

Article

« Développements récents dans la modélisation de la persistance à long terme »

C. Jimenez, W. Hipel et A. McLeod

Revue des sciences de l'eau / Journal of Water Science, vol. 3, n° 1, 1990, p. 55-81.

Pour citer cet article, utiliser l'information suivante :

URI: <http://id.erudit.org/iderudit/705065ar>

DOI: 10.7202/705065ar

Note : les règles d'écriture des références bibliographiques peuvent varier selon les différents domaines du savoir.

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter à l'URI <https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. Érudit offre des services d'édition numérique de documents scientifiques depuis 1998.

Pour communiquer avec les responsables d'Érudit : info@erudit.org

Développements récents dans la modélisation de la persistance à long terme

Developments in modelling
long term persistence *

C. JIMENEZ¹, W. HIPEL², A. McLEOD¹

RÉSUMÉ

Afin de modéliser efficacement la persistance dans les séries chronologiques rencontrées en hydrologie, des développements récents autour du modèle fractionnaire autorégressif à moyenne mobile (FARMA) (fractional autoregressive-moving average model) sont présentés. On s'intéresse particulièrement ici à de nouvelles procédures permettant d'estimer les paramètres du modèle FARMA d'une manière efficace au point de vue calcul. Pour obtenir les distributions d'échantillons des estimateurs des paramètres à partir de petits échantillons, une technique faisant appel au bootstrap peut être utilisée. Des applications pratiques à des séries de débits en rivière, de précipitations et de températures, montrent l'utilité des modèles FARMA.

Mots clés : *modélisation, persistance à long terme, modèles FARMA, hydrologie.*

SUMMARY *

In order to model effectively persistence in hydrologic time series, recent developments in fractional autoregressive-moving average (FARMA) models are presented. A time series possesses persistence or long memory if it has an autocorrelation structure that attenuates slowly to zero with increasing lags. Based on the controversy surrounding the Hurst phenomenon, some hydrologists claim that it is important to employ stochastic models which have the ability to model long memory when it is present in a given time series. Fractional Gaussian noise models and approximations thereof were developed within the field of hydrology in order to be able to model long memory. However, a particularly flexible set of models having the capability to describe long memory is the FARMA family of models, which constitutes a direct generalization of autoregressive integrated moving average (ARIMA) models.

1. Department of Statistics and Actuarial Sciences, The University of Western Ontario, London, Ontario, Canada.

2. Departments of Systems Design Engineering and Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada.

* Original English version of this paper is available. Please address your request to : *Revue des sciences de l'eau*, INRS-Eau, 2800 rue Einstein, Québec (Qué), Canada G1X 4N8.

In particular, like an ARIMA model, a FARMA model contains autoregressive and moving average parameters. Whereas the differencing operator d is restricted to be zero or take on positive integer values in an ARIMA model, the parameter d in a FARMA model can have real values and is estimated along with the other model parameters. For a specified range of values for the d parameter, a FARMA model has long memory. Besides reviewing the background and main theoretical properties of FARMA models, simulation and forecasting techniques are presented. Additionally, procedures for estimating the parameters of a FARMA model are given and a bootstrapping technique is described to obtain the small sample distributions of the estimated parameters.

To explain how to apply FARMA models in practice and demonstrate their usefulness, they are fitted to riverflow, precipitation and temperature time series.

