d,2) on a, par exemple :

Domaine de stationnarité :

$$\begin{cases} |\varphi_2| < 1 \\ \varphi_2 + \varphi_1 < 1 \\ \varphi_2 - \varphi_1 < 1 \end{cases}$$



* BOX & COX, 1964

** GRANGER & NEWBOLD, 1976

Domaine d'inversibilité .

$$\begin{cases} |\theta_2| < 1 \\ \theta_2 + \theta_1 < 1 \\ \theta_2 - \theta_1 < 1 \end{cases}$$

Les identités (2.3.2) et (2.3.4) permettent alors d'établir des relations récursives entre les coefficients φ_i, θ_i, ψ_i et π_i.

2.4. - Parcimonie

Il est utile de considérer des modèles ayant le plus petit nombre possible de paramètres à estimer (Principe de parcimonie).

2.5. - Modèles SARIMA pour séries saisonnières

Pour les séries saisonnières, on définit un modèle multiplicatif SARIMA $(p,d,q) \times (P,D,Q)_s$:

$$\phi(B)\ \Phi(B^s)\ \nabla^d\ \nabla_s^D\ z_t = \theta(B)\ \Theta(B^s)\ a_t \qquad (2.5.1)$$

avec

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \ldots - \Phi_P B^{Ps}$$

$$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \ldots - \Theta_Q B^{Qs}$$

$$\nabla_s = 1 - B^s \quad \text{(opérateur de différence saisonnière)}$$

Tout ce qui a été dit pour les modèles non saisonniers s'applique encore ici, notamment ce qui concerne l'introduction éventuelle d'une constante θ_o et les conditions de stationnarité [pour $\Phi(B^s)$] et d'inversibilité [pour $\Theta(B^s)$].

2.6. - Séries doublement saisonnières

Des séries avec deux périodicités se rencontrent parfois : périodicité hebdomadaire et annuelle, par exemple, dans la consommation

journalière d'énergie électrique* . Ces séries comportent alors souvent des perturbations, dues à la présence de fêtes fixes ou mobiles et aux périodes de vacances (mois d'août). Nous ne les aborderons pas ici.

3. - ESTIMATION DES PARAMETRES

3.1. - Moindres carrés

Le principe du maximum de vraisemblance permet d'estimer les paramètres du modèle le mieux adapté pour représenter une série donnée $\{z_t\}$, ceci au sein d'une classe particulière de modèles ARIMA ou SARIMA, c'est-à-dire pour p,d,q, P,D,Q,s fixés.

Pour des séries suffisamment longues ($N \geq 40$ dans le cas non-saisonnier ou 60 dans le cas saisonnier), ce "meilleur modèle" peut être obtenu par la méthode des moindres carrés. Le problème à résoudre est le suivant :

$$\left| \begin{array}{l} \min S\ (x|w) \\ \\ \text{avec } S = \Sigma\ a_t^2 \end{array} \right. \qquad (P)$$

où w représente la série différenciée et x le vecteur des paramètres :

$$w = \{w_1, w_2 \ldots w_n\} \qquad n = N - d - s\,D$$

$$x = (\varphi_1 \ldots \varphi_p, \theta_o, \theta_1 \ldots \theta_q, \Phi_1 \ldots \Phi_P, \Theta_1 \ldots \Theta_Q)$$

3.2. - Le problème de l'initialisation

Dans la somme S, les résidus sont calculés par récurrence, par exemple à partir de la formule suivante, dans le cas non-saisonnier :

$$a_t = w_t - \varphi_1 w_{t-1} \ldots - \varphi_p w_{t-p} - \theta_o + \theta_1 a_{t-1} \ldots + \theta_q a_{t-q} \qquad (3.2.1)$$

---

* ABADIE & MESLIER, 1977 et 1979; MESLIER, 1976

On voit tout de suite que la résolution du problème (P) nécessite une évaluation des premiers résidus a_* nécessaires à l'initialisation de l'équation de récurrence (3.2.1).

3.2.1. - Premiers résidus nuls

Une première méthode consiste à considérer ces premiers résidus nuls et à résoudre le problème approché

$$\min S\ (x \mid w,\ a_* = 0) \qquad (P_o)$$

Mais la perturbation introduite n'est pas toujours suffisamment vite absorbée au travers de l'équation de récurrence, surtout pour les séries saisonnières.

3.2.2. - Prévisions à rebours

Une deuxième méthode consiste à effectuer des sortes de "prévisions à rebours" ("back forecasting" ou "backcasting") à l'aide du même modèle écrit en remontant dans le temps avec l'opérateur d'avance F ($Fz_t = z_{t+1}$). Dans le cas non-saisonnier on écrit, par exemple :

$$\phi(F)\ w_t = \theta(F)\ e_t \qquad (3.2.2)$$

On effectue alors une suite de prévisions en arrière et en avant. Le processus converge normalement assez rapidement et permet d'estimer les premiers résidus a_*. Le problème à résoudre peut ainsi s'écrire :

$$\min S\ (x \mid w,\ a_* = a_{Rebours}) \qquad (P_R)$$

Bien que ce soit la méthode la plus généralement employée, nous ne la décrirons pas plus en détail ici car nous ne l'utilisons pas nous-mêmes. Le lecteur en trouvera l'exposé dans l'ouvrage de référence de Box & Jenkins *.

---

* BOX & JENKINS, 1970, pp. 209-220

La méthode des prévisions à rebours présente l'inconvénient sérieux de reposer entièrement sur la parfaite stationnarité de la série différenciée $\{w_t\}$ et donc sur l'identification exacte des degrés de différenciation d et D.

3.2.3. - Une troisième voie : l'optimisation simultanée des premiers résidus a_*

Nous proposons d'utiliser plutôt une méthode plus naturelle qui consiste à considérer les premiers résidus a_* comme de simples paramètres supplémentaires à optimiser :

$$\min S(x, a_* \mid w) \qquad (P_1)$$

$$\text{avec } S = \Sigma a_t^2 + \Sigma a_*^2$$

(les premiers résidus a_* sont inclus dans le calcul de la somme S).

La variance résiduelle à l'optimum $\hat{\sigma}_a^2$ peut alors être estimée en divisant la somme résiduelle $\hat{S}$ à l'optimum par le nombre ν de résidus calculés par récurrence, c'est-à-dire non compris les premiers résidus a_* :

$$\hat{\sigma}_a^2 = \frac{\hat{S}}{\nu} \qquad \text{avec } \nu = N - d - sD - p - sP$$

3.3. - Le programme d'optimisation

Le problème (P_1) peut comporter un assez grand nombre de paramètres en raison des q + Qs premiers résidus a_* introduits.

Nous utilisons un programme d'optimisation non-linéaire* qui comporte une recherche unidimensionnelle par la méthode de Coggin et l'approximation de l'inverse du Hessien par la méthode de Davidon-Fletcher-Powell. L'optimisation se fait en deux temps : résolution du problème (P_o) puis résolution du problème (P_1) à partir de la solution approchée précédente.

---

* HIMMELBLAU, 1972

Ce programme, comme on le verra par la suite, se comporte bien, même au voisinage des limites du domaine d'inversibilité et de stationnarité. En effet, aucune hypothèse n'a été faite sur la stationnarité de la série $\{w_t\}$ et l'estimation des paramètres s'effectue correctement même dans le cas où les polynômes $\phi(B)$, $\theta(B)$, $\Phi(B^s)$ ou $\Theta(B^s)$ possèdent des racines proches de 1 ou égales à 1.

La somme S ne doit jamais être calculée en des points de l'espace des paramètres situés à l'extérieur des domaines de stationnarité et d'inversibilité du modèle. Si le cas se présente, on calcule S sur la frontière du domaine et on ajoute une pénalité proportionnelle à la distance du point considéré à la frontière.

3.4. - Précision des estimations

Pour des séries suffisamment longues, on montre* que les valeurs estimées des paramètres sont des variables aléatoires approximativement distribuées selon une loi multinormale de moyennes égales aux vraies valeurs des paramètres, avec une matrice de variances-covariances :

$$V(\hat{x}) = \begin{pmatrix} V(\hat{x}_1) & \cdots & \mathrm{Cov}(\hat{x}_1\ \hat{x}_m) \\ & \ddots & \\ \cdots & & V(\hat{x}_m) \end{pmatrix} = \frac{2}{n}\, S(\hat{x})\ .\ S''(\hat{x})^{-1} \qquad (3.4.1)$$

avec

$$\begin{cases} S(\hat{x}) = \min \sum a_t^2 \\ \\ S''(\hat{x}) = \begin{pmatrix} \dfrac{\partial^2 S}{\partial x_1} & \cdots & \dfrac{\partial^2 S}{\partial x_1\ \partial x_m} \\ & \ddots & \\ \cdots & & \dfrac{\partial^2 S}{\partial x_m^2} \end{pmatrix}_{x = \hat{x}} \end{cases}$$

---

* BOX & JENKINS, 1970, p. 227

On peut estimer $S''(\hat{x})$ en calculant les dérivées secondes par différences finies. Dans les cas simples, on peut utiliser les formules analytiques suivantes :

- ARIMA (0,d,1) : $\quad Var(\hat{\theta}) \simeq \frac{1}{n}(1 - \hat{\theta}^2)$

- ARIMA (0,d,2) :

$$V(\hat{\theta}_1, \hat{\theta}_2) \simeq \frac{1}{n} \begin{pmatrix} 1 - \hat{\theta}_2^2 & -\hat{\theta}_1(1 + \hat{\theta}_2) \\ -\hat{\theta}_1(1 + \hat{\theta}_2) & 1 - \hat{\theta}_2^2 \end{pmatrix}$$

- ARIMA (1,d,0) : $\quad Var(\hat{\varphi}) \simeq \frac{1}{n}(1 - \hat{\varphi}^2)$

- ARIMA (2,d,0) :

$$V(\hat{\varphi}_1, \hat{\varphi}_2) \simeq \frac{1}{n} \begin{pmatrix} 1 - \hat{\varphi}_2^2 & -\hat{\varphi}_1(1 + \hat{\varphi}_2) \\ -\hat{\varphi}_1(1 + \hat{\varphi}_2) & 1 - \hat{\varphi}_2^2 \end{pmatrix}$$

- ARIMA (1,d,1) :

$$V(\hat{\varphi}, \hat{\theta}) \simeq \frac{1}{n} \frac{1 - \hat{\varphi}\hat{\theta}}{(\hat{\varphi} - \hat{\theta})^2} \begin{pmatrix} (1-\hat{\varphi}^2)(1-\hat{\varphi}\hat{\theta}) & (1-\hat{\varphi}^2)(1-\hat{\theta}^2) \\ (1-\hat{\varphi}^2)(1-\hat{\theta}^2) & (1-\hat{\varphi}\hat{\theta})(1-\hat{\theta}^2) \end{pmatrix}$$

On constate que la précision des estimations ne dépend que du nombre de termes de la série différenciée $\{w_t\}$ et des valeurs $\hat{\varphi}_i$ et $\hat{\theta}_i$ des paramètres à l'optimum. Ceci permet de construire facilement des intervalles de confiance (par exemple à 95%) autour des valeurs estimées.

4. - CHOIX DU MODELE

4.1. - Identification à l'aide des corrélogrammes

Avant d'estimer les paramètres, il faut spécifier les valeurs de p, d, q, P, D, Q, s. La méthode habituelle, appelée identification, repose sur l'analyse des corrélogrammes et fonctions d'autocorrélation partielle de la série étudiée $\{z_t\}$ et de ses différences $\{\nabla^d \nabla_s^D z_t\}$ (ces notions sont définies plus loin aux § 5.1 et 5.3). Les fonctions

observées sont rapprochées des fonctions théoriques connues correspondant aux modèles ARIMA ou SARIMA simples et on retient finalement le modèle qui semble le mieux approprié.

En pratique, il s'agit d'une opération assez délicate car les fonctions observées sont rarement typiques et l'identification nécessite une assez grande expérience, surtout pour les modèles mixtes ($p \neq 0$ et $q \neq 0$) et les modèles saisonniers.

Pour identifier les degrés de différenciation d et D, certains auteurs* proposent simplement de retenir le ou les couples (d, D) correspondant à la série différenciée de variance minimum :

$$\min \left[\hat{\sigma}^2_{w_t}(d, D) \mid w_t = \nabla^d \nabla_s^D z_t \right]$$

avec $0 \leq d \leq 2$

$0 \leq D \leq 2$

Nous allons voir que l'utilisation d'un programme d'estimation avec optimisation simultanée des premiers résidus permet d'envisager une toute autre approche sans faire appel aux notions de corrélogramme et des fonctions d'autocorrélation partielle.**

4.2. - Equivalence des modèles

Etant donnée une série observée $\{z_t\}$, l'estimation des paramètres d'un modèle quelconque (donc a priori mal identifié) a-t'elle un sens ?

Examinons, par exemple, les estimations correspondant à différents modèles pour une série de cours d'actions IBM (cf. § 7, annexe). Les résultats sont rassemblés dans le tableau 1. Les chiffres entre crochets désignent l'écart-type de l'estimation.

---

* Cf. par exemple, ANDERSON, O.D., 1976, p. 125 & LENZ, 1978, p. 7.

** Tout ce qui suit ne serait pas applicable avec la méthode des prévisions à rebours puisque celle-ci suppose que le degré de différenciation a déjà été correctement identifié.

Chaque modèle peut être considéré comme la "meilleure" approximation du "vrai modèle" sous-jacent à l'intérieur d'une classe particulière ARIMA (p,d,q).

Quelques opérations simples (factorisations, simplifications, abandon des paramètres non significativement différents de zéro) permettent de se rendre compte que la plupart de ces modèles sont équivalents. Les quelques exceptions se reconnaissent facilement à une somme résiduelle $\hat{S}$ nettement plus élevée.

