

GUY MÉLARD

OLIVIER ROULAND

Sélection d'une méthode de prévision par l'emploi du modèle ARIMA sous-jacent

RAIRO. Recherche opérationnelle, tome 20, n° 2 (1986),
p. 89-113

http://www.numdam.org/item?id=RO_1986__20_2_89_0

© AFCET, 1986, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Recherche opérationnelle » 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/>

SÉLECTION D'UNE MÉTHODE DE PRÉVISION PAR L'EMPLOI DU MODÈLE ARIMA SOUS-JACENT (*)

par Guy MÉLARD ⁽¹⁾ et Olivier ROULAND ⁽²⁾

Résumé. — *Dans l'ensemble des méthodes de prévision à court terme, les méthodes extrapolatives, faisant intervenir un petit nombre de paramètres et facilement interprétables, se distinguent de la méthode de Box et Jenkins basée sur les modèles ARIMA. Des travaux ont montré que des modèles ARIMA sous-tendent la plupart des méthodes extrapolatives. L'idée de cet article est d'exploiter ces relations (détaillées en annexe) pour estimer les paramètres inclus dans ces méthodes de prévision à l'aide d'un logiciel conçu pour l'estimation des modèles ARIMA. Cette approche permet d'intégrer les deux types de méthodes. Les modèles sont généralisés pour tenir compte d'un mouvement saisonnier, d'hétéroscédasticité, d'une transformation paramétrée de la variable et d'interventions. La procédure est illustrée sur une série chronologique publiée.*

Mots clés : Méthode de Box et Jenkins, méthodes de prévision extrapolatives, lissage exponentiel, maximum de vraisemblance, modèles ARIMA.

Abstract. — *Among short term forecasting methods, extrapolative methods are often opposed to the Box-Jenkins method based on ARIMA models. The extrapolative methods depend on a small number of parameters and provide easy interpretation. It is known that, for most of them, ARIMA processes are underlying. The main contribution of this paper is to use these relations (given in the appendix) in order to estimate the parameters of the forecasting methods by means of a package made for fitting ARIMA models. Both types of methods are thus included in an integrated approach. Models are generalized to cope with a deterministic seasonal component, heteroskedasticity, parametrized transformation of the variate, and interventions. The procedure is illustrated on a published time series.*

Keywords: Box-Jenkins method, extrapolative forecasting methods, exponential smoothing, maximum likelihood, ARIMA models.

(*) Reçu août 1984.

Une première version de cet article a été présentée lors des Journées de Statistique de l'Association des Statisticiens Universitaires (A.S.U.) (Paris, 28-31 mai 1979).

Nous tenons à remercier le lecteur pour les remarques pertinentes qui ont contribué grandement à l'amélioration du texte.

⁽¹⁾ Centre d'Économie Mathématique et d'Économétrie et Institut de Statistique, C.P. n° 210, Campus de la Plaine, boulevard du Triomphe, B-1050 Bruxelles, Belgique.

⁽²⁾ Département d'Économie Appliquée (D.U.L.B.E.A.), C.P. n° 140, avenue F.-D. Roosevelt, 50, B-1050 Bruxelles, Belgique.

1. INTRODUCTION

Dans l'ensemble des méthodes de prévision à court terme (Wheelwright et Makridakis [36]) un certain nombre de méthodes extrapolatives occupent une place particulière (Kaufman et Groboillot [16], Montgomery et Johnson [30]). Elles se caractérisent par une formule de prévision spécifique, faisant intervenir un petit nombre de paramètres facilement interprétables et dont la mise en œuvre ne requiert qu'un minimum d'information sur le passé.

A cette approche s'oppose, *a priori*, celle basée sur la représentation d'une série chronologique par un modèle ARIMA qui est déterminé à l'issue d'une procédure itérative d'identification, d'estimation et de validation (Box et Jenkins [7]). Les deux principaux inconvénients sont la difficulté d'interpréter intuitivement les paramètres et la formule de prévision, et l'intervention humaine répétée dans la procédure de sélection du modèle, bien qu'il existe des versions partiellement automatiques de la méthode (Abadie et Travers [1] et Libert [18]). En revanche, la méthode de Box et Jenkins garantit des estimations correctes des paramètres, permet d'investiguer la validité statistique du modèle, fournit des intervalles de prévision.

Un certain nombre de travaux ont montré qu'un processus stochastique ARIMA d'un certain ordre sous-tend la plupart des méthodes de prévision à court terme. L'objectif de notre travail est d'exploiter ces liens pour estimer les paramètres inclus dans ces méthodes de prévision et tester l'adéquation du modèle ainsi réalisé, en employant un logiciel d'analyse de séries chronologiques par des modèles ARIMA. On combine ainsi les avantages des deux approches. L'utilisateur a la possibilité d'employer les mêmes outils d'analyse et peut donc comparer les résultats de manière rigoureuse. Nous allons plus loin que Makridakis et Wheelwright [21] qui se contentent de fournir les deux approches en parallèle puisqu'ici il y a intégration complète même au niveau de l'estimation des paramètres.

Dans une première partie (§2) nous présentons les méthodes de prévision retenues, les liens avec les modèles ARIMA ainsi que quatre généralisations de ces derniers : le mouvement saisonnier déterministe, l'hétéroscédasticité, la transformation paramétrée et les interventions. La deuxième partie est consacrée à l'estimation des paramètres (§3), à la procédure de sélection d'une méthode de prévision (§4) et à la détermination des prévisions (§5). Dans chaque cas on exploite la relation entre les méthodes de prévisions retenues et les processus ARIMA sous-jacents correspondants. La troisième partie (§6) consiste à illustrer la procédure sur une série chronologique publiée. Les

détails techniques à la base de cet article sont donnés en annexe de sorte que chaque possesseur d'un logiciel de Box et Jenkins peut, s'il le désire, réaliser l'intégration qui fait l'objet de cet article.

2. PRÉSENTATION DES MÉTHODES DE PRÉVISION

Les méthodes de prévision retenues sont celles que sous-tendent un processus ARIMA (p, d, q) où les ordres p, d, q et les coefficients $(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ des polynômes autorégressifs (ϕ) et moyenne mobile (θ) peuvent être déterminés à partir des paramètres de l'algorithme. Ces algorithmes qui sont présentés en détail à l'annexe 1 sont subdivisés en 5 classes :

(1) lissage exponentiel à un paramètre : lissages simple, double, triple et lissage dit du second ordre (Brown [8], [9], Gross et Ray [14], Montgomery et Johnson [30]);

(2) moyenne mobile simple (Montgomery et Johnson [30]);

(3) méthode de Holt-Winters en version additive (Winters [37], Montgomery et Johnson [30], Chatfield [10]);

(4) modèle à deux aléas de Bachelet et Morlat [4] et ses généralisations (Kaufman et Groboillot [16]);

(5) modèles polynomiaux (Harrison [15]).

Les méthodes de prévision considérées donnent toutes lieu à une interprétation simple, soit au travers de la formule de calcul des prévisions, soit par la signification des paramètres, soit encore pour les deux raisons comme dans le cas des méthodes de lissage exponentiel.

Les paramètres du modèle général à deux aléas et des modèles polynomiaux ont une interprétation économique intéressante. Par exemple, le modèle de croissance linéaire appliqué à des données de ventes peut décrire un marché dont la pente de la tendance évolue selon un processus de promenade aléatoire (ou marche au hasard). Les cas particuliers du modèle général à deux aléas, à savoir le modèle de Muth [31], le modèle de Theil et Wage [33] et le modèle de Morlat et Bachelet peuvent s'interpréter selon un schéma analogue.

Une méthode de prévision de plus en plus utilisée est basée sur la représentation de la série chronologique par un modèle ARIMA, défini par l'équation :

$$\phi(B) \nabla^d X_t = \theta(B) \varepsilon_t,$$

où

$$\begin{aligned}\phi(B) &= 1 - \phi_1 B - \dots - \phi_p B^p, \\ \theta(B) &= 1 - \theta_1 B - \dots - \theta_q B^q, \\ \nabla &= 1 - B,\end{aligned}$$

avec B , l'opérateur de retard, tel que $BX_t = X_{t-1}$ et les ε_t qui sont une suite de variables aléatoires non corrélées, de moyenne 0 et de variance σ_ε^2 . Sauf exception, les modèles ARIMA sont moins aisés à interpréter.