Key-words : *modelling, long term persistence, FARMA models, hydrology.*

INTRODUCTION

Certains processus géophysiques présentent un comportement persistant. Le terme « persistant » ou « mémoire à long terme » est utilisé pour décrire une série chronologique qui possède soit une structure d'auto-corrélation qui décroît lentement jusqu'à zéro lorsque le décalage augmente, ou, de façon équivalente, une série pour laquelle la densité spectrale est fortement concentrée aux fréquences proches de zéro. Cette structure d'auto-corrélation indique que l'état actuel du processus doit être fortement dépendant des valeurs passées depuis un temps très long et que par conséquent, pour modéliser ces processus, toute l'information passée doit être intégrée dans la description du processus. Puisqu'un processus à mémoire courte tel qu'un modèle autorégressif à moyenne mobile (ARMA) ne considère qu'une portion finie du passé à travers sa structure d'autocorrélation qui a une décroissance exponentielle, plusieurs chercheurs pensent que ce type de modèle est inadéquat pour décrire des séries géophysiques persistantes qui peuvent avoir des fonctions d'autocorrélation à décroissance plus lente. Le concept de persistance a été introduit en hydrologie à la suite d'une controverse concernant le phénomène du Hurst (HURST, 1951). Plusieurs auteurs, entre autres KLEMES (1974), McLEOD et HIPEL (1978a), et SALAS *et al.* (1979), décrivent le phénomène de Hurst et suggèrent différentes approches pour modéliser la persistance des séries hydrologiques.

Plusieurs définitions de la persistance ont été données. Celle présentée par McLEOD et HIPEL (1978a) en relation avec le phénomène de Hurst, est particulièrement intéressante en raison de sa simplicité et parce qu'elle capte l'essence de la notion de « persistance ». Ces auteurs ont classé un processus de série chronologique en fonction du comportement de sa mémoire qui est définie par :

$$M = \sum_0^{\infty} |\rho_n| \quad (1)$$

où ρ_n est la valeur théorique de la FAC (fonction d'autocorrélation) du processus pour un décalage n . Un processus « mémoire à long terme » est défini comme un processus qui possède une valeur $M = \infty$, alors qu'un processus « mémoire à court terme » possède une valeur $M < \infty$. Le terme M est souvent utilisé comme coefficient global de dépendance pour les séries stationnaires (BRILLINGER, 1976) pour indiquer dans quelle mesure les valeurs présentes de la série sont indépendantes des valeurs passées. L'indépendance asymptotique entre les valeurs d'une série qui sont bien espacées dans le temps et pour lesquelles le coefficient de dépendance est donné par $M < \infty$, a été utilisée pour prouver des résultats tels que : la normalité, le comportement asymptotique d'une quantité comme la valeur empirique de la FAC, les estimations des paramètres (obtenues par la méthode du maximum de vraisemblance ou par la méthode des moments), les propriétés des tests d'hypothèses, les tests portemanteau, etc. Donc la plupart des résultats couramment utilisés dans l'analyse des séries chronologiques et présentés dans des ouvrages tels ceux de BRILLINGER (1976) ou de HANNAN (1970), ne sont pas nécessairement vrais pour les processus « mémoire à long terme » (processus ayant une mémoire infinie).

En plus de son utilisation pour des séries géophysiques, la classification « mémoire à long terme » et « mémoire à court terme » est aussi utile dans plusieurs autres domaines (COX, 1984 ; PARZEN, 1982) et particulièrement en économie (GRANGER, 1988 ; GRANGER et JOYEUX, 1980). Cette classification a aussi été utilisée pour d'autres types de processus stochastiques (COX, 1984) avec des définitions différentes pour la mémoire. Une autre définition de la « mémoire à long terme », équivalente à la précédente, consiste à considérer les processus de séries chronologiques pour lesquels la FAC décroît de la manière suivante :

$$\rho_n = O(n^{-a}) \quad (2)$$

où a se situe dans l'intervalle $(0,1)$.