Tableau 1 - Série B - MODELES ESTIMES, FACTORISES, SIMPLIFIES

ARIMA (p,d,q)	$\hat{S} = \Sigma a_t^2$ à l'optimum	Modèles
(0,0,2)	274.572	$\tilde{z}_t = (1 + 1{,}5\ B + 0{,}87\ B^2)\ a_t$ (S très grand) [0,03] [0,03]
(2,0,0)	19.201	$(1 - 1{,}09\ B + 0{,}09\ B^2)\ \tilde{z}_t = a_t$ [0,05] [0,05] $\rightarrow (1 - B)\ (1 - 0{,}09\ B)\ \tilde{z}_t = a_t$
(1,0,1)	19.211	$(1 - 0{,}999\ B)\ \tilde{z}_t = (1 + 0{,}09\ B)\ a_t$ [0,03] [0,05]
(2,0,2)	19.067	$(1 - 0{,}056\ B - 0{,}94\ B^2)\tilde{z}_t = (1 + 1{,}04\ B - 0{,}1\ B^2)a_t$ $\rightarrow \cancel{(1 + 0{,}9\ B)}\ (1 - 0{,}99\ B)\tilde{z}_t = \cancel{(1 + 0{,}9\ B)}(1 + 0{,}13\ B)a_t$
(0,1,2)	19.215	$\nabla z_t = (1 + 0{,}088\ B \cancel{+ 0{,}009\ B^2})a_t$ [0,05] [0,05]
(2,1,0)	19.184	$(1 - 0{,}087\ B \cancel{+ 0{,}007\ B^2})\ \nabla z_t = a_t$ [0,05] [0,05]
(1,1,1)	19.207	$(1 - 0{,}09\ B)\ \nabla z_t = \cancel{(1 - 0{,}005\ B)}\ a_t$ [0,60] [0,60]
(2,1,2)	18.161	$\cancel{(1 - 0{,}92\ B + 0{,}92\ B^2)}\nabla z_t = \cancel{(1 - 0{,}85\ B + 0{,}87\ B^2)}a_t$ racines complexes racines complexes
(0,2,2)	19.224	$\nabla^2 z_t = (1 - 0{,}9\ B - 0{,}08\ B^2)\ a_t$ [0,05] [0,05] $\rightarrow \nabla^2 z_t = (1 - 0{,}99\ B)\ (1 + 0{,}08\ B)\ a_t$
(2,2,0)	25.762	$(1 + 0{,}58\ B + 0{,}28\ B^2)\nabla^2 z_t = a_t$ (S très grand) [0,05] [0,05]
(1,2,1)	19.215	$(1 - 0{,}08\ B)\nabla^2 z_t = (1 - 0{,}99\ B)\ a_t$ [0,05] [0,01]
(2,2,2)	19.101	$(1 + 0{,}84\ B - 0{,}1\ B^2)\nabla^2 z_t = (1 - 0{,}07\ B - 0{,}91\ B^2)a_t$ $\rightarrow \cancel{(1 + 0{,}94\ B)}(1 - 0{,}1\ B)\ \nabla^2 z_t = \cancel{(1 + 0{,}92\ B)}(1 - 0{,}99B)a_t$

4.3. - Identification par élimination progressive

4.3.1. - Principe général

L'exemple de la série B suffit à montrer qu'une "erreur" dans le choix du degré de différenciation d se traduit simplement par l'estimation d'un facteur 1-B du côté autorégressif ou du côté des moyennes mobiles, ce qui permet une rectification très facile.

On voit alors qu'il est possible de concevoir une démarche progressive qui permette de trouver le modèle idéal sans pour autant estimer tous les modèles possibles, ce qui serait coûteux en temps de calcul.

Une première idée consisterait à partir d'un modèle largement suridentifié pour n'avoir plus qu'à le simplifier progressivement pour arriver au modèle définitif. En fait une telle approche est un peu brutale : l'interprétation d'un modèle ARIMA (2,d,2) n'est pas toujours facile et le temps de calcul nécessaire peut être assez long, surtout si un même facteur est présent des deux côtés à la fois.

Finalement, nous proposons d'adopter la ligne de conduite suivante, pour les séries non-saisonnières, en partant d'une valeur de d donnée quelconque ($0 \leq d \leq 2$, par exemple 0).

(a) Estimer le modèle (1,d,1)

(a_1) si $\hat{\varphi} \simeq 1$, remplacer d par d + 1 et retourner en (a);
sinon aller en (a_2).

(a_2) si $\hat{\theta} \simeq 1$, remplacer d par d - 1 et retourner en (a);
sinon aller en (b).

(b) Suridentifier ensuite dans la direction suggérée par le modèle retenu précédemment, c'est-à-dire :

(b_1) si $\hat{\theta} \simeq 0$ estimer (2,d,0)

(b_2) si $\hat{\varphi} \simeq 0$ estimer (0,d,2)

(b_3) si $\hat{\varphi} \simeq \hat{\theta}$ estimer (2,d,0) et (0,d,2) et retenir le côté correspondant à la plus petite somme $S = \Sigma a_t^2$

(b_4) si $\hat{\varphi} \neq 0$, $\hat{\theta} \neq 0$, $\hat{\varphi} \neq \hat{\theta}$ estimer (2,d,2)

Factoriser, si possible, les polynômes et aller en (c).

(c) Réduire (c'est-à-dire diminuer p ou q) ou développer (c'est-à-dire augmenter p ou q) progressivement le modèle en tenant compte :

(c_1) de la précision des estimations pour décider si un paramètre est non significativement différent de zéro ou si deux facteurs sont identiques.

(c_2) de la valeur de $\hat{S} = \Sigma a_t^2$ à l'optimum et du principe de parcimonie pour le nombre de paramètres utilisés.

En cas de doute, on peut toujours essayer quelques modèles supplémentaires, mais généralement le choix final s'opère assez rapidement. Dans le cas relativement fréquent où certains paramètres prennent des valeurs tout juste significatives c'est à l'analyste de décider si l'abaissement de la somme $\hat{S}$ est suffisamment important pour justifier l'emploi de ces paramètres supplémentaires.

4.3.2. - Premier exemple : série B (suite)

En suivant la méthode ci-dessus pour la série IBM nous aurions successivement essayé :

$(1,0,1)$: $(1 - 0{,}999\ B)\ \tilde{z}_t = (1 + 0{,}09\ B)\ a_t$ $\quad \hat{S} = 19.211$
[0,03] [0,05]

↓

$(1,1,1)$: $(1 - 0{,}09\ B)\nabla z_t = (1 - 0{,}005\ B)\ a_t$ $\quad \hat{S} = 19.207$
[0,6] [0,6]

↓

$(2,1,0)$: $(1 - 0{,}09\ B + 0{,}007\ B^2)\nabla z_t = a_t$ $\quad \hat{S} = 19.184$
[0,05] [0,05]

↓

$(1,1,0)$: $(1 - 0{,}09\ B)\nabla z_t = a_t$ $\quad \hat{S} = 19.207$
[0,05]

Le modèle (1,1,1) comporte un paramètre φ voisin de zéro et on peut donc également essayer :

$$(0,1,2) : \nabla z_t = (1 + \underset{[0,05]}{0,09}\ B + \underset{[0,05]}{0,009}\ B^2)a_t \qquad \hat{S} = 19.215$$

$$(0,1,1) : \nabla z_t = (1 + \underset{[0,05]}{0,09}\ B)\ a_t \qquad \hat{S} = 19.217$$

Remarquons que les deux modèles finaux (1,1,0) et (0,1,1) sont pratiquement indiscernables malgré la longueur de la série étudiée :

$$(1 - 0,09\ B)^{-1} \simeq (1 + 0,09\ B).$$

Ils sont d'ailleurs très proches du modèle de marche au hasard $\nabla z_t = a_t$ (puisque $\hat{\varphi} \simeq 0$ et $\hat{\theta} \simeq 0$) que l'on rencontre souvent pour les séries de cours boursiers.

Signalons enfin que l'estimation du modèle (1,2,1) au départ, nous aurait immédiatement conduit à envisager d = 1. On trouve en effet :

$$(1,2,1) : \underset{[0,05]}{(1 - 0,09\ B)}\nabla^2 z_t = \underset{[0,01]}{(1 - 0,999\ B)}a_t \qquad \hat{S} = 19.215$$

4.3.3. - Modèles comportant un terme constant

Normalement, la moyenne résiduelle est très petite par rapport à l'écart-type : $\bar{a} << \hat{\sigma}_a$. Dans certains cas, il peut pourtant arriver que la moyenne résiduelle ne soit pas suffisamment petite. Il est alors préférable d'utiliser un modèle de la forme (2.1.4) comportant un terme constant θ_o : $\phi(B)\ \nabla^d z_t = \theta_o + \theta(B)\ a_t$. Si l'adoption d'un tel modèle est justifiée, elle se traduit alors par une réduction très sensible de $\bar{a}$ et de la somme $\hat{S}$.

Pour ces modèles, les simplifications s'opèrent de la même manière que pour les autres modèles, à condition de modifier la constante au passage.

Par exemple : $(1 - \alpha B)(1 - \varphi B)\nabla z_t = \theta_o + (1 - \alpha B)\, a_t$

se simplifie en : $(1 - \varphi B)\nabla z_t = \frac{\theta_o}{1-\alpha} + a_t \qquad \alpha \neq 1$

4.3.4. - Deuxième exemple : série N

Un deuxième exemple permettra de mieux se familiariser avec la méthode d'identification par élimination progressive (tableau 2). Il s'agit du Produit National Brut aux USA (série N, cf. § 7, annexe).

4.4. - Identification des séries saisonnières

4.4.1. - Les modèles envisageables

Pour les séries saisonnières, le problème de l'identification à l'aide des corrélogrammes devient encore plus difficile que dans le cas non-saisonnier. L'identification par élimination progressive peut donc rendre des services encore plus grands. Mais la méthode doit être appliquée de manière judicieuse, car les modèles envisageables sont beaucoup plus nombreux : déjà 213 modèles SARIMA $(p,d,q) \times (P,D,Q)s$ en se limitant pourtant à $p \leq 2$, $d \leq 2$, $q \leq 2$, $P \leq 1$, $D \leq 1$, $Q \leq 1$.

De plus, des modèles très complexes du genre $(2,d,2) \times (2,D,2)s$ sont à proscrire absolument car ils sont trop peu parcimonieux et la convergence du programme d'estimation risque d'être vraiment trop lente.

Dans le cas de séries mensuelles, il faut éviter d'envisager des modèles qui réduisent trop le nombre de points utilisables pour l'estimation. En pratique on peut se limiter à $P + D \leq 2$.

Quant à l'opérateur de moyennes mobiles saisonnières $\theta(B^{12})$, on se limite aussi à $Q \leq 1$ pour ne pas introduire trop de paramètres supplémentaires à optimiser. En effet, l'estimation se fait avec optimisation simultanée des premiers résidus et $Q = 2$ correspondrait à 24 paramètres supplémentaires.

TABLEAU 2 - SERIE N - CHOIX DU MODELE

	ARIMA	$\hat{S} = \Sigma a_t^2$	Modèles estimés, factorisés, simplifiés	Remarques
d=0	(1,0,1)	3.593,0	$(1-0{,}99997\ B)\tilde{z}_t = (1+0{,}64\ B)a_t$ [0,001] [0,09]	$\hat{\sigma}_a \simeq \bar{a} \rightarrow$ introduire θ_o
	avec cte θ_o	2.011,4	$(1-0{,}99998\ B)\tilde{z}_t = 6{,}98 + (1+0{,}47\ B)a_t$ [0,001] [0,1]	$\varphi = 1 \Longrightarrow d > 0$
d=1	(1,1,1)	1.959,7	$(1-0{,}89\ B)\nabla z_t = (1-0{,}16\ B)a_t$ [0,06] [0,12]	$\hat{\sigma}_a \simeq 5\ \bar{a} \rightarrow$ introduire θ_o
	avec cte θ_o	1.738,9	$(1-0{,}62\ B)\nabla z_t = 2{,}71 + (1-0{,}01\ B)a_t$ [0,14] [0,2]	$\varphi \neq 1$ $\theta \simeq 0$ → développer à gauche
	(2,1,0)	1.737,9	$(1-0{,}61\ B - 0{,}01\ B^2)\nabla z_t = 2{,}74 + a_t$ [0,11] [0,11]	$\varphi_2 \simeq 0 \rightarrow p = 1$
	avec cte θ_o		$(1+0{,}02\ B)(1-0{,}62\ B)\nabla z_t = 2{,}74 + a_t$	
	(1,1,0) avec cte θ_o	1.739,1	$(1-0{,}62\ B)\nabla z_t = 2{,}78 + a_t$ [0,09]	choix final
d=2	(1,2,1)	1.754,3	$(1-0{,}56\ B)\nabla^2 z_t = (1-0{,}95\ B)a_t$ [0,11] [0,04]	$\hat{\sigma}_a \simeq 5\ \bar{a} \rightarrow$ introduire θ_o
	avec cte θ_o	1.660,5	$(1-0{,}55\ B)\nabla^2 z_t = 0{,}04 + (1-0{,}9999\ B)a_t$ [0,09] [0,001]	$\theta \simeq 1 \Longrightarrow d < 2$

4.4.2. - Degré de différenciation

Pour les séries qui nous intéressent, le choix de s ne pose généralement pas de problème. Il s'agira le plus souvent de s = 12 (séries mensuelles) ou s = 4 (séries trimestrielles).

Le choix de l'opérateur de différences $\nabla^d \nabla_s^D$ est plus délicat. Comme pour les séries non-saisonnières, il se fait à partir de l'estimation des paramètres de différents modèles $(1,d,1) \times (1,D,1)_s$. Plusieurs cas se présentent :

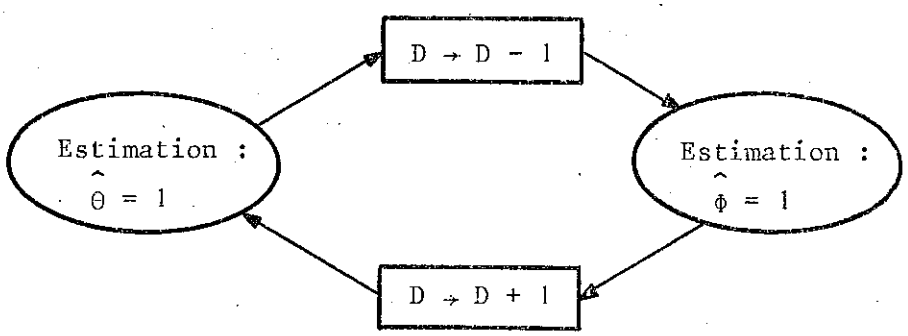
$\hat{\varphi} \simeq 1$: différencier une fois de plus $(d \rightarrow d + 1)$

$\hat{\theta} \simeq 1$: différencier une fois de moins $(d \rightarrow d - 1)$

$\hat{\Phi} \simeq 1$: différencier une fois de plus $(D \rightarrow D + 1)$

$\hat{\Theta} \simeq 1$: différencier une fois de moins $(D \rightarrow D - 1)$

Malheureusement, on peut se retrouver dans un cercle apparemment vicieux du genre :



En fait, l'opérateur $\nabla^d \nabla_s^D$ ne correspond alors vraiment à une surdifférenciation que si l'estimation du modèle $(1,d,1) \times (0,D,1)_s$ conduit aussi à $\hat{\Theta} = 1$.