Plusieurs auteurs ont établi les liens qui existent entre les méthodes de prévision considérées et les modèles ARIMA. Les résultats synthétisés en annexe sont dus à Muth [31] pour le lissage exponentiel simple, à Cogger [11], Godolphin et Harrison [12], Goodman [13] et McKenzie [23] pour les méthodes de lissage exponentiel de degré supérieur, à Ward [35], Smith [32] et McKenzie [24] pour la méthode de Holt-Winters, dans sa version additive, à Harrison [15] et Van Winkel [34] pour les modèles polynomiaux, à Van Winkel [34] pour le modèle de Theil et Wage et Libert [17] pour le modèle général à deux aléas.

Par exemple plusieurs méthodes correspondent au modèle ARIMA (0, 2, 2):

$$\nabla^2 X_t = (1 - \theta_1 B - \theta_2 B^2) \varepsilon_t.$$

La mise sous forme ARIMA permet notamment de faire apparaître les conditions d'inversibilité qui sont nécessaires à l'utilisation de ces méthodes. Les graphiques des figures 1 et 2 résument les contraintes imposées aux paramètres θ_1 et θ_2 des modèles ARIMA (0, 2, 2) sous-jacents aux algorithmes de lissage exponentiel du second degré, de lissage du second ordre de Brown [8], de lissage exponentiel avec tendance de Holt, au modèle de Theil et Wage et au modèle de croissance linéaire.

En fait, les processus sous-jacents aux trois premières classes de méthodes sont de type ARIMAG (Mélard [28]). Les processus ARIMAG sont définis comme les processus ARIMA, sauf que les innovations n'ont pas nécessairement une variance constante.

Dans la suite de cet article, nous envisagerons des extensions des méthodes comportant les caractéristiques suivantes :

1° L'inclusion d'un mouvement saisonnier déterministe (Abraham et Box [2]) pour les méthodes ne prenant pas en compte le mouvement saisonnier;

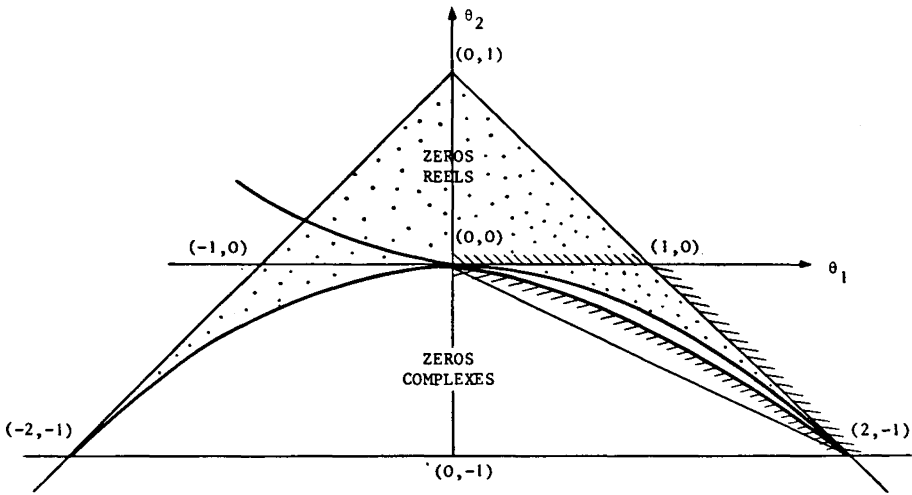


Figure 1. — Les modèles ARIMA (0, 2, 2).

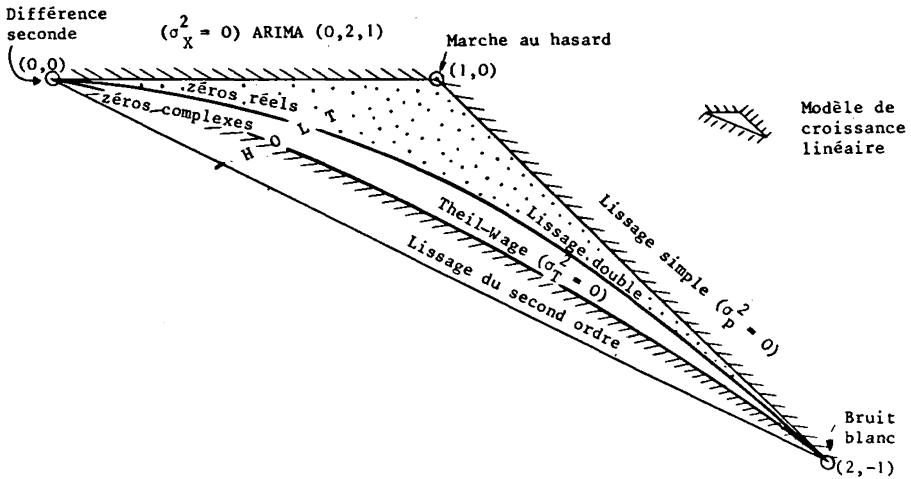


Figure 2. — Extrait de la figure 1.

2° la possibilité de faire varier la variance des erreurs de prévision, y compris dans les classes (4) et (5) ci-dessus;

3° une transformation de la variable, soit choisie *a priori*, soit dépendant d'un paramètre qui est estimé simultanément avec les autres paramètres (Box et Cox [5]);

4° un sous-modèle d'intervention qui représente, de façon paramétrée, l'effet transitoire ou permanent d'une intervention extérieure (Box et Tiao [6]).

Il est utile d'illustrer les liens avec les processus ARIMA et la prise en compte de ces extensions sur l'exemple du lissage exponentiel simple.

Le lissage exponentiel simple est défini par la relation de récurrence :

$$\tilde{X}_t = \alpha X_{t-1} + (1-\alpha) \tilde{X}_{t-1} \quad (1)$$

où α est la constante de lissage, X_t est l'observation au temps t et \tilde{X}_t , la valeur lissée est la prévision faite à l'instant $t-1$ pour l'instant t . Notons que le problème du choix d'une valeur initiale pour \tilde{X}_t sera traité au paragraphe 3.

Si les erreurs de prévision $\varepsilon_t = X_t - \tilde{X}_t$ sont des variables aléatoires de moyenne nulle et non autocorrélées entre elles, on peut écrire d'une part :

$$X_t = \varepsilon_t + \alpha X_{t-1} + (1-\alpha) \tilde{X}_{t-1}$$

et, d'autre part :

$$(1-\alpha) X_{t-1} = (1-\alpha) \varepsilon_{t-1} + (1-\alpha) \tilde{X}_{t-1}.$$

Par soustraction membre à membre il vient :

$$X_t - (1-\alpha) X_{t-1} = \varepsilon_t + \alpha X_{t-1} - (1-\alpha) \varepsilon_{t-1}$$

et donc

$$\nabla X_t = (1 - (1-\alpha)B) \varepsilon_t. \quad (2)$$

Rien n'impose que la variance g_t^2 des ε_t soit constante. Si l'on préfère normaliser les ε_t pour que leur variance soit constante, (2) peut s'écrire sous la forme :

$$\nabla X_t = (1 - (1-\alpha)B) (g_t \varepsilon_t). \quad (3)$$

Cette équation correspond à un cas particulier de processus ARIMAG. La forme de la fonction g_t peut être choisie et dépendre de paramètres inconnus. Les autres méthodes font l'objet d'un traitement similaire par Mélard [28].

On peut évidemment appliquer une méthode comme le lissage exponentiel sur les logarithmes des données ou, plus généralement, sur X_t^λ . En appliquant la méthode de Winters en version additive sur $\log X_t$, on retrouve la méthode de Winters dans une version multiplicative qui ne coïncide toutefois pas avec la méthode (hybride, car utilisant une moyenne arithmétique pondérée de rapports) connue sous ce nom.

Pour les méthodes qui ne sont pas adaptées à l'existence de variations saisonnières, il est souvent utile (Makridakis et Hibon [20]) de désaisonnaliser les données. Cela peut se faire en remplaçant X_t par $X_t - m_t$, où m_t est l'estimation du saisonnier en t . Les équations (1) et (3) deviennent alors respectivement :

$$\begin{aligned}\tilde{X}_t &= \alpha(X_{t-1} - m_{t-1}) + (1 - \alpha)\tilde{X}_{t-1} \\ \nabla X_t - \nabla m_t &= (1 - (1 - \alpha)B)(g_t \varepsilon_t).\end{aligned}$$

Evidemment \tilde{X}_t représente ici la prévision désaisonnalisée. La prévision proprement dite vaut $\tilde{X}_t + m_t$.

3. DESCRIPTION DE L'ALGORITHME D'ESTIMATION

Toutes les méthodes de prévision dépendent de paramètres inconnus. Trois attitudes sont possibles :

(A) Fixer la valeur des paramètres.