Dans le passé, plusieurs modèles capables de mettre en évidence la persistance ont été proposés. Ces modèles à mémoire longue incluent les modèles « bruit Gaussien » (MANDELBROT et WALLIS, 1969a, b, c), de « ligne brisée » (MEJIA *et al.*, 1972), « niveau variable » (BOES et SALAS, 1978 ; BALLERINI et BOES, 1985), et « autorégressif et moyennes mobiles sur différences fractionnaires » (GRANGER, 1980 ; GRANGER et JOYEUX, 1980 ; HOSKING, 1981, 1984, 1985 ; LI et McLEOD, 1986). Plusieurs de ces modèles à mémoire longue sont aussi présentés dans des ouvrages d'hydrologie stochastique (KOTTEGODA, 1980 ; SALAS *et al.*, 1980 ; BRASS et RODRIGUEZ-ITURBE, 1984 ; HIPEL et McLEOD, 1990). Un avantage des modèles FARMA sur les autres modèles est que leurs structures d'autocorrélation sont capables de mettre en évidence un comportement de « mémoire à court terme » semblable à ceux des modèles ARMA mais aussi un comportement de « mémoire à long terme ». Un autre avantage des modèles FARMA est que les méthodes d'ajustement des paramètres de ces modèles peuvent tirer partie des nombreuses connaissances disponibles sur les procédures d'ajustement utilisées dans les modèles ARMA, comme le montrent BOX et JENKINS (1976).

Le but de cette étude est de présenter certains développements récents concernant l'utilisation du modèle FARMA et de montrer son intérêt pour la représentation des phénomènes hydrologiques.

MODÉLISATION DES SÉRIES CHRONOLOGIQUES PAR LE MODÈLE FARMA (AUTORÉGRESSIF ET MOYENNES MOBILES SUR DIFFÉRENCES FRACTIONNAIRES)

Définition

Pour modéliser des séries chronologiques on utilise souvent un procédé qui consiste à différencier la série si on croit que sa moyenne dépend du temps. Une moyenne qui dépendrait du temps pourrait produire des auto-corrélations empiriques qui ne décroissent pas vers zéro de façon exponentielle, comme dans le cas des modèles ARMA mais qui décroissent vers zéro beaucoup plus lentement. En pratique, lorsque la vitesse de décroissance de la FAC semble dépendre de façon linéaire du décalage, l'approche habituellement utilisée consiste à considérer les différences premières de la série. Pour le type de processus examiné dans cette étude, la FAC décroît plus lentement vers zéro qu'une fonction exponentielle mais cette vitesse de décroissance est plus élevée que celle d'une fonction linéaire. Cette constatation suggère l'utilisation d'un procédé semblable à celui utilisant l'opérateur de différentiation ordinaire, afin de modéliser une série « mémoire longue », ayant une FAC qui décroît lentement. De fait, les modèles FARMA permettent de généraliser de façon naturelle la notion de modèle ARMA différencié.

Avant de définir les modèles FARMA, il faut d'abord généraliser la notion de différenciation à l'aide du filtre :

$$\nabla^d(B) = (1 - B)^d = \sum_0^{\infty} c_n B^n \quad (3)$$

où $c_n = (-1)^n \binom{d}{n}$ et B représente l'opérateur du décalage d'un pas en arrière.

Alors on peut dire qu'un processus de série chronologique FARMA est le résultat de l'application du filtre ∇^d et du filtre ARMA ordinaire $\theta(B)/\Phi(B)$ à un processus Gaussien à bruit blanc $\{a_t\}$. Ici $\theta(B)$ et $\Phi(B)$ sont respectivement les opérateurs de moyenne mobile et d'auto-régression. Le paramètre d contrôle la mémoire du processus. Si d est non-positif, la mémoire est finie, sinon le processus possède une « mémoire à long terme ». Aussi, lorsqu'on considère le phénomène de Hurst, d est plus petit que le coefficient du Hurst de 0,5.

La notation FARMA (p,d,q) est utilisée pour représenter un modèle FARMA non saisonnier. Dans ce cas p et q sont respectivement les ordres des opérateurs d'auto-régression et de moyenne mobile alors que d est un paramètre du filtre de l'équation 3. Ce dernier peut prendre des valeurs réelles. Lorsque d

est un entier positif, le modèle FARMA (p,d,q) est équivalent à un modèle ARIMA (p,d,q) où l'abréviation ARIMA est utilisée pour désigner un processus « autoregressive integrated moving average ». Si $d = 0$, le modèle FARMA (p,d,q) est identique à un modèle (p,q) ayant une « mémoire à court terme ». Des modèles FARMA saisonniers peuvent être convenablement définis de la même façon que sont définis les modèles ARMA et ARIMA saisonniers.