De même, l'opérateur $\nabla^d \nabla_s^D$ ne correspond vraiment à une sous-différenciation que si l'estimation du modèle $(1,d,1) \times (1,D,0)_s$ conduit aussi à $\hat{\Phi} = 1$.

4.4.3. - Forme du modèle

L'identification consiste à passer progressivement d'un modèle à l'autre jusqu'à arriver au modèle définitif. Ce passage s'effectue en tenant compte des résultats de chaque estimation à l'aide des opérations suivantes :

(a) Modification de l'opérateur de différence chaque fois qu'apparaît un facteur proche de $(1 - B)$ ou $(1 - B^s)$ dans l'estimation (s'il apparaît des deux côtés à la fois, voir j)).

(b) Abandon des paramètres voisins de 0 (compte tenu de la précision des estimations).

(c) Factorisation des polynômes et simplifications éventuelles chaque fois qu'apparaissent des facteurs égaux (compte tenu de la précision des estimations) des deux côtés à la fois.

(d) Développement du modèle du côté où les paramètres ne sont pas nuls.

Pourtant certaines situations peuvent sembler assez difficiles au premier abord et les indications suivantes peuvent s'avérer utiles :

(e) Partir d'un modèle $(1,d,1) \times (1,D,1)_s$, par exemple $(1,0,1) \times (1,0,1)_s$.

(f) Ne pas envisager inutilement de modèles correspondant à $P > 1$, $D > 1$ ou $Q > 1$. En pratique, on se limitera à $P + D \leq 2$ et $Q \leq 1$.

(g) Ne pas tout changer à la fois : modifier la partie saisonnière ou la partie non-saisonnière du modèle.

(h) Si plusieurs possibilités s'offrent, les explorer toutes.

(i) En cas de simplification (cf. (c)) essayer des modèles dissymétriques $(0,d,q) \times (P,D,Q)$ et $(p,d,0) \times (P,D,Q)$ ou $(p,d,q) \times (0,D,Q)$ et $(p,d,q) \times (P,D,0)$.

(j) En cas de cercle vicieux (cf. § 4.4.2) essayer aussi des modèles dissymétriques. Il s'agit en effet, le plus souvent, d'un cas particulier de (i) où des termes voisins de $(1 - B)$ ou $(1 - B^S)$ apparaissent simultanément des deux côtés.

(k) Si l'opérateur $(1 + B)$ ou $(1 + B^S)$ apparaît, essayer de développer le modèle de l'autre côté. Le résultat est en général un abaissement sensible de la somme S.

(l) Si plusieurs modèles restent finalement en concurrence, on pourra éventuellement les départager en comparant les sommes $\hat{S}$ et les diagnostics.

4.4.4. - Exemples (séries G et Z)

Deux exemples vont éclairer le genre de cheminement qu'il faut réaliser.

a) Série G : Trafic aérien

L'identification de la série transformée $\{\text{Log } z_t\}$ (cf. § 7, annexe) est résumée dans le tableau 3. Elle ne soulève aucune difficulté .

b) Série Z : ventes mensuelles d'une entreprise

L'identification de la série Z (voir données et références au § 7, annexe) est résumée dans le tableau 4. Elle conduit au modèle :

$$(0,1,1) \times (2,0,0) : \underset{[0,1]}{(1 - 0,604\ B^{12} - 0,289\ B^{24})} \nabla z_t = \underset{[0,1]}{(1 - 0,615\ B)}\ a_t$$

L'opérateur saisonnier se factorise :

$$(1 + 0,314\ B^{12})(1 - 0,918\ B^{12}) \nabla z_t = (1 - 0,615\ B)\ a_t$$

et on remarque alors la présence de l'opérateur $(1 - 0,918\ B^{12})$ très proche de $\nabla_{12} = 1 - B^{12}$. De fait, l'estimation d'un modèle SARIMA

TABLEAU 3 - SERIE G - CHOIX DU MODELE

	SARIMA	$\hat{S} = \Sigma a_t^2$	Modèles estimés, factorisés, simplifiés	Remarques
1	$(1,0,1)\times(1,0,1)_{12}$	0,182	$(1-0{,}998\ B)(1-0{,}9999\ B^{12})\ \widetilde{\text{Log}}\ z_t = (1-0{,}381\ B)(1-0{,}568\ B^{12})a_t$ [0,005] [0,0009] [0,08] [0,07]	$\varphi \simeq 1 \Longrightarrow d > 0$ $\Phi \simeq 1 \Longrightarrow D > 0$
2	$(1,0,1)\times(1,1,1)_{12}$	0,163	$(1-0{,}990\ B)(1+0{,}213\ B^{12})\nabla_{12}\ \text{Log}\ z_t = (1-0{,}466\ B)(1-0{,}412\ B^{12})a_t$ [0,01] [0,15] [0,08] [0,14]	$\varphi \simeq 1 \Longrightarrow d > 0$
3	$(1,1,1)\times(1,1,1)_{12}$	0,163	~~$(1-0{,}177\ B)(1+0{,}119\ B^{12})$~~ $\nabla\nabla_{12}\ \text{Log}\ z_t = (1-0{,}592\ B)(1-0{,}533\ B^{12})a_t$ ~~[0,19]~~ ~~[0,14]~~ [0,15] [0,12]	$\Phi \simeq 0 \rightarrow$ développer à droite
4	$(1,1,1)\times(0,1,1)_{12}$	0,175	~~$(1-0{,}299\ B)$~~ $\nabla\nabla_{12}\ \text{Log}\ z_t = (1-0{,}665\ B)(1-0{,}628\ B^{12})a_t$ ~~[0,18]~~ [0,14] [0,07]	$\varphi \simeq 0 \rightarrow$ développer à droite
5	$(0,1,2)\times(0,1,1)_{12}$	0,175	$\nabla\nabla_{12}\ \text{Log}\ z_t = (1-0{,}391\ B - 0{,}047\ B^2)(1-0{,}616\ B^{12})a_t$ [0,09] [0,09] [0,07]	$\theta_2 \simeq 0 \rightarrow q = 1$
6	$(0,1,1)\times(0,1,1)_{12}$	0,182	$\nabla\nabla_{12}\ \text{Log}\ z_t = (1-0{,}396\ B)(1-0{,}614\ B^{12})a_t$ [0,08] [0,07]	choix final

TABLEAU 4 - SERIE Z - CHOIX DU MODELE

	SARIMA	$\hat{S} = \Sigma a_t^2$	Modèles estimés, factorisés, simplifiés	Remarques
1	$(1,0,1)\times(1,0,1)_{12}$	$28{,}06.10^6$	$(1-0{,}970\ B)(1-0{,}999\ B^{12})\tilde{z}_t = (1-0{,}570\ B)(1-0{,}581\ B^{12})a_t$ $[0{,}03]\quad[0{,}001]\quad[0{,}11]\quad[0{,}1]$	$\Phi \simeq 1$ $\Theta \neq 1$ → essayer D = 1
2	$(1,0,1)\times(1,1,1)_{12}$	$14{,}54.10^6$	$(1-0{,}933\ B)(1+0{,}494\ B^{12})\nabla_{12}\ z_t = (1-0{,}532\ B)(1-0{,}999\ B^{12})a_t$ $[0{,}06]\quad[0{,}12]\quad[0{,}15]\quad[0{,}002]$	$\Theta \simeq 1$: essayer P = 0 pour voir si ∇_{12} est vraiment inadéquat
3	$(1,0,1)\times(0,1,1)_{12}$	$26{,}03.10^6$	$(1-0{,}900\ B)\nabla_{12}\ z_t = (1-0{,}601\ B)(1-0{,}9999\ B^{12})a_t$ $[0{,}02]\quad[0{,}12]\quad[0{,}0015]$	$\Theta \simeq 1$: ∇_{12} ne convient pas essayer d = 1, D = 0
4	$(1,1,1)\times(1,0,1)_{12}$	$28{,}40.10^6$	$(1+0{,}084\ B)(1-0{,}9999\ B^{12})\nabla z_t = (1-0{,}541\ B)(1-0{,}573\ B^{12})a_t$ $[0{,}2]\quad[0{,}001]\quad[0{,}2]\quad[0{,}1]$	$\Phi \simeq 1$: essayer Q = 0 pour voir si ∇_{12} apparaît toujours
5	$(1,1,1)\times(1,0,0)_{12}$	$29{,}75.10^6$	$(1+0{,}077\ B)(1-0{,}652\ B^{12})\nabla z_t = (1-0{,}492\ B)a_t$ $[0{,}2]\quad[0{,}1]\quad[0{,}1]$	$\varphi \simeq 0$: essayer p = 0
6	$(0,1,2)\times(1,0,0)_{12}$	$29{,}76.10^6$	$(1-0{,}657\ B^{12})\nabla z_t = (1-0{,}569\ B + 0{,}050\ B^2)a_t$ $[0{,}1]\quad[0{,}1]\quad[0{,}1]$	$\theta_2 \simeq 0$: essayer q = 1
7	$(0,1,1)\times(1,0,0)_{12}$	$29{,}81.10^6$	$(1-0{,}647\ B^{12})\nabla z_t = (1-0{,}549\ B)a_t$ $[0{,}1]\quad[0{,}1]$	P + D < 2 : on peut encore suridentifier en essayant P = 2
8	$(0,1,1)\times(2,0,0)_{12}$	$21{,}70.10^6$	$(1-0{,}604\ B^{12} - 0{,}289\ B^{24})\nabla z_t = (1-0{,}615\ B)a_t$ $[0{,}1]\quad[0{,}1]\quad[0{,}1]$ $(1+0{,}314\ B^{12})(1-0{,}918\ B^{12})\nabla z_t = (1-0{,}615\ B)a_t$	Factorisation → choix final

A titre indicatif, citons aussi :

9	$(0,1,1)\times(1,1,0)_{12}$	$21{,}85.10^6$	$(1+0{,}366\ B^{12})\nabla\nabla_{12}\ z_t = (1-0{,}633\ B)a_t$ $[0{,}1]\quad[0{,}1]$	Modèle très voisin du précédent
10	$(0,1,1)\times(0,1,1)_{12}$	$22{,}13.10^6$	$\nabla\nabla_{12}\ z_t = (1-0{,}634\ B)(1-0{,}9999\ B^{12})a_t$ $[0{,}1]\quad[0{,}002]$	$\Theta \simeq 1$: essayer D = 0

$(0,1,1) \times (1,1,0)$ aboutit à un modèle très voisin :

$$(0,1,1) \times (1,1,0) : \underset{[0,1]}{(1 + 0{,}366\ B^{12})} \nabla\nabla_{12}\ z_t = \underset{[0,1]}{(1 - 0{,}633\ B)}\ a_t .$$

4.5. - Une identification controversée : la série P

4.5.1. - Historique

La série P représente les ventes mensuelles d'une certaine entreprise X. Cette série à été choisie à dessein pour faire l'objet d'une étude plus détaillée car son identification est assez difficile et a déjà fait couler beaucoup d'encre à la suite de l'article initial de Chatfield & Prothero (1973). Résumons les faits.

Chatfield & Prothero utilisent une transformation en logarithmes décimaux et identifient puis estiment le modèle :

$$(1 + 0{,}47\ B) \nabla\nabla_{12}\ \log_{10} z_t = (1 - 0{,}81\ B^{12})\ a_t \qquad (a)$$

Ce modèle se comporte manifestement assez mal en prévision, même en prenant soin d'utiliser les premiers résidus fournis par la méthode des prévisions à rebours.

Chatfield & Prothero estiment alors trois autres modèles en $\nabla\nabla_{12}$:

$$(1 + 0{,}51\ B)\ (1 + 0{,}47\ B^{12}) \nabla\nabla_{12}\ \log_{10} z_t = a_t \qquad (b)$$

$$(1 + 0{,}56\ B^{12}) \nabla\nabla_{12}\ \log_{10} z_t = (1 - 0{,}49\ B)\ a_t \qquad (c)$$

$$\nabla\nabla_{12}\ \log_{10} z_t = (1 - 0{,}44\ B)\ (1 - 0{,}85\ B^{12})\ a_t \qquad (d)$$

pour finalement préférer le modèle (b). Compte tenu de toutes les difficultés rencontrées dans le choix entre les quatre modèles (a), (b), (c), (d), ils concluent en apportant de sérieuses réserves quant à l'utilisation de la méthode de Box & Jenkins.

Box & Jenkins (1973) répondent en incriminant la transformation logarithmique. Ils suggèrent plutôt une transformation en puissance et identifient la même forme de modèle qu'ils estiment à leur tour :

$$(1 + 0{,}5\ B)\nabla\nabla_{12}\ z_t^{1/4} = (1 - 0{,}8\ B^{12})\ a_t \qquad (a')$$

O.D. Anderson (1975a) propose le modèle :

$$\nabla_{12}\ \mathrm{Log}\ z_t = (1 + 0{,}467\ B + 0{,}715\ B^2)\ a_t \qquad (e)$$

et Melard (1977) propose une transformation sous forme d'une fonction exponentielle du temps :

$$(1 - 0{,}50\ B)\nabla\nabla_{12}\ z_t = 0{,}50\ \exp.(0{,}016\ t)\ a_t \qquad (f)$$

4.5.2. - Comparaison des estimations

Ces auteurs ont estimé tous ces modèles par la méthode des prévisions à rebours.