Dans le cas où la prévision de la demande doit être effectuée pour des milliers d'articles, c'est souvent la seule attitude raisonnable. Pour le lissage exponentiel simple, par exemple, ceci permet de mémoriser la seule valeur lissée. De plus le calcul de \tilde{X}_t par la formule $\tilde{X}_{t-1} + \alpha \varepsilon_{t-1}$ nécessite une multiplication, une addition et une soustraction.

(B) Faire varier les valeurs des paramètres aux sommets d'une grille rectangulaire et sélectionner le sommet optimal, au sens d'un critère à déterminer. C'est ainsi qu'a été formulée la méthode de Winters [37]. Même pour le lissage exponentiel simple, l'entièreté de l'historique de la série est nécessaire si le caractère optimal des valeurs des paramètres est revu périodiquement. De plus l'ensemble des calculs est considérablement plus lourd qu'au (A).

(C) Faire varier les valeurs des paramètres dans un sous-espace de l'espace des paramètres, sous le contrôle d'un algorithme d'optimisation, au sens d'un critère à déterminer.

L'application de cette procédure dans les méthodes de prévision est essentiellement due à Box et Jenkins [7].

Dans cet article, nous adopterons l'attitude (C) qui présente l'avantage suivant par rapport à (B) : sauf dans le cas d'un paramètre, elle revient souvent à examiner moins de points et est donc plus efficace. Cet avantage est encore renforcé lorsque les extensions mentionnées au paragraphe 2 sont employées. Ceci compense la complexité accrue de la méthode.

Les attitudes (B) et (C) ci-dessus impliquent le choix d'un critère parmi les suivants :

(1) Critère des moindres carrés conditionnels : minimisation de la somme des carrés des erreurs.

C'est l'approche utilisée le plus souvent (par exemple Winters [37]). L'inconvénient essentiel réside dans le choix de valeurs initiales permettant d'amorcer les récurrences.

(2) Critère des moindres carrés non conditionnels.

Cette procédure réduit, sans l'éliminer complètement, l'incidence du choix de valeur initiales. Elle emploie la méthode de prévision rétrospective (Box et Jenkins [7]). Elle a été employées par Makridakis et Hibon [20] notamment.

(3) Critère du maximum de vraisemblance.

Cette procédure est basée sur l'évaluation, sous l'hypothèse de normalité, de la fonction de vraisemblance exacte d'un processus ARMA. Il existe à cette fin des algorithmes très performants, dont celui de Mélard [27] qui s'avère plus rapide que celui du (2).

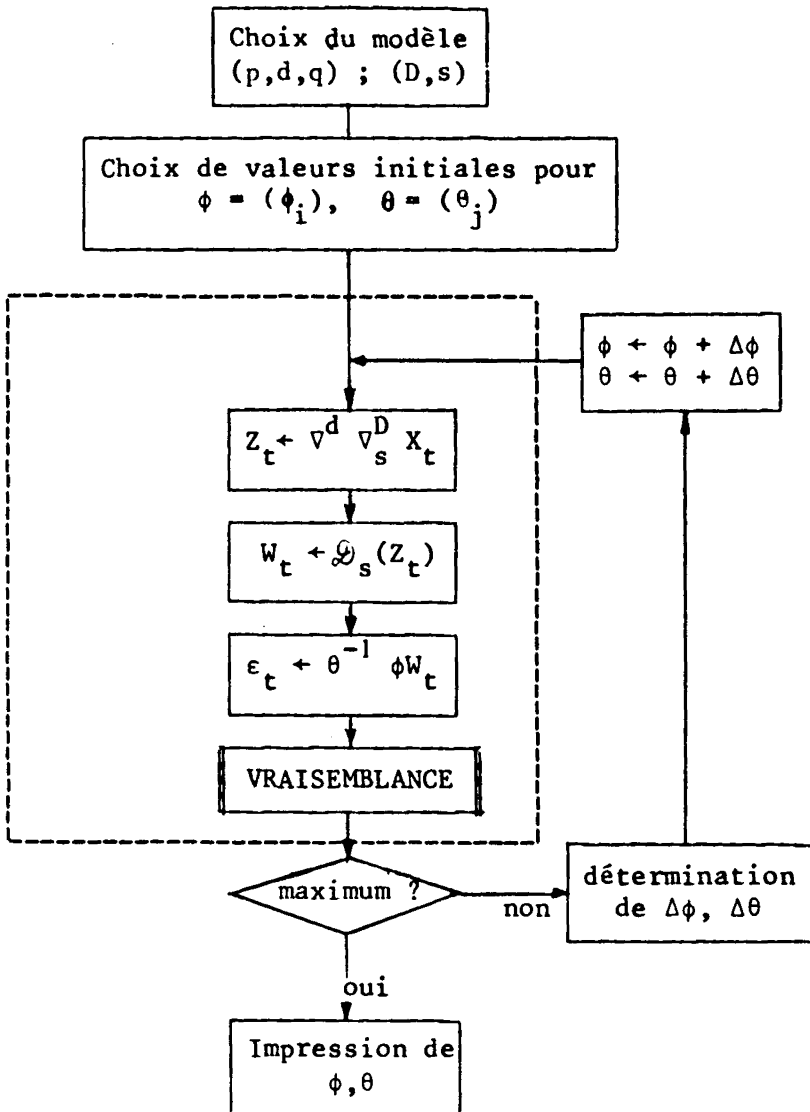
Il faut noter l'existence de procédures *ad hoc* pour la plupart des méthodes examinées (Kaufman et Groboillot [16]).

Dans le présent article nous préconisons, pour les méthodes de prévision retenues, l'estimation des paramètres par maximum de vraisemblance. Il est évident que la détermination du modèle ARIMA sous-jacent à une quelconque des méthodes de prévision joue un rôle crucial. Cette approche a pour avantage d'éliminer entièrement la nécessité de valeurs initiales. Notons qu'une approche différente mais visant le même objectif a été développée par Abadie et Travers [1]. Par rapport à l'application de l'algorithme du maximum de vraisemblance pour les modèles ARIMA la différence essentielle est que les paramètres ne sont pas les coefficients des polynômes $\phi(B)$ et $\theta(B)$ mais le(s) paramètre(s) naturel(s) associé(s) à la méthode.

Nous présentons maintenant le schéma de la procédure pour l'estimation des paramètres $\alpha = (\alpha_1, \dots, \alpha_m)$ des méthodes (fig. 4) en parallèle avec la procédure de Box et Jenkins (fig. 3). Le critère utilisé est celui du maximum de vraisemblance sous le contrôle d'un algorithme d'optimisation décrit sommairement.

On en déduit qu'il est aisé de modifier un programme implémentant la méthode de Box et Jenkins. En effet, pour au moins deux raisons (les polynômes ϕ et θ s'obtiennent souvent comme produits de polynômes; possibilité d'omettre des termes des polynômes ϕ et θ) il n'y a pas coïncidence entre les paramètres d'une part et les coefficients des polynômes d'autre part.

Fig. 3 - Box et Jenkins (généralisé)



Notations : s = période saisonnière
 D = ordre de différence saisonnière
 \mathcal{L}_s = désaisonnalisation si $D = 0$, identité si $D \neq 0$.

Figure 3. - Box et Jenkins (généralisé).