Propriétés

Les propriétés fondamentales des processus FARMA ont été présentées par HOSKING (1981) et par GRANGER et JOYEUX (1980). Ces auteurs ont entre autres montré que :

- a) pour que le processus soit stationnaire, il faut que $d < 0,5$
- b) pour que le processus soit inversible il faut que $d > -0,5$
- c) la FAC se comporte comme :

$$\rho_n = O(n^{-1+2d}) \quad (4)$$

d) ce processus est semblable à lui-même ce qui signifie que les propriétés stochastiques du processus sont invariantes lors de changements d'échelle.

Plusieurs probabilistes (ROSENBLATT 1961, 1979, 1981 ; TAQQU 1975) ont étudié le comportement de statistiques obtenues à partir de ces processus pour lesquels la FAC a un comportement décrit par l'équation 4 et pour lesquels d est positif. Ils ont montré que :

(e) la multiplication de la moyenne empirique par $N^{1/2+d}$, où N est le nombre d'observations, converge en loi vers une variable aléatoire normale.

(f) les auto-covariances empiriques ne convergent pas asymptotiquement vers des variables aléatoires normales.

Le résultat présenté en (e) concernant la moyenne est intéressant pour les hydrologues. En effet, on a trouvé que des processus pour lesquels on pensait qu'ils possédaient une « mémoire à long terme » ont une moyenne empirique qui semble indiquer de lents changements de tendance (BOES et SALAS, 1978). Cette incohérence de la moyenne empirique peut être expliquée par le résultat (e) ci-dessus. Par conséquent, pour démontrer que la persistance dans les processus géophysiques peut être attribuée à un lent changement dans la tendance, on ne peut se baser uniquement sur l'analyse des observations disponibles du processus. On devrait utiliser aussi l'information tirée des caractéristiques physiques du ou des processus de génération. En effet, des processus à mémoire longue peuvent eux aussi présenter ce qui peut sembler être un lent changement de la moyenne dans les cas où le nombre d'observations utilisées pour calculer la moyenne empirique n'est pas très élevé. Mentionnons aussi que la validité des résultats présentés ci-haut ne se restreint pas uniquement aux cas des processus FARMA puisque ces résultats sont valides aussi pour n'importe quelle série ayant une FAC donnée par l'équation 4.

Une extension importante, quoique en apparence évidente, que l'on peut apporter à la définition originale des processus FARMA fournie par HOSKING (1981) et GRANGER et JOYEUX (1980), consiste à ne pas tenir compte de l'hypo-

thèse que la moyenne de la série soit nulle. L'extension du modèle présenté précédemment qui correspond au cas où la moyenne est non-nulle, est obtenue directement. Pourtant il est très important de remarquer que si une constante, en particulier la moyenne, passe par le filtre ∇^d , la sortie, dans le cas où d est positif, est nulle. Il en résulte qu'il n'est pas nécessaire que la moyenne figure dans les équations qui définissent le processus. Cependant, il faut mentionner que la moyenne est une quantité parfaitement définie pour ce processus lorsque $d < 0,5$.

La propriété mentionnée précédemment est très importante pour déterminer les propriétés stochastiques des estimateurs des paramètres. Cette importance vient du fait que la moyenne empirique peut être utilisée comme estimateur de la moyenne de la série. De plus la faible vitesse de convergence de la moyenne empirique, (voir le résultat donné en (e)) n'affecte pas la vitesse de convergence asymptotique vers une variable aléatoire Gaussienne des estimateurs des autres paramètres, cette vitesse étant N^{-1} .

Autre fait intéressant à noter, le filtre ∇^d peut lisser certaines tendances particulières comme on peut facilement le constater dans le cas $d = 1$ lorsque la tendance est une ligne droite. Lorsque $0 \leq d < 0,5$, le filtre ∇^d lisse les tendances qui changent lentement. Donc, même si la moyenne du processus change lentement, les modèles FARMA pourraient être utilisés pour représenter les séries chronologiques à peu près de la même façon que les modèles ARMA sont utilisés avec une composante déterministe de dérive.