De notre côté, l'estimation avec optimisation simultanée des premiers résidus, au lieu de l'équation a), nous conduit à l'équation :

$$\underset{[0,11]}{(1 + 0{,}423\ B)}\nabla\nabla_{12}\ \log_{10} z_t = \underset{[0,00001]}{(1 - 0{,}999999\ B^{12})}\ a_t \qquad (a_1)$$

de sorte que l'opérateur $\nabla\nabla_{12}$ paraît surdifférencié.

La transformation logarithmique n'est pas responsable de cette surdifférenciation. En effet, au lieu de l'équation (a'), nous trouvons encore :

$$\underset{[0,11]}{(1 + 0{,}475\ B)}\nabla\nabla_{12}\ z_t^{1/4} = \underset{[0,002]}{(1 - 0{,}9999\ B^{12})}\ a_t \qquad (a_1')$$

Pour les trois autres modèles de Chatfield & Prothero, nous trouvons respectivement :

$$(1 + \underset{[0,10]}{0,601}\ B)\ (1 + \underset{[0,12]}{0,229}\ B^{12})\nabla\nabla_{12}\ \log_{10} z_t = a_t \qquad (b_1)$$

$$(1 + \underset{[0,11]}{0,349}\ B^{12})\nabla\nabla_{12}\ \log_{10} z_t = (1 - \underset{[0,12]}{0,508}\ B)\ a_t \qquad (c_1)$$

$$\nabla\nabla_{12}\ \log_{10} z_t = (1 - \underset{[0,11]}{0,373}\ B)\ (1 - \underset{[0,001]}{0,9999}\ B^{12})\ a_t \qquad (d_1)$$

Quant au modèle d'Anderson nous trouvons (en logarithmes décimaux pour faciliter les comparaisons) :

$$\nabla_{12}\ \log_{10} z_t = (1 + \underset{[0,09]}{0,432}\ B + \underset{[0,09]}{0,717}\ B^2)\ a_t \qquad (e_1)$$

4.5.3. - Identification

La situation apparaît pour le moins confuse. Essayons donc la méthode par élimination progressive à partir de la série transformée en logarithmes décimaux (cf. tableau 5).

Elle conduit à l'identification et à l'estimation du modèle suivant :

$$(1 + \underset{[0,09]}{0,561}\ B)\ (1 - \underset{[0,11]}{0,631}\ B^{12} - \underset{[0,11]}{0,237}\ B^{24})\nabla\ \log_{10} z_t = a_t$$

dans lequel l'opérateur saisonnier se factorise :

$$(1 + 0,561\ B)\ (1 + 0,265\ B^{12})\ (1 - 0,896\ B^{12})\nabla\ \log_{10} z_t = a_t \qquad (g_1)$$

Ce modèle est très proche du modèle (b_1). C'est donc ces deux modèles (g_1) et (b_1) que nous retiendrons au stade de l'identification.

TABLEAU 5 - SERIE P - CHOIX DU MODELE

	SARIMA	$\hat{S} = \Sigma a_t^2$	Modèles estimés, factorisés, simplifiés	Remarques
1	$(1,0,1)\times(1,0,1)_{12}$	0,245	$(1-0{,}983\ B)(1-0{,}927\ B^{12})\log_{10}^{\sim} z_t = (1-0{,}467\ B)(1-0{,}9999\ B^{12})a_t$ [0,02] [0,04] [0,11] [0,001]	$\varphi \simeq 1 \Longrightarrow d > 0$ $\Phi \simeq \Theta \simeq 1 \rightarrow$ garder D = 0
2	$(1,1,1)\times(1,0,1)_{12}$	0,235	$(1+0{,}642\ B)(1-0{,}921\ B^{12})\nabla\log_{10} z_t = (1+0{,}234\ B)(1-0{,}9999\ B^{12})a_t$ [0,18] [0,04] [0,23] [0,0007]	$\Phi \simeq \Theta \simeq 1 \rightarrow$ essayer P = 0
3	$(1,1,1)\times(0,0,1)_{12}$	0,861	$(1-0{,}434\ B)\nabla\log_{10} z_t = (1-0{,}484\ B)(1+0{,}9999\ B^{12})a_t$ [1,6] [1,6] [0,0001]	$\Theta \simeq -1$, $\hat{S}$ grand $\rightarrow$ essayer plutôt Q = 0
4	$(1,1,1)\times(1,0,0)_{12}$	0,388	$(1+0{,}683\ B)(1-0{,}841\ B^{12})\nabla\log_{10} z_t = (1+0{,}065\ B)a_t$ [0,12] [0,06] [0,18]	$\theta \simeq 0 \Longrightarrow q = 0$
5	$(2,1,0)\times(1,0,0)_{12}$	0,384	$(1+0{,}609\ B - 0{,}059\ B^2)(1-0{,}832\ B^{12})\nabla\log_{10} z_t = a_t$ [0,11] [0,11] [0,06]	$\varphi_2 \simeq 0 \Longrightarrow p = 1$
6	$(1,1,0)\times(1,0,0)_{12}$	0,389	$(1+0{,}643\ B)(1-0{,}844\ B^{12})\nabla\log_{10} z_t = a_t$ [0,09] [0,06]	P + D < 2 : on peut encore suridentifier en essayant P = 2
7	$(1,1,0)\times(2,0,0)_{12}$	0,263	$(1+0{,}561\ B)(1-0{,}631\ B^{12} - 0{,}237\ B^{24})\nabla\log_{10} z_t = a_t$ [0,09] [0,11] [0,11] $(1+0{,}561\ B)(1+0{,}265\ B^{12})(1-0{,}896\ B^{12})\nabla\log_{10} z_t = a_t$	Factorisation choix final (g_1)

A titre indicatif, citons aussi :

8	$(1,1,0)\times(1,1,0)_{12}$	0,290	$(1+0{,}601\ B)(1+0{,}229\ B)\nabla\nabla_{12}\log_{10} z_t = a_t$ [0,10] [0,12]	Modèle très voisin (b_1) du précédent
9	$(0,1,1)\times(0,1,1)_{12}$	0,290	$\nabla\nabla_{12}\log_{10} z_t = (1-0{,}373\ B)(1-0{,}9999\ B^{12})a_t$ [0,11] [0,001]	$\Theta \simeq 1$ essayer D = 0

Nous comparerons les validations et les prévisions correspondant à ces modèles aux paragraphes 5.5 et 6.2.

4.5.4. - Bilan

L'identification à partir de divers modèles $(1,d,1) \times (1,D,1)_{12}$ nous permet de conclure que seul l'opérateur ∇ convient vraiment. L'opérateur ∇_{12} est inadéquat et le modèle (e_1) présente, de toutes façons, une somme résiduelle beaucoup trop forte. Enfin, parmi tous les modèles en $\nabla\nabla_{12}$, on peut éventuellement retenir le modèle (b_1) qui constitue une bonne approximation de notre modèle identifié (g_1).

Toutes les difficultés antérieures proviennent du fait que la méthode d'estimation avec prévisions à rebours n'est pas suffisamment précise lorsque $\hat{\Theta}$ est proche de 1. En effet, le modèle n'est alors plus vraiment réversible et les allers-retours caractéristiques de la méthode des prévisions à rebours convergent trop lentement.

C'est ce que confirment Box & Jenkins eux-mêmes dans leur réponse à Chatfield & Prothero (Box & Jenkins, 1973, remarque en bas de page 339) en signalant qu'une estimation plus poussée du modèle (a') conduit à $\hat{\Theta} = 0,9$.

5. - VALIDATION

5.1. - Corrélogramme résiduel

Nous avons vu comment il était possible d'éviter l'interprétation difficile des corrélogrammes au stade de l'identification en pratiquant l'identification par élimination progressive. Au stade de la validation, alors qu'il ne s'agit plus que de vérifier que les résidus constituent bien un bruit blanc, le corrélogramme redevient un outil simple et utile.

Rappelons d'abord les définitions des grandeurs suivantes, indépendantes du temps (processus stationnaires).

a) Autocovariances

$$\gamma_k = \gamma_{-k} = \text{cov}(\tilde{z}_t, \tilde{z}_{t-k}) = E(\tilde{z}_t, \tilde{z}_{t-k}) \qquad (5.1.1)$$

$$\gamma_o = E(\tilde{z}_t^2) = \sigma_z^2 \qquad \text{avec } \tilde{z}_t = z_t - \mu \text{ et } \mu = E(z_t)$$

b) Autocorrélations

$$\rho_k = \rho_{-k} = \frac{E(\tilde{z}_t \cdot \tilde{z}_{t-k})}{\sqrt{E(\tilde{z}_t^2) \cdot E(\tilde{z}_{t-k}^2)}} = \frac{\gamma_k}{\gamma_o} \leq 1 \qquad (5.1.2)$$

$$\rho_o = 1$$

c) Fonction d'autocorrélation ou corrélogramme

Il s'agit de la fonction symétrique $\rho_k = f(k)$ dont on ne représente que la partie correspondant à $k \geq 0$. Pour un bruit blanc, le corrélogramme théorique se réduit à $\rho_k = 0 \qquad \forall k \geq 1$.

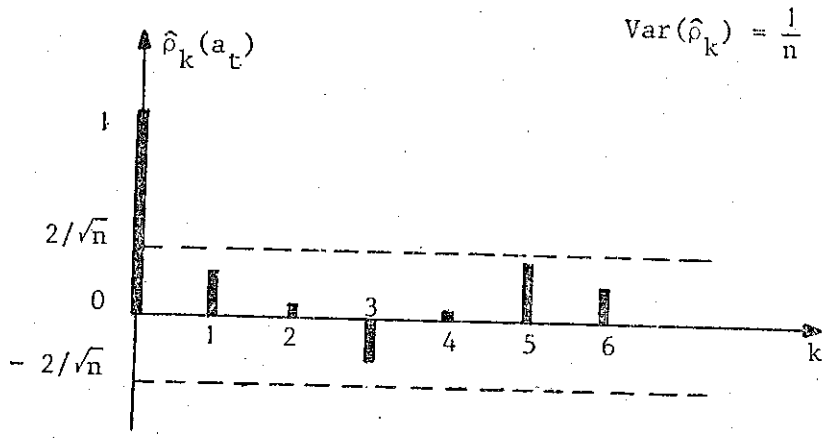
Dans la pratique les coefficients γ_k et ρ_k sont estimés à partir des formules :

$$\hat{\rho}_k = r_k = \frac{C_k}{C_o} \qquad (5.1.3)$$

$$C_k = \hat{\gamma}_k = \frac{1}{N} \sum_{t=k+1}^{N} (z_t - \bar{z})(z_{t-k} - \bar{z}) \qquad (5.1.4)$$

Même si les ρ_k théoriques sont nuls, les $\hat{\rho}_k$ calculés ne sont évidemment jamais exactement nuls. Il faut donc également calculer $\text{Var}(\hat{\rho}_k)$, de façon à déceler les $\hat{\rho}_k$ qui seraient significativement différents de zéro. On prend habituellement la formule approchée $\text{Var}(\hat{\rho}_k) = 1/n$.

Le corrélogramme d'un bruit blanc a donc l'allure suivante :



Les lignes pointillées représentent un intervalle de confiance à 95% autour de la valeur théorique zéro.

On devrait normalement s'attendre à trouver tous les $\hat{\rho}_k$ (sauf $\rho_o = 1$) à l'intérieur de cet intervalle. Pourtant, la probabilité de 95% s'applique à chaque coefficient pris individuellement et il serait donc tolérable que sur un ensemble, disons, d'une vingtaine de coefficients, un ou deux d'entre eux sortent des limites prévues.

5.2. - Test du Chi 2

Pour construire un test global plus rigoureux, on peut calculer la quantité :

$$Q_K = n \sum_{k=1}^{K} \hat{\rho}_k^2 (\hat{a}_t) \qquad (5.2.1)$$

On montre* que, pour un modèle ARIMA(p,d,q), Q_K se comporte approximativement comme un χ^2 à K-p-q degrés de liberté, avec n = N - d, sous réserve que n >> K et que K soit lui-même suffisamment grand (par exemple K = 25 pour une série non-saisonnière avec n > 200). On peut donc vérifier que la valeur de Q_K n'est pas trop élevée en la comparant aux valeurs théoriques lues dans une table du χ^2.

---

* BOX & PIERCE, 1970.

En fait, ce test ne s'avère pas très sensible et on peut utiliser à la place de Q_K^* :

$$Q'_K = n(n+2) \sum_{k=1}^{K} \frac{1}{n-k} \hat{\rho}_k^2(\hat{a}_t)$$

qui possède une variance supérieure.

5.3. - Fonction d'autocorrélation partielle

L'étude de la fonction d'autocorrélation va de pair avec celle de la fonction d'autocorrélation partielle et nous avons vu comment on peut se passer de ces notions au stade de l'identification.

Au stade de la validation, on peut aussi calculer la fonction d'autocorrélation partielle des résidus. C'est ce que nous avons fait pour un grand nombre de séries et de modèles sans que ce test n'apporte d'information qui ne soit déjà contenue dans le corrélogramme.

Pour cette raison, nous proposons de nous passer complètement de la notion d'autocorrélation partielle que nous ne décrirons donc pas ici et nous renvoyons le lecteur intéressé à l'ouvrage de Box & Jenkins**.

5.4. - Périodogramme cumulé

Dans le cas des séries saisonnières, on peut éventuellement effectuer aussi le test du périodogramme cumulé pour s'assurer que la périodicité de la série a été bien prise en compte ***. Nous n'avons pas utilisé ce test.

5.5. - Exemples (séries Z et P)

Examinons, à titre d'exemple les corrélogrammes résiduels de quelques uns des modèles estimés pour les séries Z et P.