Figure 4 - Algorithmes de prévision

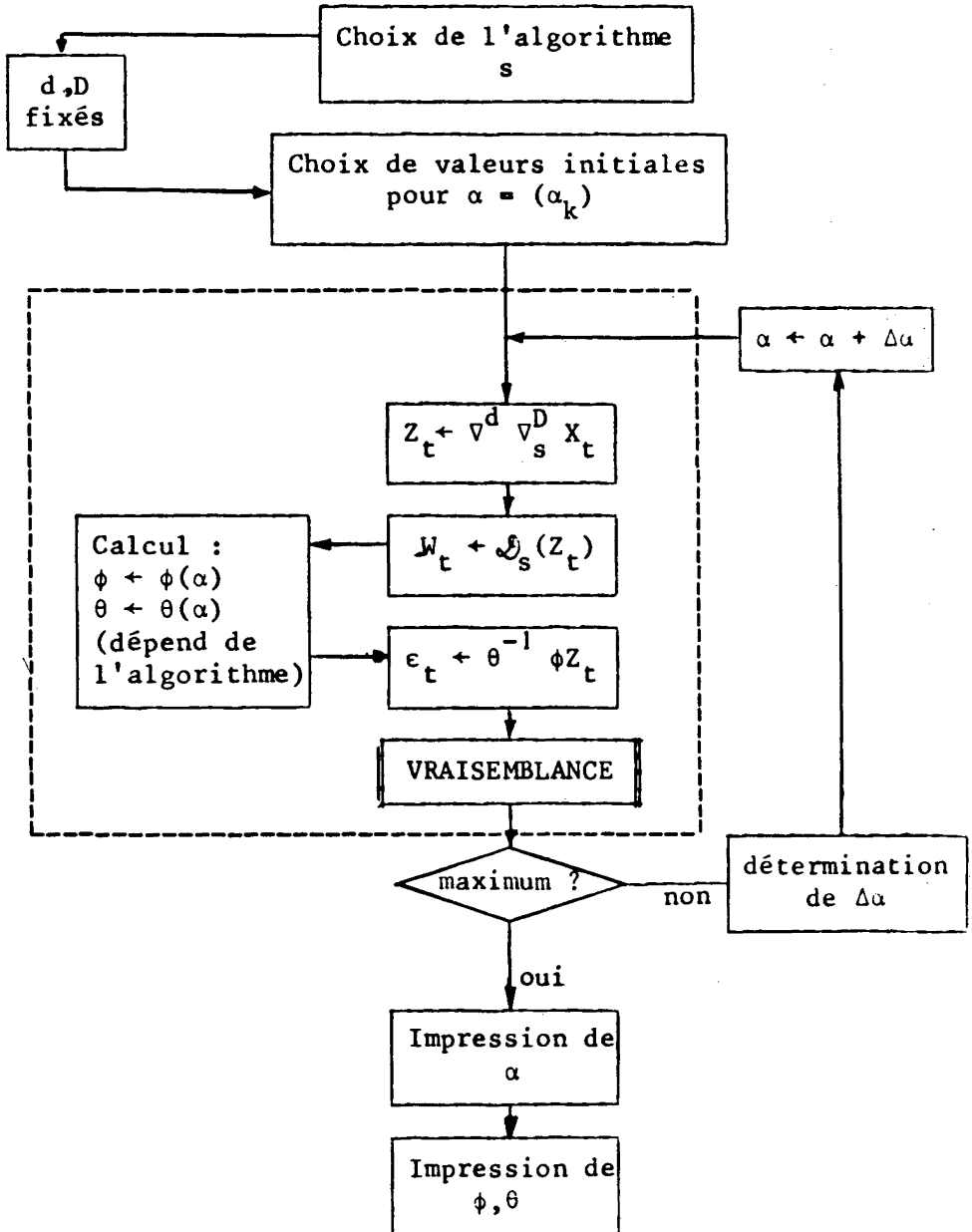


Figure 4. - Algorithmes de prévision.

Il suffit donc d'insérer, pour chaque méthode de prévision, les instructions de calcul des coefficients des polynômes φ et θ en fonction des paramètres α propres à la méthode. Par exemple, pour le lissage exponentiel simple : $\theta_1 = 1 - \alpha$.

Pour le présent article, le logiciel utilisé est ANSECH (Mélard [26]). Il permettait déjà, dans le cadre des modèles ARIMA, les quatre extensions (1°-4°) citées au paragraphe 2, ainsi que les trois attitudes (A-C) et les trois critères (1-3) ci-dessus. Il a été modifié pour inclure les modèles sous-jacents aux méthodes de prévision présentées en annexe, dans la paramétrisation donnée.

4. DESCRIPTION DE LA PROCÉDURE DE SÉLECTION

La sélection d'un modèle est toujours une tâche délicate en statistique. A part dans des cas particuliers (régression polynomiale et régression harmonique, notamment) où une procédure séquentielle praticable existe, on procède souvent par essai d'un nombre plus ou moins grand de modèles, élimination des modèles inadéquats et élection d'un modèle (ou de plusieurs) parmi les modèles restants.

La classe des modèles est celle des modèles sous-jacents aux méthodes de prévision données en annexe, complétée éventuellement par une transformation de variable, la possibilité de faire varier la variance des erreurs de prévision ou des interventions. L'élimination des méthodes est basée sur l'aptitude du modèle à représenter la série, ce qui se traduit essentiellement par l'absence d'autocorrélation des erreurs de prévision d'horizon 1, les résidus du modèle, et l'homogénéité de ces résidus. Plus concrètement l'analyse des résidus inclut le test de Ljung et Box [19] et les tests d'homogénéité (Mélard [25]).

L'élection d'un modèle et donc d'une méthode parmi les méthodes restantes est basé sur le critère d'information de Akaike. Cela consiste à choisir la méthode qui donne lieu à la plus petite valeur du critère d'Akaike (Akaike [3]) que nous calculons par la formule :

$$AIC = -2 \log_e L + 2(P+1) \frac{n}{n-l},$$

où L est la valeur maximale de la fonction de vraisemblance, n est la longueur de la série, l , le nombre d'observations perdues à cause des différences et P est le nombre de paramètres estimés (y compris la constante additive et les coefficients saisonniers éventuels);

Parmi les avantages liés à l'utilisation d'un programme Box-Jenkins pour l'analyse des différentes méthodes, on peut mentionner les suivants :

- la procédure d'estimation est unique et statistiquement valable; elle n'est pas conditionnée par le choix de valeurs initiales grâce à l'utilisation de la fonction de vraisemblance exacte;
- de par la procédure d'estimation choisie, la matrice de covariance asymptotique des estimateurs peut être estimée;
- les tests de validité des modèles sont identiques;
- on peut aisément passer d'une méthode à l'autre;
- si un modèle ne convient pas, on peut rechercher, sur base de la fonction d'autocorrélation et de la fonction d'autocorrélation partielle estimées sur les résidus, le modèle ARIMA « le plus proche » qui convienne le mieux;
- si on peut construire un sous-modèle simple pour les résidus issus de l'algorithme de prévision le moins mauvais, on peut garder une partie du pouvoir explicatif du modèle de départ.

5. LA DÉTERMINATION DES PRÉVISIONS

Du point de vue pratique, le fait de travailler dans un cadre plus large présente un énorme avantage. Au lieu de devoir assimiler des méthodes particulières, différentes les unes des autres (ou de se limiter à une seule d'entre-elles), on les place dans un ensemble homogène, celui des modèles ARIMA. Pour la prévision proprement dite, le gain est très important pour les classes de méthodes (4) et (5) car une fois le modèle ARIMA connu, les prévisions s'obtiennent comme une combinaison linéaire des observations passées et des erreurs de prévision passées.

Bien que pour chaque méthode de prévision il existe une formule de prévision spécifique à celle-ci, le principe de la méthode proposée ici est d'utiliser le prédicteur du modèle ARIMA correspondant (Box et Jenkins [7]). Ainsi, par exemple, pour le modèle sous-jacent à la méthode de Holt de lissage exponentiel avec tendance, la prévision d'horizon h faite en $t-1$, $\hat{X}_{t-1}(h)$ s'écrit :

$$\hat{X}_{t-1}(h) = 2\hat{X}_{t-1}(h-1) - \hat{X}_{t-1}(h-2) - \theta_1 \hat{\varepsilon}_{t+h-2} - \theta_2 \hat{\varepsilon}_{t+h-3},$$

avec

$$\begin{aligned} \hat{\varepsilon}_{t-i} &= 0 & \text{si } i \leq 0, \\ \hat{\varepsilon}_{t-i} &= X_{t-i} - \hat{X}_{t-1-i}(1) & \text{si } i \geq 1, \end{aligned}$$

θ_1, θ_2 = fonctions des constantes de lissage estimées (cf. annexe), tandis que la formule de prévision propre à cette méthode s'écrit :

$$\hat{X}_{t-1}(h) = N_t + h P_t,$$

où N_t et P_t sont respectivement le niveau et la pente au temps t , telles qu'on peut les déterminer à partir des relations de récurrence rappelées en annexe. Cette possibilité présente plusieurs avantages dont celle de comparer aisément les prédicteurs des différentes méthodes et, notamment pour les modèles à deux aléas, la simplicité de la forme analytique du prédicteur.

De plus, un sous-produit intéressant est l'obtention de la variance des erreurs de prévisions pour des horizons quelconques (en supposant $g_t = 1$, pour simplifier). En effet, si l'on écrit :

$$X_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots,$$

la variance des erreurs de prévision d'horizon h s'obtient par :

$$V(h) = \left(1 + \sum_{i=1}^{h-1} \psi_i^2 \right) \sigma^2. \quad (4)$$

Par exemple, pour le modèle ARIMA (0, 2, 2) ci-dessus, on a :

$$(1 - 2B + B^2)(1 + \psi_1 B + \psi_2 B^2 + \dots) = (1 - \theta_1 B - \theta_2 B^2)$$

et donc, en identifiant les termes en B de même puissance :

$$\psi_1 = 2 - \theta_1, \psi_2 = 3 - 2\theta_1 - \theta_2, \psi_3 = 4 - 3\theta_1 - 2\theta_2, \dots$$

On peut donc évaluer (4) de la même manière que dans la méthode de Box et Jenkins. Il est donc possible d'assortir les prévisions issues des différentes méthodes d'un intervalle de prévision à $100(1 - \alpha)\%$, ce qui n'est pas possible, semble-t-il, par l'approche directe de ces méthodes de prévision, ou n'a jamais été essayé.