Une autre conséquence due au fait que :

$$\nabla^d (x_t - \mu) = \nabla^d x_t, d > 0 \quad (5)$$

où x est la valeur du processus au temps t et où μ est la moyenne théorique, est que le comportement du processus est indépendant de la moyenne alors que, dans le cas d'un processus ARMA stationnaire, le comportement local du processus dépend de la moyenne. On peut le démontrer en considérant la valeur du processus, où la dépendance du passé est donnée par $E\{x_{t+1} | x_s, s \leq t\}$. Dans le cas des processus ARMA, cette quantité dépend de $\mu = E\{x_t\}$ mais dans le cas des processus FARMA avec $d > 0$, elle n'en dépend pas. A moins d'indications contraires, nous supposerons dans ce qui suit que la moyenne μ est nulle.

Une importante conséquence du fait que la vitesse de convergence vers zéro de la FAC est faible (équation 4), est que la formule de BARTLEH (1946) donnant les variances et les covariances de la fonction d'auto-covariance estimée (FACV) $\{\hat{\gamma}_k\}$ doit être modifiée. En fait, la formule exacte pour la variance est donnée par :

$$\text{var}(\hat{\gamma}_k) = N^{-1} \sum_{m=-(N-k)+1}^{(N-k)-1} \left\{ 1 - \frac{|m|+k}{N} \right\} \{ \gamma_m^2 + \gamma_m + k\gamma_m - k \} \quad (6)$$

et on peut en déduire à partir de l'équation 4

$$\text{var}(\hat{\gamma}_k) = \begin{cases} O(N^{-1}) & \text{si } d \leq 0,25 ; \\ O(N^{4d-2}) & \text{si } d > 0,25. \end{cases} \quad (7)$$

Donc, si $d < 0,25$ on a $\text{var}(\hat{y}_k) = O(N^{-1})$. Cette expression est du même ordre que dans le cas d'un processus à mémoire courte. Toutefois, si $0,25 < d < 0,5$, l'ordre de $\text{var}(\hat{y}_k)$ est plus grand que N^{-1} , et en fait, lorsque d tend vers $0,5$, la variance tend vers une quantité d'ordre 1. Ceci implique que la variabilité stochastique de la FACV estimée est plus élevée pour des processus à mémoire longue avec $0,25 < d < 0,5$ que pour des processus à mémoire courte. De plus, l'ordre de la variance dépend de la quantité inconnue d . Finalement, des résultats semblables sont valables pour les covariances de la FAC estimée.

Le processus FARMA $(0,d,0)$, qu'on appelle « modèle différencié fractionnaire » (fractional differencing model) est un sous-ensemble particulièrement intéressant de la famille des processus FARMA (p,d,q) . Ce modèle a été étudié en détail et on connaît les relations donnant la FAC, la fonction partielle d'autocorrélation (FPAC), les coefficients de régression linéaire partielle, ainsi que les autocorrélations inverses (HOSKING, 1981). Une caractéristique importante du comportement stochastique d'un processus FARMA $(0,d,0)$ est que toutes ses autocorrélations sont positives si d est positif et qu'elles sont négatives autrement. De plus, toutes les autocorrélations partielles du modèle FARMA $(0,d,0)$ ont le même signe que le paramètre de persistance d et leur taux de décroissance vers zéro est du même ordre que l'inverse du décalage. Ces limites inhérentes à la structure de la FAC obligent à filtrer le bruit différencié fractionnaire à travers un modèle ARMA, cette opération ayant pour but d'obtenir une structure d'autocorrélation plus riche à l'intérieur même du cadre d'un processus FARMA (p,d,q) .