---

* LJUNG & BOX, 1976 (cité par O.D. ANDERSON (1977a) p. 283)

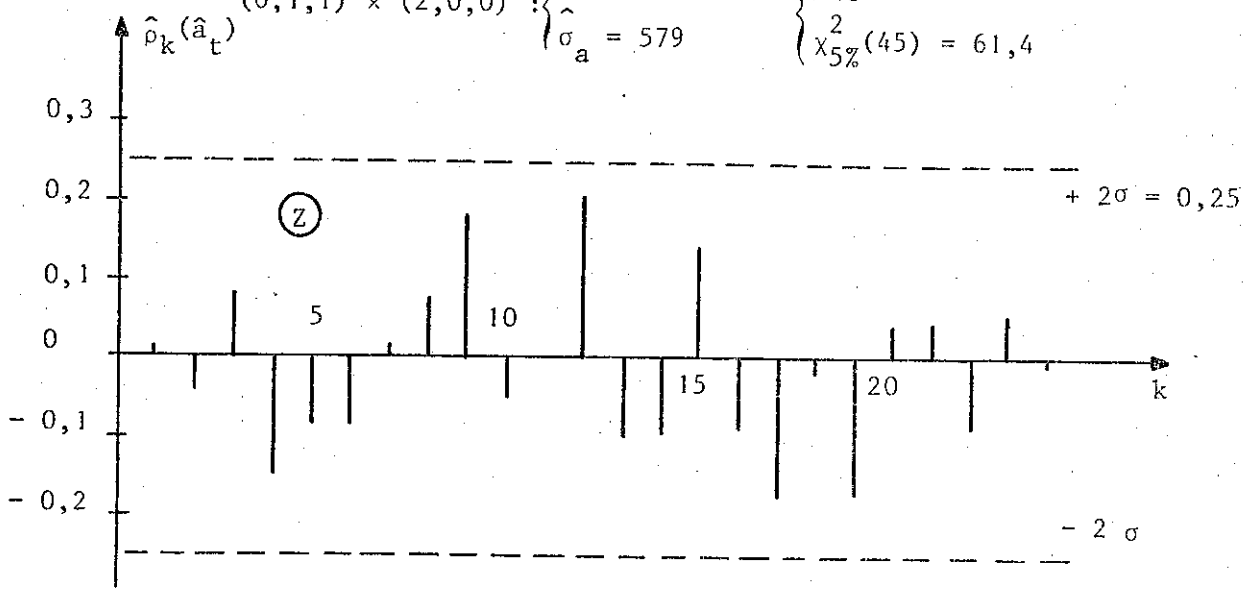
** BOX & JENKINS, 1970, pp. 64-78 et 174-178

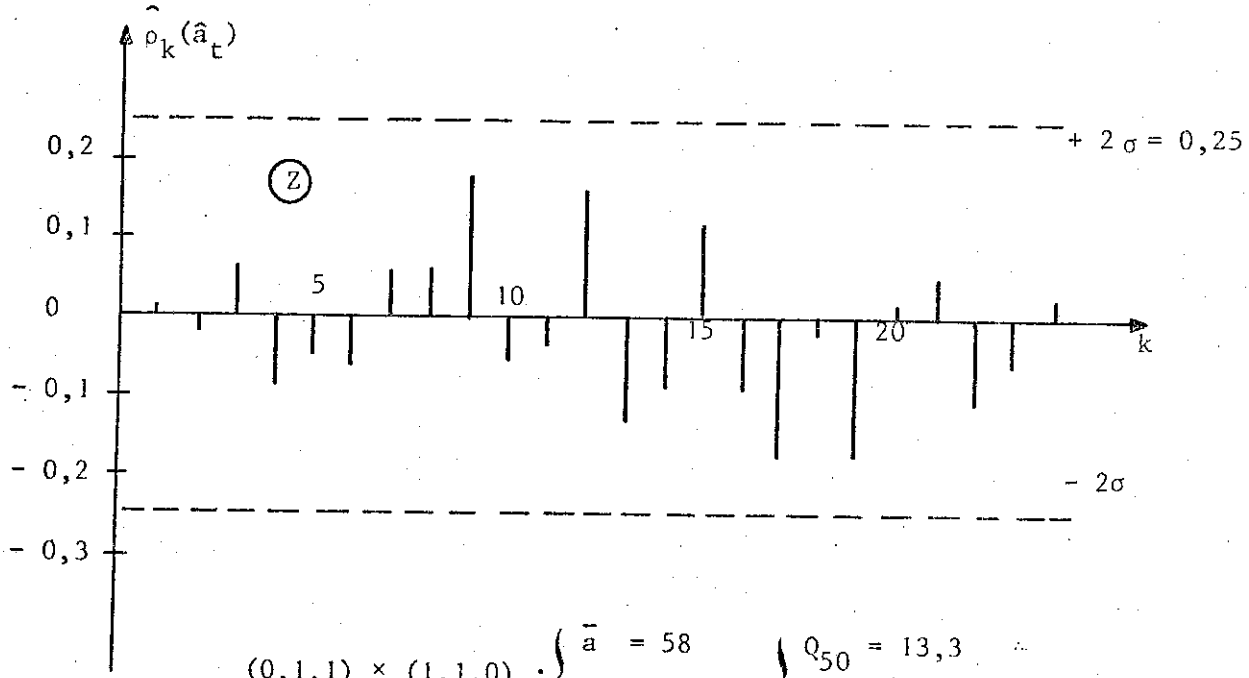
*** BOX & JENKINS, 1970, p. 320.

a) Série Z $\begin{cases} (1 + 0{,}314\ B^{12})\ (1 - 0{,}918\ B^{12})\nabla z_t = (1 - 0{,}615\ B)a_t \\ (1 + 0{,}366\ B^{12})\nabla\nabla_{12}\ z_t = (1 - 0{,}633\ B)\ a_t \end{cases}$

Le diagnostic est excellent dans les deux cas :

$(0,1,1) \times (2,0,0)$: $\begin{cases} \bar{a} = 62 \\ \hat{\sigma}_a = 579 \end{cases}$ $\begin{cases} Q_{48} = 17{,}8 \\ \chi^2_{5\%}(45) = 61{,}4 \end{cases}$

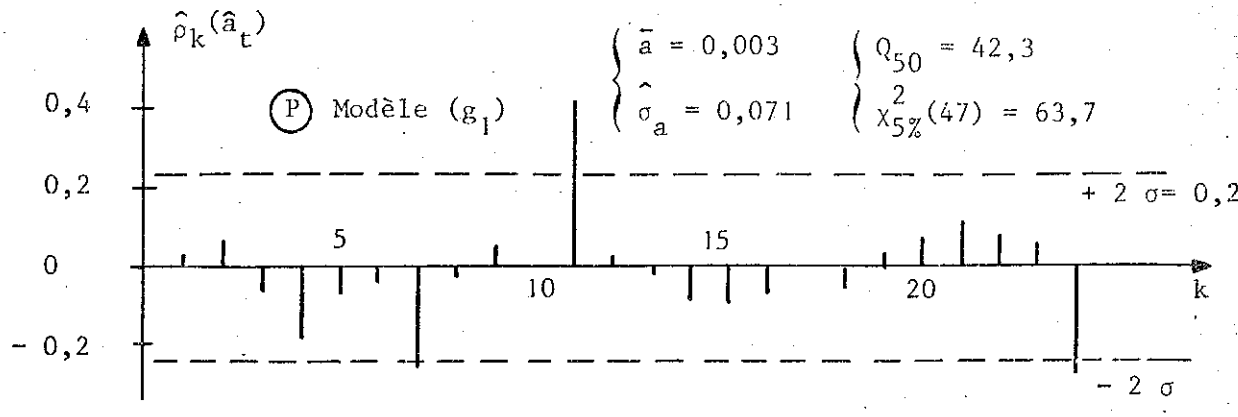




$(0,1,1) \times (1,1,0)$: $\begin{cases} \bar{a} = 58 \\ \hat{\sigma}_a = 581 \end{cases}$ $\begin{cases} Q_{50} = 13{,}3 \\ \chi^2_{5\%}(48) = 64{,}9 \end{cases}$

b) Série P

Le diagnostic du modèle (g_1) est le suivant :



On trouve trois valeurs fortes : $\hat{\rho}_7 = - 0,25$; $\hat{\rho}_{11} = 0,42$; $\hat{\rho}_{24} = - 0,27$; mais, sur un ensemble de 24 coefficients, ce n'est pas trop surprenant et on peut donc considérer le modèle comme acceptable.

Le diagnostic du modèle (b_1) est très voisin de celui du modèle $g_1)$, avec $\bar{a} = 0,002$, $\hat{\rho}_a = 0,067$ et $Q_{50} = 34$. Il présente, lui aussi, deux valeurs fortes $\hat{\rho}_7 = - 0,29$ et $\hat{\rho}_{11} = 0,40$.

Notons au passage que les modèles (a) et (e_1), par exemple, ont eux aussi des corrélogrammes résiduels qui comportent quelques valeurs fortes :

Modèle (a)	: $\hat{\rho}_7 = - 0,29$	$\hat{\rho}_{11} = 0,34$	
Modèle (e_1)	: $\hat{\rho}_3 = 0,25$	$\hat{\rho}_{11} = 0,23$	$\hat{\rho}_{12} = - 0,23$

Quel que soit le modèle, on trouve toujours une valeur résiduelle $\hat{\rho}_{11}$ élevée.

5.6. - Conclusion

L'utilisateur familier des notions de corrélogramme peut trouver dans un diagnostic négatif des informations quant à la direction dans laquelle la forme du modèle doit être modifiée. Mais il ne s'agit, en pratique que d'un cas relativement rare, car la validation ne semble pas être une épreuve très difficile. Bien des modèles ont été rejetés au cours de notre processus d'identification par élimination progressive qui présentaient pourtant un diagnostic favorable. Et la validation ne permet souvent pas de départager les modèles finaux (cf. série P) cependant fort divers entre lesquels on peut hésiter.

Ceci renforce l'importance du choix du modèle : c'est de lui que dépend directement la forme de la fonction qui sera utilisée en prévision.

6. - CONSEQUENCES DU CHOIX D'UN MODELE SUR LA PREVISION

6.1. - Calcul des prévisions

L'analyse d'une série temporelle (choix du modèle et validation) ne constitue que rarement une fin en soi. Le véritable objectif poursuivi est la prévision.

Les prévisions $\hat{z}_t(h)$ effectuées à la date t pour l'horizon h (c'est-à-dire pour la date t + h) s'obtiennent à partir du modèle écrit sous sa forme générale en remplaçant chaque terme par son espérance mathématique, soit :

$$\left.\begin{aligned} E[z_j] &= \hat{z}_t(j-t) \\ E[a_j] &= 0 \end{aligned}\right\} \quad \text{pour } j > t$$

$$\left.\begin{aligned} E[z_j] &= z_j \\ E[a_j] &= a_j \end{aligned}\right\} \quad \text{pour } j \leq t$$

Par $a_j (j \leq t)$, on entend ici les valeurs $a_j = \hat{a}_j$ calculées par récurrence et qui correspondent à l'optimum $\hat{S} = \min (\sum a_t^2)$. Ces valeurs doivent donc être conservées en mémoire sur l'ordinateur.

A défaut, on pourra reconstituer les $\hat{a}_j$ utiles à partir des $\hat{a}_*$ optimisés en même temps que les autres paramètres.

L'ensemble des prévisions effectuées à une même date pour différents horizons h s'appelle fonction de prévision. Chaque modèle possède une fonction de prévision caractéristique.

La prévision ainsi obtenue est sans biais. La variance de l'erreur de prévision d'horizon h est donnée par :

$$V(h) = (1 + \psi_1^2 + \psi_2^2 + \ldots + \psi_{h-1}^2)\ \hat{\sigma}_a^2 \qquad (6.1.1)$$

(voir § 2.3 pour la définition des ψ_j).

En particulier, les résidus a_t s'interprètent comme les erreurs de prévision d'horizon 1 commises avec le modèle considéré : la méthode des moindres carrés ($\min S = \Sigma\, a_t^2$) est optimale lorsqu'il s'agit de prévision à l'horizon 1.

6.2. - Exemples de fonctions de prévision

Examinons quelques formes typiques de fonctions de prévision pour différents modèles ARIMA (p,d,q).

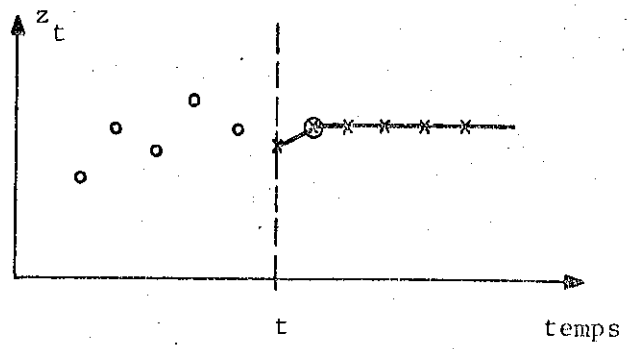
6.2.1. - ARIMA(0,1,1) $\qquad \nabla z_t = (1 - \theta B)a_t$

L'équation aux différences s'écrit :

$$z_{t+h} = z_{t+h-1} + a_{t+h} - \theta a_{t+h-1}$$

soit, en prenant l'espérance mathématique à la date t :

$$\left|\begin{array}{ll} \hat{z}_t(1) = z_t - \theta a_t & \\ \ldots \quad\quad \ldots & \\ \hat{z}_t(h) = \hat{z}_t(h-1) = \hat{z}_t(1) & h \geq 2 \end{array}\right.$$



La fonction de prévision est constante :

$$\hat{z}_t(h) = z_t - \theta a_t$$

a_t est l'erreur de prévision d'horizon 1 faite à la date t-1 :

$$a_t = z_t - \hat{z}_{t-1}(1) = z_t - \hat{z}_{t-1}(h)$$

donc $$\hat{z}_t(h) = (1 - \theta) z_t + \theta\, \hat{z}_{t-1}(h)$$

On reconnaît la formule de mise à jour très simple caractéristique du lissage exponentiel (avec la valeur θ qui correspond à l'optimum pour l'horizon 1).