6. ILLUSTRATION DE LA MÉTHODE

Le but de ce paragraphe est simplement d'exposer l'application de la procédure proposée et de préciser certains points qui n'ont été abordés ni dans le teste, ni dans l'annexe. La série traitée est celle de ventes mensuelles (en millions de bouteilles) de champagne (par exemple : Wheelwright et Makridakis [36]). Dans un premier stade, nous traitons des données brutes.

Ultérieurement, nous envisagerons les logarithmes des données, et puis l'application de la méthode de Box et Cox [5]. Étant donné l'importance du mouvement saisonnier dans cette série, toutes les méthodes ont été munies d'une fonction périodique m_t , à l'exception de la méthode de Winters qui effectue un traitement stochastique de la composante annuelle. De plus, les méthodes qui se ramènent à un modèle ARIMA ne comportant pas plus d'une différence (non saisonnière ou saisonnière) reçoivent un paramètre supplémentaire sous forme d'une constante additive.

Le lissage simple donne lieu au modèle :

$$\nabla X_t - \nabla m_t - 0,0254 = (1 - 0,932 B) \varepsilon_t,$$

où la constante de lissage vaut donc $1 - 0,932 = 0,068$. Nous ne détaillons pas les coefficients saisonniers ∇m_t . Ils ont été calculés à partir des différences premières des observations, par comparaison des moyennes de chaque mois à la moyenne générale. Le programme utilisé permet aussi d'estimer les 11 coefficients simultanément avec la constante de lissage. Il est également possible d'employer une représentation de Fourier de la composante saisonnière en sélectionnant les périodes qui apportent une contribution significative. Enfin, on peut représenter le mouvement saisonnier par une combinaison linéaire de fonction trigonométriques et de « variables muettes » saisonnières, les coefficients étant donc estimés simultanément avec la constante de lissage. L'estimation non biaisée de la variance, $\hat{\sigma}_\varepsilon^2$, vaut ici 0,873. Les autocorrélations d'ordre 1 et 12 valent 0,31 et 0,42, ce qui montre la faible performance du modèle. Le critère d'Akaike vaut $AIC = 1589$.

Le lissage double fournit une constante de lissage optimale de 0,054 et $AIC = 1598$. Le lissage triple est moins bon avec $AIC = 1614$. Si on exprime la prévision désaisonnalisée comme moyenne des 12 observations désaisonnalisées les plus récentes, on obtient $AIC = 1586$. Le résultat est meilleur mais il subsiste de l'autocorrélation d'ordre 1 et 12 dans les résidus.

Les constantes de lissage de la méthode de Winters sont $\alpha = 0,005$, $\beta = -0,890$ et $\gamma = 0,774$. Notons qu'il est nécessaire de fixer des valeurs initiales de ces paramètres telles que le modèle est inversible. L'écart-type associé à β vaut 1,500 de sorte que la constante de tendance n'est pas significativement différente de 0. Il est également intéressant de présenter la forme ARIMA du modèle :

$$\nabla \nabla_{12} X_t = [(1 - 0,9994 B + 0,0049 (B^2 + \dots + B^{11}) - 0,235 B^{12} + 0,225 B^{13}) \varepsilon_t].$$

Le modèle est apparemment très bon avec $AIC = 1561$, une seule autocorrélation partielle résiduelle significative à 5%, celui d'ordre 13 qui vaut 0,22, et

un niveau de signification supérieur à 1/2 pour le test de Ljung et Box (effectué jusqu'à l'ordre 24).

Les résultats sont moins probants pour les modèles à deux aléas et les modèles polynomiaux. Ces modèles sont sensibles au choix des valeurs initiales pour les paramètres. Il est possible qu'une part de responsabilité incombe à l'algorithme d'optimisation utilisé (Marquardt [22]). Profitons-en pour aborder un petit problème dû à l'utilisation d'optimisation sans contrainte alors que certains paramètres sont des rapports d'écart-type, donc non négatifs. Le sous-programme de calcul des coefficients du modèle ARIMA correspondant à ces algorithmes de prévision prend la valeur absolue du paramètre à estimer, ce qui évite toute difficulté sauf au voisinage de zéro. Des valeurs nulles des paramètres sont d'ailleurs peu recommandées ici parce qu'elles peuvent donner lieu à un modèle non inversible.

Pour le modèle à deux aléas général, on trouve, avec les notations présentées en annexe, $\rho_1 = 0,38$, $\rho_2 = -0,52$ et le rapport de l'écart-type de l'aléa permanent à celui de l'aléa transitoire vaut 3.10^{-6} , ces deux derniers étant d'ailleurs non significatifs à 5%. On obtient $AIC = 1581$ et les résidus possèdent un seul coefficient d'autocorrélation significatif à 5%, pour l'ordre 12. Le modèle ARIMA correspondant est défini par l'équation :

$$(1 + 0,138 B - 0,196 B^2)(\nabla X_t - m_t) = (1 - 0,484 B - 0,517 B^2) \varepsilon_t.$$

Le modèle de base de Morlat et Bachelet, cas particulier du modèle précédent où $\rho_1 = \rho_2 = 0$, est identique au lissage exponentiel simple, mais la paramétrisation est différente. De même que le modèle de Theil et Wage, il correspond mal à l'équation ci-dessus. Les échecs de ces deux modèles ne sont donc pas surprenants. Le modèle polynomial de degré 0 (de « croissance constante ») est identique au modèle de Morlat-Bachelet et au lissage exponentiel simple sauf que la constante additive est omise. La constante de lissage associée au paramètre naturel du modèle vaut 0,241.

Le modèle polynomial de degré 1 (de « croissance linéaire ») a posé quelques problèmes. Le choix de valeurs initiales pour les paramètres est très important. Nous obtenons de bons résultats en prenant : 1° 0 pour le rapport des écarts-types tendance sur bruit; 2° le paramètre obtenu lors de l'estimation du modèle de Theil et Wage pour le rapport des écarts-types pente sur bruit. On profite ainsi de ce que le modèle de Theil et Wage est un cas particulier du modèle de croissance linéaire lorsque le premier paramètre de celui-ci est nul. Le résultat de l'estimation est un modèle IMA (2, 2) :

$$\nabla^2 X_t - m_t = (1 - 1,751 B + 0,756 B^2) \varepsilon_t.$$

Ceci est en accord avec le lissage exponentiel avec tendance de Holt pour lequel les constantes de lissage obtenues sont de 0,127 et 0,002. Le modèle IMA (2,2) correspondant est :

$$\nabla^2 X_t - m_t = (1 - 1,873 B + 0,874 B^2) \varepsilon_t$$

Ces modèles sont tous rejetés avec AIC voisin de 1 600.

Finalement, le moins mauvais est le lissage exponentiel avec tendance et saisonnier (modèle de Winters) pour lequel la variance résiduelle, corrigée pour le biais, vaut 0,558. Montgomery et Contreras [29] ont traité la même série en utilisant la méthode de Winters de façon traditionnelle et dans la version où la composante saisonnière est multiplicative. Pour nous rapprocher de cette étude, nous avons donc repris toutes les méthodes sur les logarithmes des données. Les résultats sont meilleurs car même pour le lissage simple on atteint AIC=1 528 et il ne subsiste plus d'autocorrélation marquée dans les résidus. La méthode pour laquelle la variance résiduelle est la plus petite est celle de la prévision par moyenne mobile (d'ordre 12) avec une valeur de 0,402 (après normalisation selon Box et Cox [5]).

Le tableau I contient les principaux résultats relatifs à l'estimation des paramètres et à la prévision. Les paramètres ont été estimés sur base des 96 premières observations. On donne la valeur de AIC et l'estimation non biaisée $\hat{\sigma}_e^2$ de la variance des erreurs de prévision d'horizon 1. La colonne LB contient la valeur de la statistique de Ljung-Box suivie, entre parenthèses, du nombre de degrés de liberté. La colonne H contient le nombre de tests d'homogénéité parmi les quatre tests essayés qui conduisent au rejet de l'algorithme, au niveau de probabilité de 5%. (Les résultats de ces tests n'ont aucun sens si les résidus sont autocorrélés). Les prévisions sont calculées à partir des 96 observations et comparées à 9 données supplémentaires. Le tableau I présente le carré moyen des erreurs (MSE) et la moyenne des pourcentages d'erreur, en valeur absolue (MAPE). On constate que la transformation logarithmique fournit des résultats supérieurs à tous points de vue, sauf pour les tests d'homogénéité.