6.2.2. - <u>ARIMA (1,1,0)</u> $\qquad (1 - \varphi B)\nabla z_t = \theta_o + a_t$

Avec un terme constant θ_o pour plus de généralité (cf. § 2.1) l'équation aux différences s'écrit :

$$z_{t+h} = (\varphi + 1) z_{t+h-1} - \varphi\, z_{t+h-2} + \theta_o + a_{t+h}$$

soit, en prenant l'espérance mathématique à la date t :

$$\left|\begin{array}{l} \hat{z}_t(1) = (\varphi + 1)\, z_t - \varphi\, z_{t-1} + \theta_o \\ \hat{z}_t(2) = (\varphi + 1)\, \hat{z}_t(1) - \varphi\, z_t + \theta_o \\ \ldots \\ \hat{z}_t(h) = (\varphi + 1)\, \hat{z}_t(h-1) - \varphi\, \hat{z}_t(h-2) + \theta_o \qquad h \geq 3 \end{array}\right.$$

et, par addition :

$$\hat{z}_t(h) = \varphi\, \hat{z}_t(h-1) + z_t - \varphi\, z_{t-1} + h\, \theta_o$$

qui a pour solution :

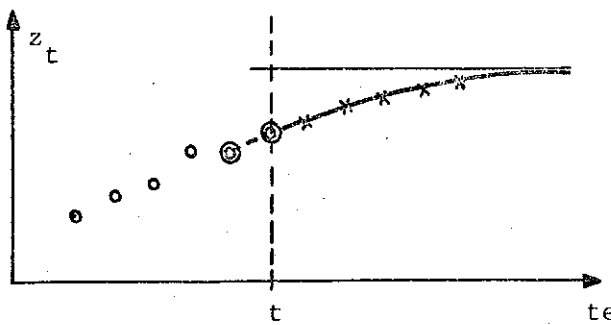
$$\hat{z}_t(h) = \varphi^h\, z_t + (z_t - \varphi\, z_{t-1} + h\, \theta_o)\, \frac{1-\varphi^h}{1-\varphi}$$

ou $$\hat{z}_t(h) = z_t + \frac{(1-\varphi^h)\varphi}{1-\varphi}\,(z_t - z_{t-1}) + h\, \theta_o\, \frac{1-\varphi^h}{1-\varphi}$$

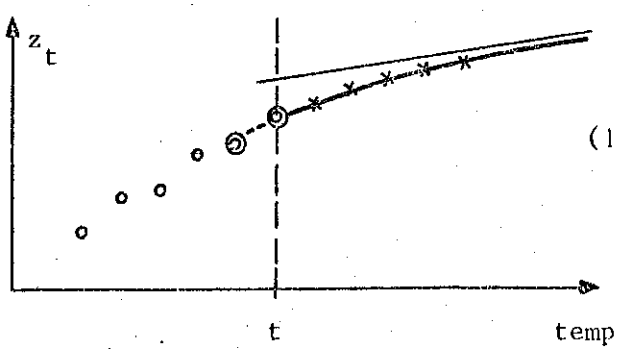
et, comme $|\varphi| < 1$, la fonction de prévision est asymptote à la droite

$$y_t(h) = z_t + \frac{\varphi}{1-\varphi}\,(z_t - z_{t-1}) + h\, \frac{\theta_o}{1-\varphi}$$

Par exemple :



$(1 - 0{,}8\ B)\ \nabla z_t = a_t$



$(1 - 0{,}8\ B)\nabla z_t = 0{,}04 + a_t$

6.2.3. - ARIMA (1,0,1) $(1 - \varphi B)\tilde{z}_t = (1 - \theta B)a_t$ avec $\tilde{z}_t = z_t - \bar{z}$

L'équation aux différences s'écrit :

$$\tilde{z}_{t+h} = \varphi\, \tilde{z}_{t+h-1} + a_{t+h} - \theta\, a_{t+h-1}$$

soit, en prenant l'espérance mathématique à la date t :

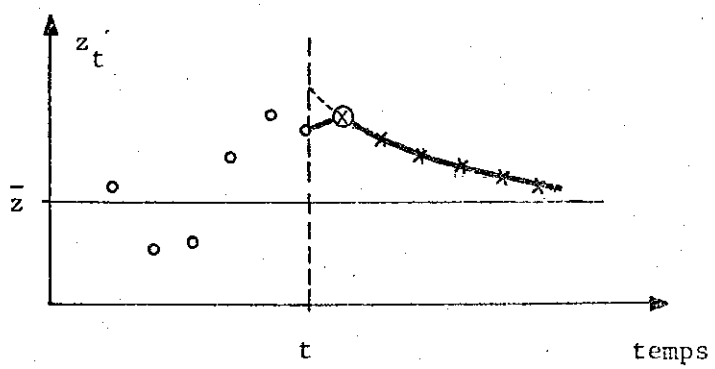
$$\begin{cases} \hat{\tilde{z}}_t(1) = \varphi\, \tilde{z}_t - \theta\, a_t \\ \dots \\ \hat{\tilde{z}}_t(h) = \varphi\, \hat{\tilde{z}}_t(h-1) \qquad\qquad h \geq 2 \end{cases}$$

La fonction de prévision est donc :

$$\hat{\tilde{z}}_t(h) = \varphi^h\, \tilde{z}_t - \theta\, \varphi^{h-1}\, a_t$$

Elle se rapproche exponentiellement de la moyenne $\bar{z}$ à partir de la première prévision $\hat{\tilde{z}}_t(1)$.

Par exemple, pour le modèle $(1 - 0{,}8\ B)\tilde{z}_t = (1 - 0{,}4\ B)a_t$, on a typiquement :



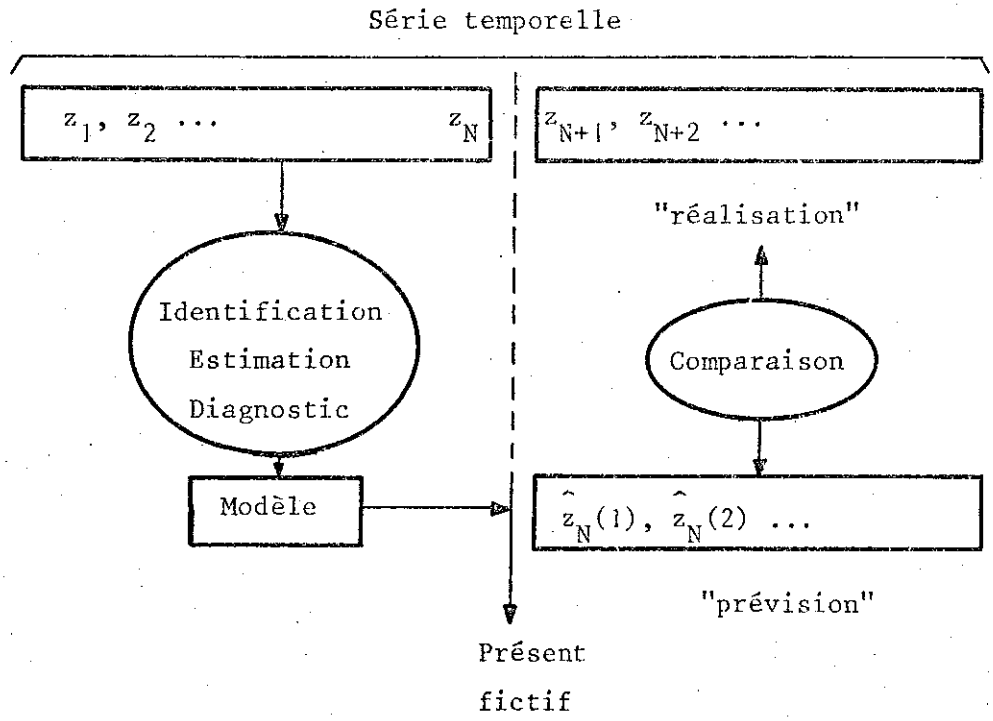
6.3. - Comparaison des prévisions

Une méthode d'estimation ne se juge pas sur le nombre de décimales qu'elle permet de retenir pour chaque paramètre. En fait, du point de vue de la prévision, une estimation avec un ou deux chiffres après la virgule suffit largement : les prévisions sont très stables par rapport au choix des valeurs des paramètres. Cette stabilité est d'ailleurs un trait fondamental de la méthode de Box & Jenkins. Elle permet de ne pas avoir besoin de réestimer le modèle chaque fois qu'une nouvelle valeur de la série devient disponible.

C'est plutôt la forme du modèle qui conditionne l'allure de la fonction de prévision et il est intéressant d'effectuer des comparaisons sur les séries pour lesquelles plusieurs modèles sont restés en concurrence (séries Z et P).

Jusqu'à présent, les différents modèles envisagés pour décrire une même série étaient comparés sur la base de leur adéquation à la série : somme résiduelle faible, indépendance des résidus, etc... Pour comparer les performances en prévision de différents modèles, il faut pouvoir rapprocher les prévisions et la réalité.

On utilisera donc un artifice qui consiste à n'employer qu'une partie de la série pour la phase d'analyse et à simuler une prévision qui pourra être comparée aux valeurs réelles gardées en réserve.



Cette comparaison peut se faire par inspection graphique, ou de manière plus formelle, à partir des écarts prévisions/réalisations $\hat{z}_N(h) - z_{N+h}$, par exemple en calculant l'écart quadratique moyen :

$$E.Q.M. = \sqrt{\frac{\sum_{h=1}^{H} (\hat{z}_N(h) - z_{N+h})^2}{H}} \qquad (6.3.1)$$

a) Série Z

1er modèle : $(1 + 0{,}314\ B^{12})$ $(1 - 0{,}918\ B^{12})\nabla z_t = (1 - 0{,}615\ B)\ a_t$
[0,1] [0,1] [0,1]

2ème modèle : $(1 + 0{,}366\ B^{12})\nabla\nabla_{12}\ z_t = (1 - 0{,}633\ B)\ a_t$
[0,1] [0,1]

N = 64 observations utilisées pour l'estimation

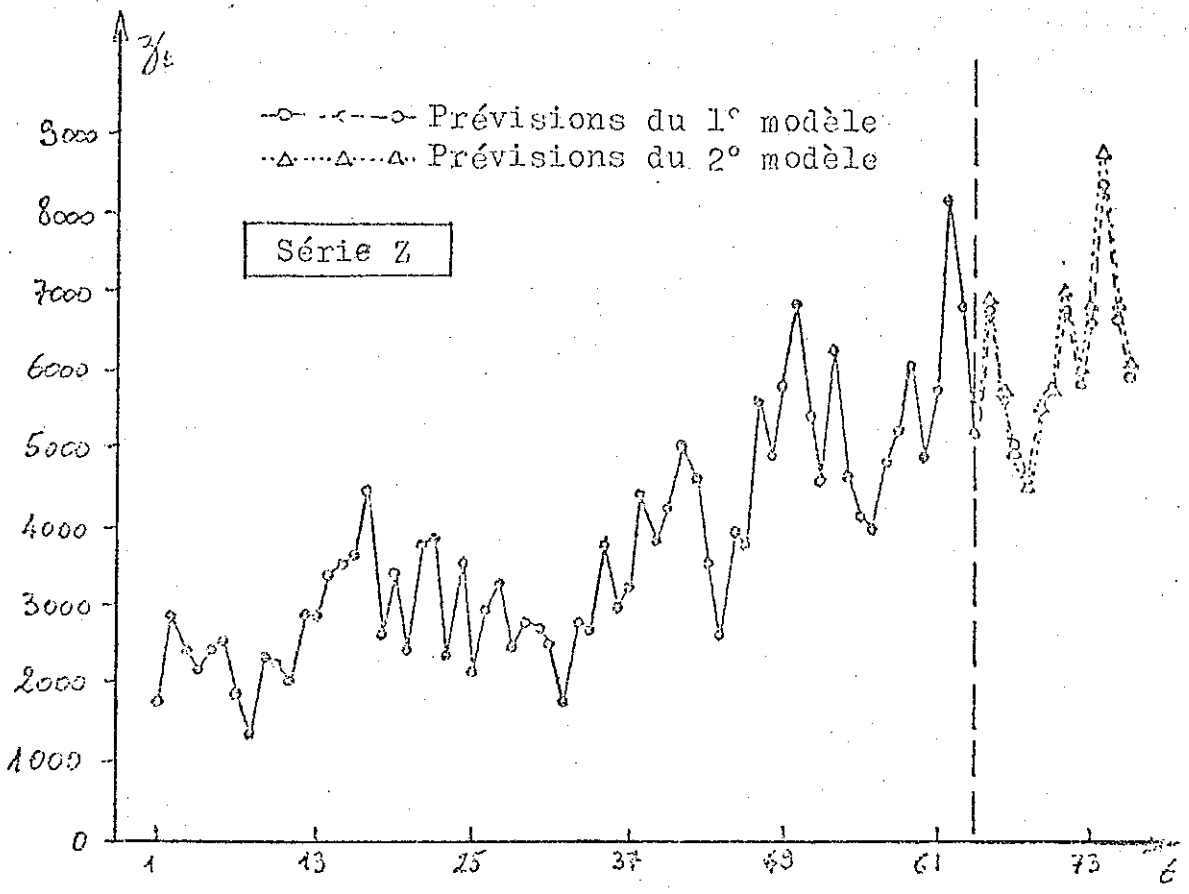
Prévisions effectuées à la date 64 pour 12 périodes.

	h	1	2	3	4	5	6
$\hat{z}_{64}(h)$	1er modèle	6736	5630	5036	4692	5556	5748
	2ème modèle	6866	5679	4989	4570	5556	5750

	h	7	8	9	10	11	12
$\hat{z}_{64}(h)$	1er modèle	6798	5848	6633	8397	6590	5938
	2ème modèle	6971	5925	6807	8723	6738	5999

Le petit nombre d'observations disponibles pour cette série ne permet pas de simuler des prévisions pour les comparer à la réalité. On constate néanmoins que les deux fonctions de prévision sont très voisines et ont l'air tout à fait satisfaisantes.

Les premières prévisions sont quasiment identiques, mais peu à peu apparaissent quelques petites différences dues à l'emploi des deux opérateurs $(1 - 0{,}9\ B^{12})\nabla$ et $\nabla_{12}\ \nabla$. Pour la prévision à plusieurs périodes, on peut préférer le deuxième modèle (en $\nabla\nabla_{12}$) dans lequel les oscillations saisonnières ne sont pas amorties à long terme.



b) Série P

N = 77 observations utilisées pour l'estimation.

Prévisions effectuées à la date 77 pour 6 périodes; nous négligeons le biais introduit par la transformation logarithmique (environ 1% pour h = 1). Tous les modèles comparés ont été introduits au § 4.5.

- Modèles estimés par la méthode des prévisions à rebours

	h	1	2	3	4	5	6	E.Q.M.
(a_o) : $(1+0{,}47\ B)\nabla\nabla_{12}\ \log z_t = (1-0{,}81\ B^{12})a_t$		305	482	673	990	1297	1387	362
(a) : idem avec $a_* \neq 0$ (voir ci-dessous)		–	–	–	–	–	1221	–
(b) : $(1+0{,}51\ B)(1+0{,}47\ B^{12})\nabla\nabla_{12}\ \log z_t = a_t$		–	–	452	–	–	990	–
(c) : $(1+0{,}56\ B^{12})\nabla\nabla_{12}\ \log z_t = (1-0{,}49\ B)a_t$		–	–	–	–	–	940	–
(d) : $\nabla\nabla_{12}\ \log z_t = (1-0{,}44\ B)(1-0{,}85\ B^{12})a_t$		–	–	–	–	–	1225	–
(a') : $(1+0{,}5\ B)\nabla\nabla_{12}\ z_t^{1/4} = (1-0{,}8\ B^{12})a_t$		286	409	511	761	966	1091	147
(f) : $(1-0{,}50\ B)\nabla\nabla_{12}\ z_t = 0{,}50\ \exp(0{,}016\ t)a_t$		262	368	428	711	887	926	69
	Réalisation	260	304	390	614	783	872	

Pour le premier modèle, les prévisions notées (a_o) sont effectuées en prenant $a_* = 0$. Cette approximation est bien sûr insuffisante, surtout avec $\hat{\Theta}$ proche de 1, et les autres prévisions utilisent les valeurs a_* prévues à rebours pour les premiers résidus. Les modèles et les prévisions sont ceux de Chatfield & Prothero (1973) [modèles (a_o), (a), (b), (c), (d)], Box & Jenkins (1973) [modèle (a')] et Melard (1977) [modèle (f)].

- Modèles estimés avec optimisation simultanée des premiers résidus

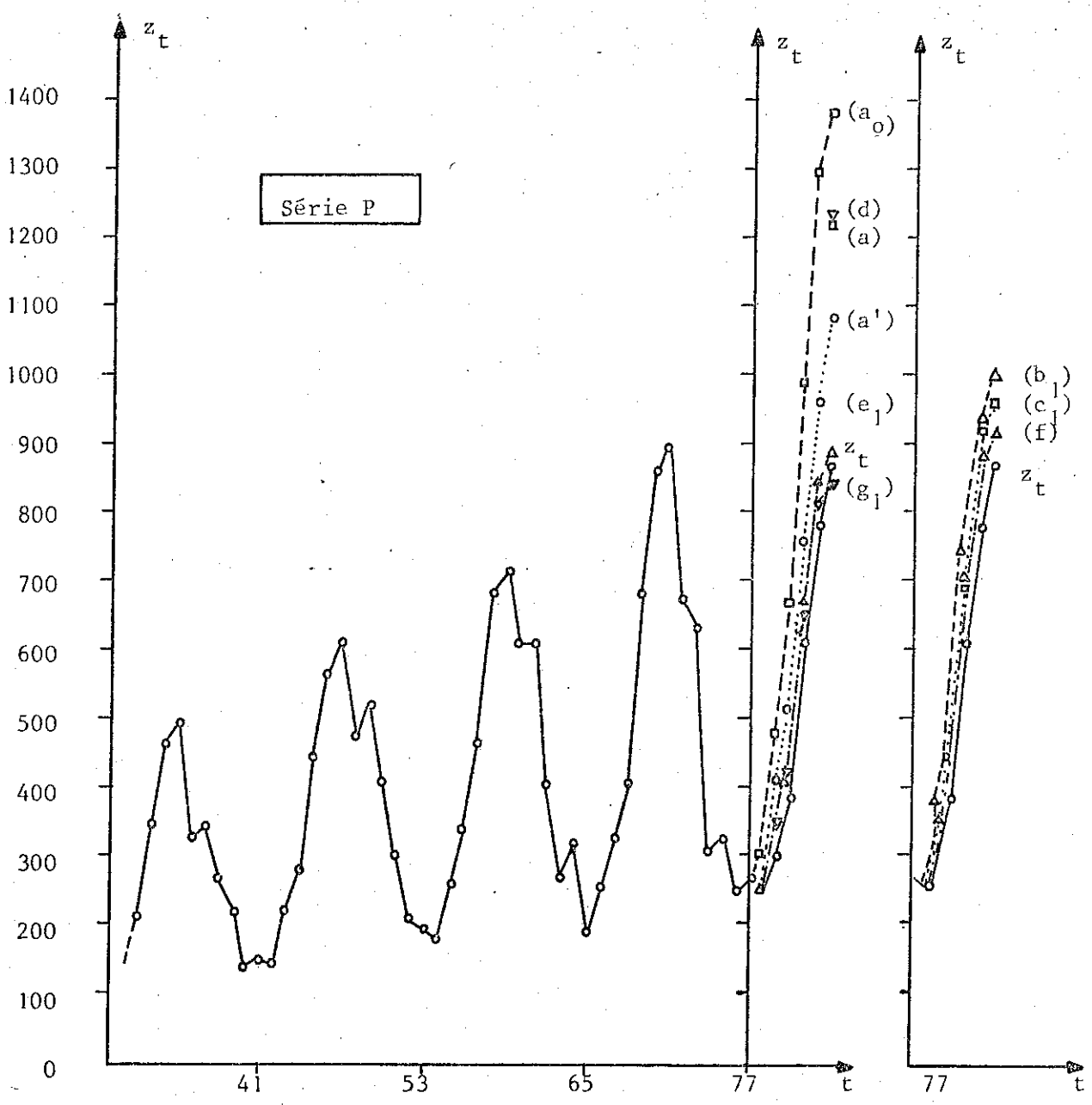
	h	1	2	3	4	5	6	E.Q.M
(b_1) : $(1+0{,}60\ B)(1+0{,}23\ B^{12})\nabla\nabla_{12} \log z_t = a_t$		249	384	438	748	941	1007	94
(c_1) : $(1+0{,}35\ B^{12})\nabla\nabla_{12} \log z_t = (1-0{,}51\ B)a_t$		269	345	447	694	926	967	83
(e_1) : $\nabla_{12} \log z_t = (1+0{,}43\ B + 0{,}72\ B^2)\ a_t$		248	402	404	677	858	895	58
(g_1) : $(1+0{,}56\ B)(1-0{,}63\ B^{12} - 0{,}24\ B^{24})\nabla \log z_t = a_t$		254	362	416	646	801	843	32

Ici, les prévisions sont toutes effectuées en utilisant les valeurs $\hat{a}_*$ des premiers résidus à l'optimum. Les modèles (a_1), (a_1') et (d_1) ont été rejetés car ils correspondent à $\hat{\Theta} = 1$. Le modèle (e_1) est très proche du modèle (e) retenu par Anderson (1975a) et le modèle (g_1) est le modèle auquel on aboutit en pratiquant l'identification par élimination progressive.

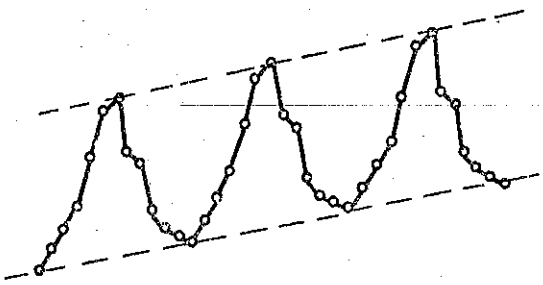
Chatfield & Prothero n'indiquent pas toutes les prévisions de h = 1 à 6, mais les valeurs de $\hat{z}_{77}(6)$ pour novembre 1971 suffisent pour permettre la comparaison de tous les modèles.

Les prévisions (a), (a_o), (a') et (d) sont manifestement "fortes". Elles proviennent toutes de modèles que nous avons trouvés mal estimés, puisque l'estimation avec optimisation simultanée des premiers résidus conduit à $\hat{\Theta} = 1$.

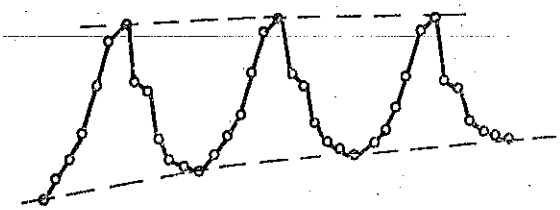
Les autres prévisions sont plus "raisonnables" et, en tous cas, plus proches de la réalisation. Le modèle (g_1) que nous avons identifié et estimé fournit les meilleures prévisions si on s'en tient au critère de l'écart quadratique moyen prévision/réalisation.



L'écart quadratique moyen n'est bien sûr pas un critère définitif et la réalisation elle-même peut sembler un peu "faible". Ceci pose le problème de la prévision à plusieurs périodes : les modèles (b_1) et (c_1) correspondent à l'opérateur $(1 - B)(1 - B^{12})$ qui produit des oscillations saisonnières entretenues, alors que le modèle (g_1) correspond à l'opérateur $(1 - B)(1 - 0,9\ B^{12})$ qui produit des oscillations amorties.



$(1 - B)(1 - B^{12})$



$(1 - B)\ (1 - 0{,}9\ B^{12})$

Il n'est donc pas étonnant que les prévisions du modèle (g_1) soient systématiquement plus faibles que celles des modèles (b_1) et (c_1), même pour des horizons modérés.

7. - ANNEXE : REVUE DES MODELES RETENUS POUR QUELQUES SERIES

Nous présentons une revue rapide des modèles que nous avons retenus pour un certain nombre de séries* , y compris celles déjà envisagées ici à titre d'exemples.

Les chiffres entre crochets au-dessous des paramètres d'un modèle indiquent l'écart-type de l'estimation correspondante. La variance résiduelle à l'optimum $\hat{\sigma}_a^2$ est calculée comme indiqué au § 3.2.3 (du moins pour les modèles que nous avons nous-mêmes estimés).

---

* TRAVERS, 1978

Série A : CONCENTRATIONS CHIMIQUES RELEVEES TOUTES LES DEUX HEURES - 197 observations

Référence : BOX & JENKINS, 1970

Deux modèles retenus :

$$\begin{cases} (1 - 0{,}91\ B)\tilde{z}_t = (1 - 0{,}60\ B)a_t & \hat{\sigma}_a^2 = 0{,}097 \\ \quad [\pm 0{,}04] \qquad\qquad [\pm 0{,}08] & \\ \nabla z_t = (1 - 0{,}71\ B)a_t & \hat{\sigma}_a^2 = 0{,}101 \\ \quad [\pm 0{,}05] & \end{cases}$$

Modèles de BOX & JENKINS :

$$\begin{cases} (1 - 0{,}92\ B)z_t = 1{,}45 + (1 - 0{,}58\ B)a_t & \hat{\sigma}_a^2 = 0{,}097 \\ \quad [\pm 0{,}04] \qquad\qquad [\pm 0{,}08] & \\ \nabla z_t = (1 - 0{,}70\ B)a_t & \hat{\sigma}_a^2 = 0{,}101 \\ \quad [\pm 0{,}05] & \end{cases}$$

Série B : COURS JOURNALIERS (FERMETURE) DES ACTIONS IBM : 17 MAI 1961 - 2 NOVEMBRE 1962 - 369 observations

Référence : BOX & JENKINS, 1970

Deux modèles retenus (cf. § 4.2 et 4.3.2)

$$\begin{cases} (1 - 0{,}09\ B)\nabla z_t = a_t & \hat{\sigma}_a^2 = 52{,}3 \\ \quad [\pm 0{,}05] & \\ \nabla z_t = (1 + 0{,}09\ B)a_t & \hat{\sigma}_a^2 = 52{,}2 \\ \quad [\pm 0{,}05] & \end{cases}$$

Modèle de BOX & JENKINS :

$$\nabla z_t = (1 + \underset{\pm 0,05}{0,09}\ B)\ z_t = a_t \qquad \hat{\sigma}_a^2 = 52,2$$

Série E : TAUX D'EPARGNE DES MENAGES EN FRANCE (%) : TRIMESTRIEL, 1962 à 1974
52 observations

Référence : Indicateurs du 7ème plan, avril 1977 (INSEE)

$$(1 - \underset{[\pm 0,10]}{0,72}\ B)\tilde{z}_t = a_t \qquad \hat{\sigma}_a^2 = 0,80$$

Série F : MATERIEL FERROVIAIRE : INDICE MENSUEL CORRIGE DES VARIATIONS SAISONNIERES, JANVIER 1963 à SEPTEMBRE 1975
153 observations

Référence : Bulletin mensuel de statistique (INSEE)
Les indices de la production industrielle base 100 en 1970, Les Collections de l'INSEE, E 35, février 1976.

$$\nabla z_t = (1 - \underset{[\pm 0,05]}{0,74}\ B)a_t \qquad \hat{\sigma}_a^2 = 153$$

Série G : TRAFIC AERIEN INTERNATIONAL MENSUEL (Milliers de passagers) : JANVIER 1949 - DECEMBRE 1960.
144 observations

Référence : BOX & JENKINS, 1970

Modèle retenu (cf. § 4.4.4.a)) :

$$\nabla\nabla_{12} \text{Log } z_t = (1 - \underset{[\pm 0,08]}{0,396}\ B)\ (1 - \underset{[\pm 0,07]}{0,614}\ B^{12})a_t \qquad \hat{\sigma}_a^2 = 0,00139$$

Modèle de BOX & JENKINS :

$$\nabla\nabla_{12} \text{ Log } z_t = (1 - 0{,}396\ B)\ (1 - 0{,}614\ B^{12})a_t \qquad \hat{\sigma}_a^2 = 0{,}00134^*$$
$$[\pm 0{,}08] \qquad [\pm 0{,}07]$$

Modèle d'ANDERSON, 1976 :

$$\nabla\nabla_{12} \text{ Log } z_t = (1 - 0{,}362\ B)\ (1 - 0{,}624\ B^{12})a_t \qquad \hat{\sigma}_a^2 = 0{,}00132$$
$$[\pm 0{,}08] \qquad [\pm 0{,}07]$$

Modèle de LENZ, 1978 :

$$\nabla\nabla_{12} \text{ Log } z_t = (1 - 0{,}4\ B - 0{,}05\ B^2 - 0{,}2\ B^3)(1 - 0{,}65\ B^{12})a_t$$
$$\hat{\sigma}_a^2 = 0{,}00129$$