Appliquons maintenant la méthode de sélection proposée au paragraphe 4, d'abord sur les données non transformées puis sur les logarithmes des données. Pour les données non transformées, la méthode avec le plus petit AIC est celle de Winters (3.2) pour laquelle les tests sur les résidus sont satisfaisants (c'est d'ailleurs le seul). On remarque aussi que c'est le modèle pour lequel les prévisions sont les meilleures. Pour les logarithmes des données, plusieurs méthodes fournissent des résultats très voisins, à une exception près. Le test de Ljung-Box conduit à l'acceptation de chacune de ces méthodes. En revan-

TABLEAU I

Comparaison des méthodes de prévision à court termes (identifiés d'après le code utilisé en annexe) appliqués aux données brutes, aux logarithmes des données, et aux données après transformation de Box et Cox et interventions.

Transformation	Méthode	Estimation				Prévision	
		AIC	$\hat{\sigma}_e^2$	LB	H	MSE	MAPE
Néant	1.1	1 589	0,762	54,1 (23)	4	0,505	21,0
	1.2	1 598	0,784	59,1 (23)	2	0,516	21,4
	1.3	1 614	0,840	61,7 (23)	2	0,833	27,3
	2 (k=12)	1 586	0,746	51,8 (24)	4	0,491	20,8
	3.1	1 599	0,781	58,9 (22)	3	0,673	24,3
	3.2	1 561	0,558	18,6 (21)	0	0,122	9,7
	4.1	1 581	0,674	27,3 (21)	4	0,742	25,0
	4.2	1 589	0,762	54,2 (23)	4	0,504	21,0
	4.3	1 599	0,789	58,8 (23)	2	0,414	19,2
	5.1	1 590	0,791	50,5 (23)	2	0,524	21,1
5.2	1 600	0,798	53,6 (22)	2	1,015	30,0	
Logarithmique (avec normalisation)	1.1	1 528	0,406	14,1 (23)	4	0,100	9,0
	1.2	1 536	0,420	16,1 (23)	4	0,093	8,7
	1.3	1 552	0,443	17,5 (23)	4	0,116	9,6
	2 (k=12)	1 526	0,402	11,5 (24)	4	0,103	9,2
	3.1	1 537	0,413	15,9 (22)	4	0,105	9,2
	3.2	1 537	0,406	12,5 (21)	1	0,104	8,5
	4.1	1 532	0,416	13,0 (23)	4	0,099	9,0
	4.2	1 528	0,406	14,1 (23)	4	0,100	9,0
	4.3	1 537	0,427	16,7 (23)	4	0,104	9,3
	5.1	1 529	0,415	13,6 (23)	4	0,099	8,0
5.2	1 537	0,417	16,3 (22)	4	0,107	9,3	
Box-Cox (avec normalisation) et interventions	1.1	1 451	0,176	16,1 (23)	4	0,135	9,5
	1.2	1 462	0,189	21,2 (23)	2	0,141	9,5
	1.3	1 477	0,201	22,5 (23)	0	0,196	11,1
	2 (k=12)	1 458	0,188	26,2 (24)	4	0,146	9,4
	3.1	1 459	0,179	14,4 (22)	4	0,135	9,4
	3.2	1 466	0,206	11,4 (21)	1	0,122	9,3
	4.1	1 453	0,177	12,9 (23)	4	0,141	9,8
	4.2	1 451	0,176	16,0 (23)	4	0,135	9,5
	4.3	1 466	0,195	24,9 (23)	1	0,166	9,9
	5.1	1 451	0,176	14,1 (23)	0	0,124	8,7
5.2	1 460	0,180	14,3 (23)	4	0,133	9,4	

che les tests d'homogénéité suggèrent au contraire de rejeter toutes les méthodes sauf celle de Winters. On serait donc tenté de choisir celle-ci par application de la méthode du paragraphe 4. Notons qu'ici les performances des prévisions sont similaires pour toutes les méthodes. Un autre choix ne serait donc pas critique. Cette concordance massive des tests d'homogénéité à rejeter les modèles s'explique par la présence de valeurs aberrantes. La vente de champagne du mois d'août de la première année n'est en effet pas aussi basse que durant les années suivantes. Des anomalies sont également observées en janvier 1965, décembre 1967, janvier 1968 et mai 1968. On peut s'interroger à ce propos sur la validité du choix de la transformation logarithmique.

Pour répondre à ces différentes critiques nous avons tenté une transformation de variable paramétrée combinée à une analyse d'intervention. Au lieu de X_t on emploie :

$$\tilde{X}_t = \begin{cases} (X_t^\lambda - 1)/\lambda, & \lambda \neq 0, \\ \log X_t, & \lambda = 0. \end{cases}$$

En fait la transformation est normalisée, en multipliant \tilde{X}_t par $G^{1-\lambda}$, où G est la moyenne géométrique des données. De cette façon, les différents $\hat{\sigma}_e^2$ et AIC sont comparables entre eux.

L'analyse d'intervention consiste à remplacer X_t par $X_t - Y_t$, où $Y_t = 0$ pour tout t sauf aux cinq mois atypiques signalés ci-dessus pour lesquels un paramètre est chaque fois utilisé. Cette approche est similaire à celle des variables muettes employées en régression multiple, sauf qu'ici le modèle est non linéaire. Les principaux résultats sont repris dans la troisième partie du tableau I. Globalement l'ajustement est évidemment meilleur. Notons que la valeur estimée de λ oscille entre $-0,20$ et $-0,04$. Les cinq paramètres introduits dans l'analyse d'intervention se révèlent significativement non nuls. On constate toutefois que les prévisions ne sont pas améliorées par rapport à la transformation logarithmique. L'explication est simple : les observations anormales se présentent assez loin de la fin de la série.

Enfin, comme alternative à une transformation, on a essayé de rendre la variance des innovations dépendantes du temps. Puisque les fluctuations semblent s'amplifier on a choisi g_t de forme exponentielle du temps. Les résultats sont aussi mauvais que sans transformation. Ils ne sont pas donnés ici.

BIBLIOGRAPHIE

1. J. ABADIE et D. TRAVERS, *Une approche simplifiée de la méthode de Box et Jenkins pour l'analyse et la prévision des séries temporelles unidimensionnelles*, (I) R.A.I.R.O., Rech. Op., vol. 14, 1980, p. 355-380; (II) R.A.I.R.O., Rech. Op., vol. 15, 1981, p. 51-71.
2. B. ABRAHAM et G. E. P. BOX, *Deterministic and Forecast-Adaptative Time-Dependent Models*, J. Roy. Statist. Soc. Ser. C. Appl. Statist., vol. 27, n° 2, 1981, p. 120-130.
3. H. AKAIKE, *On the Likelihood of a Time Series Model*, The Statistician, vol. 27, 1978, p. 217-235.
4. D. BACHELET et G. MORLAT, *Modèles à deux aléas pour des chroniques économiques*, Rev. Fr. Rech. Op., vol. 40, 1966, p. 275-289.
5. G. E. P. BOX et D. R. COX, *An Analysis of Transformations*, J. Roy. Statist. Soc., Ser. B, vol. 26, 1964, p. 211-243.
6. G. E. P. BOX et G. C. TIAO, *Intervention Analysis with Applications to Economic and Environmental Problems*, J. Amer. Statist. Assoc., vol. 70, 1975, p. 70-79.

7. G. E. P. BOX et G. M. JENKINS, *Time Series Analysis: Forecasting and Control*, Holden Day, San Francisco (Revised edition), 1976.
8. R. G. BROWN, *Statistical Forecasting for Inventory Control*, McGraw-Hill, New York, 1959.
9. R. G. BROWN, *Smoothing, Forecasting and Prediction of Discrete Time Series*, Prentice Hall, Englewood Cliffs N.J., 1962.
10. C. CHATFIELD, *The Holt-Winters Forecasting Procedure*, J. Roy. Statist. Soc., Ser. C. Appl. Statist., vol. 27, n° 3, 1978, p. 264-279.
11. K. O. COGGER, *The Optimality of General-Order Exponential Smoothing*, Op. Res., vol. 22, n° 4, 1974, p. 858-867.
12. E. J. GODOLPHIN et P. J. HARRISON, *Equivalence Theorems for Polynomial-Projecting Predictors*, J. Roy. Statist. Soc., Ser. B., vol. 37, n° 2, 1975, p. 205-215.
13. M. L. GOODMAN, *A New Look at Higher-Order Exponential Smoothing for Forecasting*, Op. Res., vol. 22, n° 4, 1974, p. 880-888.
14. D. GROSS et J. L. RAY, *A General Purpose Forecast Simulator*, Managmnt Sc., vol. 11, n° 6, 1965, p. 119-135.
15. P. J. HARRISON, *Exponential Smoothing and Short-Term Sales Forecasting*, Managmnt Sc., vol. 13, 1967, p. 821-842.
16. H. KAUFMAN et J.-L. GROBILLOT, *Les techniques de la prévision à court terme : lissage, modèles, simulation*, Dunod, Paris, 1968.
17. G. LIBERT, *Comparaison et interprétation de deux méthodes de traitement de séries chronologiques*, 10th European Meeting of Statisticians, 22-26 août, 1977, Louvain, Belgium.
18. G. LIBERT, *Construction automatique et continue de prévision*, R.A.I.R.O., Rech. Op., vol. 16, n° 4, 1982, p. 333-348.
19. G. M. LJUNG et G. E. P. BOX, *On a Measure of Lack of Fit in Time Series Models*, Biometrika, vol. 65, n° 2, 1978, p. 297-303.
20. S. MAKRIDAKIS et M. HIBON, *Accuracy of Forecasting: an Empirical Investigation*, J. Roy. Statist. Soc., Ser. A, vol. 142, 1979, p. 97-145 (with discussion).
21. S. MAKRIDAKIS et S. C. WHEELWRIGHT, *Interactive Forecasting: Univariate and Multivariable Methods*, Holden-Day, San Francisco, 1978.
22. D. W. MARQUARDT, *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*, J. Soc. Industr. Appl. Math., vol. 11, 1963, p. 431-441.
23. E. MCKENZIE, *A Comparison of Standard Forecasting Systems with the Box-Jenkins Approach*, The Statistician, vol. 23, n° 2, 1974, p. 107-116.
24. E. MCKENZIE, *A Comparison of Some Standard Seasonal Forecasting Systems*, The Statistician, vol. 25, n° 1, 1976, p. 3-14.
25. G. MÉLARD, *Modèles ARIMA pour des séries chronologiques non homogènes*, Statistique et Analyse des Données, vol. 2, 1979, p. 41-50.
26. G. MÉLARD, *Software for Time Series Analysis*, In H. CAUSSINUS, P. ETINGER et R. TOMASSONE eds., Compstat, 1982, Part I, Proceedings in Computational Statistics, Physica-Verlag, Vienna, 1982, p. 336-341.
27. G. MÉLARD, *A Fast Algorithm for the Exact Likelihood of Autoregressive-Moving Average Models*, J. Roy. Statist. Soc. C Appl. Statist., vol. 33, n° 1, 1984, p. 104-114.
28. G. MÉLARD, *Analyse de données chronologiques*, Coll. Séminaire de mathématiques supérieures, Presses de l'Université de Montréal, Montréal.
29. D. C. MONTGOMERY et L. E. CONTRERAS, *A Note on Forecasting with Adaptive Filtering*, Op. Res. Q., vol. 28, n° 1, 1977, p. 87-91.

30. D. C. MONTGOMERY et L. A. JOHNSON, *Forecasting and Time Series Analysis*, McGraw Hill, New York, 1975.
31. J. F. MUTH, *Optimal Properties of Exponentially Weighted Forecasts*, J. Amer. Statist. Ass., vol. 55, 1960, p. 299-306.
32. T. M. F. SMITH, *A Comparison of Some Models for Predicting Time Series Subject to Seasonal Variation*, The Statistician, vol. 17, 1967, p. 301-305.
33. H. THEIL et S. WAGE, *Some Observations on Adaptive Forecasting*, Managmnt Sc., vol. 10, 1964, p. 198-206.
34. E. G. F. VAN WINKEL, *Some More Observations on Adaptative Forecasting*, Preprints of the Second European Congress on Operations Research, M. ROUBENS ed., North-Holland, Amsterdam, 1976, p. 591-596.
35. D. H. WARD, *Comparison of Different Systems of Exponentially Weighted Prediction*, The Statistician, vol. 13, 1963, p. 173-185.
36. S. C. WHEELWRIGHT et S. MAKRIDAKIS, *Forecasting Methods for Management*, 2nd ed., Wiley, New York, 1977.
37. P. R. WINTERS, *Forecasting Sales by Exponentially Weighted Moving Averages*, Managmnt Sc., vol. 6, 1960, p. 324-342.

ANNEXE

LES MÉTHODES DE PRÉVISION

1.1. Lissage exponentiel du premier degré (lissage simple)

Relation de récurrence :

$$\tilde{X}_t = \alpha(X_{t-1} - m_{t-1}) + (1 - \alpha)\tilde{X}_{t-1} \quad (0 < \alpha < 2);$$

formule de prévision : $\hat{X}_{t-1}(h) = \tilde{X}_t + m_t$;

modèle équivalent : $\nabla X_t - \nabla m_t = [1 - (1 - \alpha)B] [g_t \varepsilon_t]$;

paramètre : $1 - \alpha$.

1.2. Lissage exponentiel du second degré (lissage double)

Relations de récurrence :

$$\begin{cases} \tilde{X}_t^{(1)} = \alpha(X_{t-1} - m_{t-1}) + (1 - \alpha)\tilde{X}_{t-1}^{(1)} & (0 < \alpha < 2), \\ \tilde{X}_t^{(2)} = \alpha\tilde{X}_{t-1}^{(1)} + (1 - \alpha)\tilde{X}_{t-1}^{(2)}; \end{cases}$$

formule de prévision : $\hat{X}_{t-1}(1) = 2\tilde{X}_t^{(1)} - \tilde{X}_t^{(2)} + m_t$;

modèle équivalent : $\nabla^2 X_t - \nabla^2 m_t = [1 - 2(1 - \alpha)B + (1 - \alpha)^2 B^2] [g_t \varepsilon_t]$;

paramètre : $1 - \alpha$.

1.3. Lissage exponentiel du troisième degré (lissage triple)

Relations de récurrence :

$$\tilde{X}_t^{(1)} = \alpha (X_{t-1} - m_{t-1}) + (1 - \alpha) \tilde{X}_{t-1}^{(1)} \quad (0 < \alpha < 2),$$

$$\tilde{X}_t^{(2)} = \alpha \tilde{X}_{t-1}^{(1)} + (1 - \alpha) \tilde{X}_{t-1}^{(2)},$$

$$\tilde{X}_t^{(3)} = \alpha \tilde{X}_{t-1}^{(2)} + (1 - \alpha) \tilde{X}_{t-1}^{(3)};$$

formule de prévision : $\hat{X}_{t-1}(1) = 3 \tilde{X}_t^{(1)} - 3 \tilde{X}_t^{(2)} + \tilde{X}_t^{(3)} + m_t$;

modèle équivalent :

$$\nabla^3 X_t - \nabla^3 m_t = [1 - 3(a - \alpha)B + 3(1 - \alpha)^2 B^2 - (1 - \alpha)^3 B^3] [g_t, \varepsilon_t];$$

paramètre : $1 - \alpha$.

1.4. Lissage exponentiel du second ordre de Brown

Relation de récurrence :

$$\tilde{X}_t = \alpha (X_{t-1} - m_{t-1}) + 2(1 - \alpha) \tilde{X}_{t-1} - (1 - \alpha) \tilde{X}_{t-2} \quad (0 < \alpha < 2);$$

formule de prévision : $\tilde{X}_{t-1}(1) = 2 \tilde{X}_t - \tilde{X}_{t-1} + m_t$;

modèle équivalent : $\nabla^2 X_t - \nabla^2 m_t = [1 - 2(1 - \alpha)B + (1 - \alpha)B^2] [g_t, \varepsilon_t]$;

paramètre : $1 - \alpha$.

2. MOYENNE MOBILE SIMPLE

Formule de prévision :

$$\hat{X}_{t-1}(1) = \frac{1}{k} \sum_{i=1}^k (X_{t-i} - m_{t-i}) + m_t \quad (1 \leq k \text{ entier});$$

modèle équivalent :

$$\left(1 - \sum_{i=1}^{k-1} \phi_i B^i\right) (\nabla X_t - \nabla m_t) = [g_t, \varepsilon_t], \quad \text{où } \phi_i = -(k-i)/k;$$

paramètre : k .