Série H : IMMATRICULATIONS AUTOMOBILES MENSUELLES AUX U.S.A (VEHICULES NEUFS) : JANVIER 1947 - DECEMBRE 1968.
264 observations

Référence : Nelson, 1975

Modèle retenu :

$$\nabla\nabla_{12} \text{ Log } z_t = (1 - 0{,}211\ B - 0{,}261\ B^2)(1 - 0{,}845\ B^{12})a_t \qquad \hat{\sigma}_a^2 = 0{,}01226$$
$$[\pm 0{,}061] \quad [\pm 0{,}061] \qquad [\pm 0{,}034]$$

Modèle de NELSON :

$$\nabla\nabla_{12} \text{ Log } z_t = (1 - 0{,}211\ B - 0{,}261\ B^2)(1 - 0{,}847\ B^{12})a_t \qquad \hat{\sigma}_a^2 = 0{,}01228$$
$$[\pm 0{,}061] \quad [\pm 0{,}061] \qquad [\pm 0{,}034]$$

Série I : MATERIEL AGRICOLE : INDICE MENSUEL CORRIGE DES VARIATIONS SAISONNIERES, JANVIER 1963 A MAI 1976.
161 observations

Référence : voir série F

---

* On remarquera les valeurs différentes de $\hat{\sigma}_a^2$ pour un même modèle : cela tient aux façons différentes de calculer la variance S/ν (différents ν)

Deux modèles retenus :

$$\begin{cases} (1 - 0{,}93\ B)\tilde{z}_t = (1 - 0{,}45\ B)a_t & \hat{\sigma}_a^2 = 52 \\ \quad [\pm 0{,}035] \qquad\quad [\pm 0{,}09] & \\ \nabla z_t = (1 - 0{,}49\ B)a_t & \hat{\sigma}_a^2 = 54 \\ \quad [\pm 0{,}07] & \end{cases}$$

Série M : CONSOMMATIONS ELECTRIQUES FRANCAISES DU JOUR OUVRABLE MOYEN MENSUEL CORRIGEES DE L'EFFET DE TEMPERATURE : JANVIER 1951 - JUIN 1958
90 observations

Référence : MOGHA, 1977

Modèle retenu

$$\nabla z_t = 0{,}85 + (1 - 0{,}26\ B)a_t \qquad \hat{\sigma}_a^2 = 2{,}22$$
$$[\pm 0{,}1]$$

Modèle de MOGHA (parmi les modèles de cet auteur, le meilleur - Méthode d'approximations successives)

$$\nabla z_t = 0{,}86 + (1 - 0{,}21\ B)a_t \qquad \hat{\sigma}_a^2 = 2{,}25$$
$$[\pm 0{,}1]$$

Série N : PRODUIT NATIONAL BRUT AUX U.S.A. AUX PRIX COURANTS DU MARCHE : 1er TRIMESTRE 1947 - 4ème TRIMESTRE 1966; MILLIARDS DE DOLLARS, TAUX ANNUELS
80 observations (corrigé des variations saisonnières)

Référence : NELSON, 1975

Modèle retenu (cf, § 4.3.4)

$$(1 - 0{,}62\ B)\nabla z_t = 2{,}78 + a_t \qquad \hat{\sigma}_a^2 = 223$$
$$[\pm 0{,}09]$$

Modèle de NELSON

$$(1 - 0{,}62\ B)\nabla z_t = 2{,}69 + a_t \qquad \hat{\sigma}_a^2 = 224$$
$$[\pm 0{,}09] \qquad [\pm 0{,}08]$$

Série P : VENTES MENSUELLES D'UNE ENTREPRISE X : JANVIER 1965 - MAI 1971
77 observations

Référence : CHATFIELD & PROTHERO, 1973

Se reporter pour le choix du modèle : § 4.5
pour la validation : § 5.5. b)
pour les prévisions : § 6.3. b)

Modèles retenus :

(g_1) $(1 + 0{,}561\ B)\ (1 - 0{,}631\ B^{12} - 0{,}237\ B^{24})\nabla\ \log_{10} z_t = a_t$ $\quad \hat{\sigma}_a^2 = 0{,}0052$
[±0,09] [±0,11] [±0,11]

(b_1) $(1 + 0{,}601\ B)\ (1 + 0{,}229\ B^{12})\nabla\nabla_{12}\ \log_{10} z_t = a_t$ $\quad \hat{\sigma}_a^2 = 0{,}0057$
[±0,10] [±0,12]

Modèles de CHATFIELD & PROTHERO *:

(a) $(1 + 0{,}47\ B)\nabla\nabla_{12}\ \log_{10} z_t = (1 - 0{,}81\ B^{12})a_t$ $\quad \hat{\sigma}_a^2 = 0{,}0053$
[±0,11] [±0,07]

(b) $(1 + 0{,}51\ B)\ (1 + 0{,}47\ B^{12})\nabla\nabla_{12}\ \log_{10} z_t = a_t$ $\quad \hat{\sigma}_a^2 = 0{,}0083$
[±0,11] [±0,11]

(c) $(1 + 0{,}56\ B)\nabla\nabla_{12}\ \log_{10} z_t = (1 - 0{,}49\ B)a_t$ $\quad \hat{\sigma}_a^2 = 0{,}0085$
[±0,10] [±0,11]

(d) $\nabla\nabla_{12}\ \log_{10} z_t = (1 - 0{,}44\ B)\ (1 - 0{,}85\ B^{12})a_t$ $\quad \hat{\sigma}_a^2 = 0{,}0054$
[±0,11] [±0,07]

Autres modèles : (a_o) (Chatfield & Prothero), (a') (Box-Jenkins), (f) (Melard) : voir § 6.3.b.

---

* Les variances résiduelles $\hat{\sigma}_a^2 = \hat{S}/n$ fournies par Chatfield & Prothero ont été recalculées ici par la formule $\hat{\sigma}_a^2 = \hat{S}/\nu$ (cf. § 3.2.3). Pour les modèles b) et b_1), par exemple, on a $n = 64$ et $\nu = 51$.

Série T : TRAFIC MARCHANDISES SNCF EN TONNES-KILOMETRES - JANVIER 1971 - DECEMBRE 1976
72 observations

Référence : Tendances de la conjoncture (INSEE)

$$(1 - \underset{[\pm 0,09]}{0,61}\ B^{12})\nabla z_t = (1 - \underset{[\pm 0,10]}{0,54}\ B)a_t \qquad \hat{\sigma}_a^2 = 198.10^3$$

Série X : TAUX DE COUVERTURE DES ECHANGES EXTERIEURS EN FRANCE (%) : TRIMESTRIEL, 1962 à 1974
52 observations

Référence : Indicateurs du 7° plan, avril 1977 (INSEE)

$$(1 - \underset{[\pm 0,10]}{0,66}\ B)\tilde{z}_t = a_t \qquad \hat{\sigma}_a^2 = 13,6$$

Série Z : VENTES MENSUELLES D'UNE ENTREPRISE FRANCAISE - JANVIER 1973 - AVRIL 1978
64 observations

Pas de référence : l'entreprise a tenu à rester anonyme.

1786.3 2928.5 2473.5 2196.4 2418.1 2471.3 1899.7 1348.4 2372.6 2307.3
2028.4 2872.8 2853.0 3440.6 3525.9 3655.2 4476.3 2593.0 3395.4 2466.1
3723.1 3844.0 2348.2 3551.1 2125.6 2967.9 3265.9 2486.4 2759.0 2674.2
2538.6 1701.7 2770.2 2684.4 3728.7 2986.4 3211.8 4419.8 3815.2 4218.6
5015.0 4612.9 3565.6 2636.1 3957.6 3771.6 5654.9 4894.1 5788.4 6803.2
5437.7 4567.0 6251.2 4611.8 4128.5 4003.5 4803.1 5209.1 6048.5 4838.2
5712.9 8148.1 5807.2 5143.0 ——————→

Deux modèles retenus [cf. § 4.4.4.b); 5.5.a) et 6.3.a)]

$$\begin{cases} (1 - \underset{[\pm 0,1]}{0,60}\ B^{12} - \underset{[\pm 0,1]}{0,29}\ B^{24})\nabla z_t = (1 - \underset{[\pm 0,1]}{0,62}\ B)a_t & \hat{\sigma}_a^2 = 556.10^3 \\ (1 + \underset{[\pm 0,2]}{0,37}\ B^{12})\nabla\nabla_{12}\, z_t = (1 - \underset{[\pm 0,2]}{0,63}\ B)a_t & \hat{\sigma}_a^2 = 560.10^3 \end{cases}$$

BIBLIOGRAPHIE

J. ABADIE, F. MESLIER (1977)

Présentation synthétique de modèles de prévision à très court terme de l'énergie journalière produite par Electricité de France et de la température moyenne journalière relevée à Paris-Montsouris. Cahier du LAMSADE n° 10, Université Paris IX Dauphine.

J. ABADIE, F. MESLIER (1979)

Etude de l'utilisation des modèles A.R.I.M.A. pour la prévision à court terme de l'énergie journalière produite par Electricité de France. R.A.I.R.O. Recherche opérationnelle, vol. 13, n° 1, février 1979, p. 37-54

O.D. ANDERSON (1975)

On the collection of time series data. Oper. Res. Quart. 26, p. 331-335.

O.D. ANDERSON (1976)

Time series analysis and forecasting : the Box-Jenkins approach. Butterworths, Londres.

O.D. ANDERSON (1977a)

A commentary on "A survey of time series". Int. Stat. Rev. 45, p. 273-297.

O.D. ANDERSON (1977b)

The interpretation of Box-Jenkins time series models. The Statistician 26, p. 127-145.

O.D. ANDERSON (1978a)

Box-Jenkins rules, O.K. - provided you know it makes sense. Management Science 24, n° 11, p. 1199.

O.D. ANDERSON (1978b)

Orthodox Box-Jenkins versus "Dynamic Data System Modelling" secession - a rejoinder. Management Science 24, n° 11, p. 1203-1205.

T.W. ANDERSON (1958)

The statistical analysis of time series. Wiley, Londres.

G.E.P. BOX, D.R. COX (1964)

An analysis of transformations. Jour. of the Roy. Stat. Soc. B26, p. 211-2

G.E.P. BOX, G.M. JENKINS (1970)

G.E.P. BOX, G.M. JENKINS (1973)
Some comments on a paper by Chatfield and Prothero and a review by Kendall. Jour. of the Roy. Stat. Soc. A135, p. 337-352.

G.E.P. BOX, D.A. PIERCE (1970)
Distribution of residual autocorrelations in autoregressive integrated moving average time series models. Jour. of the Amer. Stat. Assoc. 65, p. 1509-1526.

C. CHATFIELD (1975)
The analysis of time series : theory and practice. Chapman and Hall, Londres.

C. CHATFIELD, D.L. PROTHERO (1973)
Box-Jenkins seasonal forecasting : problems in a case-study. Jour. of the Roy. Stat. Soc. A136, p. 295-336.

G.V. GLASS, V.L. WILLSON, J.M. GOTTMAN (1975)
Design and analysis of time series experiments. Colorado Associated University Press, Boulder, Colorado.

C.W.J. GRANGER, P. NEWBOLD (1976)
Forecasting transformed series. Jour. of the Roy. Stat. Soc. B38, p. 189-203.

D.M. HIMMELBLAU (1972)
Applied nonlinear programming. Mc Graw-Hill, New-York.

M.G. KENDALL (1971)
Review of Box and Jenkins. Jour. of the Roy. Stat. Soc. A134, p. 450-453.

M.G. KENDALL (1973)
Time series. Griffin, Londres.

H.J. LENZ (1978)
Strategies for implementation of a fully automatic Box & Jenkins forecasting technique. Discussionsarbeiten n° 7/78, Institut für Quantitative Ökonomik und Statistik, Freie Universität, Berlin

A.M. MAHE, D. TRAVERS (1976)

Application de méthodes d'optimisation sans contrainte à la détermination des paramètres de modèles ARIMA. Mémoire de D.E.A., U.E.R. Sciences des organisations, Université Paris IX Dauphine.

S. MAKRIDAKIS (1976)

A survey of time series. Int. Stat. Rev. 44, p. 29-70.

S. MAKRIDAKIS (1978)

Time-series analysis and forecasting : an update and evaluation. Int. Stat. Rev. 46, p. 255-278.

S. MAKRIDAKIS, S.C. WHEELWRIGHT (1973)

Forecasting methods for management. Wiley, New-York.

S. MAKRIDAKIS, S.C. WHEELWRIGHT (1978)

Interactive forecasting. Holden Day, San Francisco.

G. MELARD (1977)

Prévision des ventes. Communication au 4° colloque international d'Econométrie Appliquée, Strasbourg. (Faculté des Sciences Sociales et Politiques, Institut de Statistique, Campus Plaine, Université libre de Belgique).

F. MESLIER (1976)

Contribution à l'analyse des séries chronologiques et application à la mise au point de modèles de prévision à court terme relatifs à la demande journalière d'énergie électrique en France et à la température relevée à Paris-Montsouris. Thèse de Doctorat de 3ème cycle, Université Paris IX Dauphine.

G.V. MOGHA (1977)

Sur un problème d'estimation du modèle ARIMA. EDF, Bulletin de la Direction des Etudes et Recherches, série C, suppl. au n° 1, p. 87-132.

T.H. NAYLOR, T.G. SEAKS, D.W. WICHERN (1972)

Box-Jenkins methods : an alternative to econometric models. Int. Stat. Rev. 40, p. 123-137.

C.R. NELSON (1973)

Applied time series analysis for managerial forecasting. Holden Day, San Francisco.

P. NEWBOLD (1975)

The principles of the Box-Jenkins approach. Oper. Res. Quart. 26, p. 397-412.

H.J. STEUDEL, S.M. WU (1977)

A time series approach to queueing systems with applications for modelling job-shop in-process inventories. Management Science 7, p. 745-755.

H.J. STEUDEL, S.M. WU (1978)

Dynamic Data System Modelling-revisited. Management Science 24, p. 1200-1202.

D. TRAVERS (1978)

Utilisation des méthodes d'optimisation dans la prévision de séries temporelles par la méthode de Box et Jenkins. Thèse de 3° cycle, U.E.R. Sciences des Organisations, Université Paris IX Dauphine.