3.1. Lissage exponentiel avec tendance (Holt)

Relations de récurrence :

$$\begin{cases} N_t = \alpha(X_{t-1} - m_{t-1}) + (1-\alpha)(N_{t-1} + P_{t-1}) & \text{(niveau),} \\ P_t = \beta(N_t - N_{t-1}) + (1-\beta)P_{t-1} & \text{(pente),} \end{cases}$$

(+ conditions d'inversibilité);

formule de prévision : $\hat{X}_{t-1}(h) = N_t + hP_t + m_{t+h-1}$;

modèle équivalent : $\nabla^2 X_t - \nabla^2 m_t = [1 - (2-\alpha-\alpha\beta)B + (1-\alpha)B^2] [g_t \varepsilon_t]$;

cas particuliers :

– si $\beta=1$ on retrouve le lissage exponentiel du second ordre de Brown (1.4);

– si $\alpha=\gamma(2-\gamma)$ et $\beta=\gamma/(2-\gamma)$, on retrouve le lissage exponentiel double de constante de lissage γ ;

paramètres : $1-\alpha$ et $1-\beta$.

3.2. Lissage exponentiel avec tendance et saisonnier (Winters, modèle additif)

Relations de récurrence :

$$\begin{cases} N_t = \alpha(X_t - S_{t-s}) + (1-\alpha)(N_{t-1} + P_{t-1}) & \text{(niveau),} \\ P_t = \beta(N_t - N_{t-1}) + (1-\beta)P_{t-1} & \text{(pente),} \\ S_t = \gamma(X_t - N_t) + (1-\gamma)S_{t-s} & \text{(saisonnier)} \end{cases}$$

(+ conditions d'inversibilité);

formule de prévision :

$$\hat{X}_t(h) = N_t + hP_t + S_{(t+h)-s} \quad (1 \leq h \leq s);$$

modèle équivalent :

$$\nabla \nabla_s X_t = [1 - (1-\alpha-\alpha\beta)B + \alpha\beta \sum_{i=2}^{s-1} B^i - (1-\alpha\beta + \alpha\gamma - \gamma)B^s + (1-\alpha)(1-\gamma)B^{s+1}] [g_t \varepsilon_t];$$

cas particuliers :

– si $s=1$ et $\gamma=0$, on retrouve le lissage exponentiel avec tendance (3.1);

– si $s=1$, $\beta=1$ et $\gamma=0$, on retrouve le lissage exponentiel du second ordre (1.4);

– si $s=1$ et $\alpha=\beta=\gamma$, on retrouve le lissage exponentiel du second degré (1.2);

paramètres : $1-\alpha$, β , $1-\gamma$.

4.1. Modèle à deux aléas général

Modèle :

$$X_t = T_t + \xi_t,$$

$$T_t = T_{t-1} + m_t + \tau_t,$$

$$\text{(transitoire)} \quad \xi_t = \rho_1 \xi_{t-1} + \varepsilon_t^X; \quad \varepsilon_t^X \text{ i.i.d. } (0, \sigma_X),$$

$$\text{(permanent)} \quad \tau_t = \rho_2 \tau_{t-1} + \varepsilon_t^P; \quad \varepsilon_t^P \text{ i.i.d. } (0, \sigma_P)$$

$$(-1 < \rho_1 < 1, -1 < \rho_2 < 1);$$

modèle équivalent :

$$[1 - (\rho_1 + \rho_2)B + \rho_1 \rho_2 B^2](\nabla X_t - m_t) = (1 - \theta_1 B - \theta_2 B^2) \varepsilon_t,$$

où θ_1 et θ_2 sont fonction de ρ_1 , ρ_2 et σ_P/σ_X ;

cas particuliers :

– si $\rho_1 = \rho_2 = 0$, on retrouve le modèle de Morlat et Bachelet (4.2);

– si $\rho_1 = m = 0$ et $\rho_2 = 1$, on retrouve le modèle de Theil et Wage (4.3);

paramètres : ρ_1 ; ρ_2 et $\sigma_P/\sigma_X = M_1$;

formules (Libert [17]) :

si $\rho_2 \neq 0$:

$$s = \text{signe}(\rho_2);$$

$$u = 2(1 + \rho_2 + \rho_2^2) + M_1^2(1 + \rho_1^2);$$

$$v = (1 + \rho_2)^2 + M_1^2 \rho_1;$$

$$a = \rho_2^2;$$

$$b = \rho_2(u - 2\rho_2);$$

$$c = 2a + v^2 - 2\rho_2 u;$$

$$d = \frac{-b - s \sqrt{b^2 - 4a(c - 2a)}}{2a};$$

$$\theta_2 = \frac{d + s \sqrt{d^2 - 4}}{2};$$

$$\theta_1 = \frac{v \theta_2}{\rho_2(\theta_2 - 1)};$$

si $\rho_2 = 0$:

$$\begin{aligned} u &= M_1^2 (1 + \rho_1^2); \\ v &= 1 + M_1^2 \rho_1; \\ a &= 2 M_1 (1 - \rho_1); \\ b &= M_1^2 (1 - \rho_1^2); \\ \theta_2 &= 0; \\ \theta_1 &= \frac{u + 2 - \sqrt{b^2 + a^2}}{2v}. \end{aligned}$$

4.2. Modèle de Morlat et Bachelet

Modèle :

$$\begin{aligned} X_t &= T_t + \varepsilon_t^X, \\ T_t &= T_{t-1} + m_t + \varepsilon_t^P, \\ (\text{transitoire}) \quad \varepsilon_t^X &\text{ i.i.d. } (0, \sigma_X), \\ (\text{permanent}) \quad \varepsilon_t^P &\text{ i.i.d. } (0, \sigma_P); \end{aligned}$$

modèle équivalent : $\nabla X_t - m_t = (1 - \theta B) \varepsilon_t$, où :

$$\theta = 1 + \frac{\sigma_P^2}{2\sigma_X^2} - \frac{\sigma_P}{\sigma_X} \sqrt{1 + \frac{\sigma_P^2}{4\sigma_X^2}};$$

paramètre : σ_P/σ_X .

4.3. Modèle de Theil et Wage

Modèle :

$$\begin{aligned} X_t &= T_t + \varepsilon_t^X, \\ T_t &= T_{t-1} + P_t, \\ (\text{transitoire}) \quad \varepsilon_t^X &\text{ i.i.d. } (0, \sigma_X), \\ (\text{permanent}) \quad P_t &= P_{t-1} + m_t + \varepsilon_t^P, \quad \varepsilon_t^P \text{ i.i.d. } (0, \sigma_P); \end{aligned}$$

modèle équivalent : $(\nabla^2 X_t - m_t) = (1 - \theta_1 B - \theta_2 B^2) \varepsilon_t$, où θ_1 et θ_2 sont fonction de σ_P/σ_X ;

paramètre : $\sigma_P/\sigma_X = M_1$;

formules : (Van Winkel [34]) :

$$c = \frac{1}{4}(4 + M_1^2 + \sqrt{M_1^4 + 16 M_1^2}),$$

$$a = c + \sqrt{c^2 - 1},$$

$$\theta_2 = -\frac{1}{a},$$

$$\theta_1 = \frac{4}{a+1}.$$

5.1. Modèle de croissance constante

Identique au modèle de Morlat et Bachelet sauf qu'on note ε_t^T et σ_T (tendance) au lieu de ε_t^P et σ_P (permanent).

5.2. Modèle de croissance linéaire

Modèle :

$$X_t = T_t + \varepsilon_t^X,$$

$$T_t = T_{t-1} + P_t + \varepsilon_t^T,$$

$$P_t = P_{t-1} + m_t + \varepsilon_t^P,$$

$$\text{(bruit)} \quad \varepsilon_t^X \text{ i.i.d. } (0, \sigma_X),$$

$$\text{(tendance)} \quad \varepsilon_t^T \text{ i.i.d. } (0, \sigma_T),$$

$$\text{(pente)} \quad \varepsilon_t^P \text{ i.i.d. } (0, \sigma_P);$$

modèle équivalent : $\nabla^2 X_t - m_t = (1 - \theta_1 B - \theta_2 B^2) \varepsilon_t$, où θ_1 et θ_2 sont fonctions de σ_T/σ_X et σ_P/σ_X ;

cas particulier : si $\sigma_T = 0$, on retrouve le modèle de Theil et Wage;

paramètres : $\sigma_T/\sigma_X = M_1$ et $\sigma_P/\sigma_X = M_2$;

formules : (Van Winkel [34]) :

$$c = \frac{1}{4}(4 + 2 M_1^2 + M_2^2 + \sqrt{M_2^4 + 4 M_1^2 M_2^2 + 16 M_2^2}),$$

$$a = c + \sqrt{c^2 - 1},$$

$$\theta_2 = -\frac{1}{a},$$

$$\theta_1 = \frac{M_1^2 + 4}{a+1}.$$