Quelques généralisations

On peut généraliser le filtre $(1 - B)$ dans une autre forme qui soit étroitement reliée au filtre $(1 - B)^d$. Il s'agit du filtre défini par $(1 + B)^d$. Notons que la fonction de transfert qui leur est associée possède aussi une racine sur le cercle unitaire au point $B = -1$. Les coefficients de ce filtre sont les mêmes que ceux du filtre $(1 - B)^d$ à l'exception du signe et par conséquent, le processus possède lui aussi une « mémoire à long terme » si $d > 0$, il est stationnaire si $d < 0,5$ et il est inversible si $d > -0,5$. Par contre il est important de noter que même si la valeur absolue des autocorrélations et la même pour les deux filtres, les autocorrélations du filtre $(1 + B)^d$ sont de signes alternés. On pourrait obtenir des structures d'autocorrélations plus générales en généralisant les filtres afin qu'ils admettent des racines complexes sur le cercle unitaire. La classe de processus considérée dans cette étude est formée de cas particuliers de processus plus généraux qui résultent d'un filtrage du bruit blanc à travers des filtres définis par $(1 - \epsilon B)^d$, où le paramètre ϵ se situe dans l'intervalle $|\epsilon| \leq 1$. Dans ce qui suit on suppose que $\epsilon = 1$ ou -1 .

SIMULATION

A partir des expressions approximatives que l'on connaît pour les coefficients $\phi_{k,t}$ de la régression linéaire partielle, on peut développer des algorithmes qui permettent de générer rapidement des séquences synthétiques à partir des modèles FARMA. Les coefficients de régression linéaire partielle sont définis comme étant les valeurs de α_k qui minimisent :

$$E\{x_t - \alpha_1 x_{t-1} - \dots - \alpha_k x_{t-k}\}^2, \quad (8)$$

où E représente l'espérance mathématique. Par conséquent, ces coefficients sont les valeurs qui minimisent les erreurs de prédiction au pas de temps suivant. On sait que le processus de série chronologique peut s'écrire avec ces modifications :

$$x_t = e_t + \phi_{1,t} x_{t-1} + \dots + \phi_{j,t} x_{t-j}, \quad (9)$$

où les $\{e_t\}$ sont une séquence de variables aléatoires gaussiennes indépendantes ayant une moyenne 0 et une variance $\prod_{i=1}^j (1 - \phi_{i,t}^2)$. Si nous considérons d'abord le cas où on a à simuler des bruits différenciés fractionnaires, les expressions pour $\phi_{k,t}$ sont données par HOSTING (1981) sous une forme récursive :

$$\begin{aligned} \phi_{t,t} &= d/(t-d) & \text{où } J = t-1, \dots, 1 \\ \phi_{j,t} &= \phi_{j+1,t} (j+1) (t-j-d) / ((j-1-d) (t-j)) \end{aligned} \quad (10)$$

Donc, pour simuler un bruit FARMA (0,d,0) il suffit de calculer de façon récursive $\phi_{k,t}$, de générer une variable aléatoire normale et d'utiliser l'équation 9.

Pour une simulation utilisant un modèle FARMA (0,d,q), le bruit différencié fractionnaire doit être généré et passé ensuite à travers le filtre de moyenne mobile. Lorsqu'on génère des données synthétiques en utilisant un modèle FARMA (p,d,o), une des méthodes qui peut être utilisée consiste à simuler le modèle FARMA (0,d,0) en utilisant l'algorithme décrit plus haut, à choisir ensuite p valeurs initiales (en utilisant la méthode présentée par McLEOD et HIPEL (1978b) par exemple), et finalement, comme on l'a montré précédemment, à générer de manière récursive les autres valeurs simulées. Finalement le cas général FARMA (p,d,q) peut être obtenu par une combinaison de ces différentes méthodes.

Pour générer des séquences synthétiques on peut aussi utiliser une autre méthode (McLEOD et HIPEL, 1978a, b) qui consiste à obtenir la décomposition de Cholesky de la matrice des autocorrélations théoriques, Σ , et à multiplier cette matrice de décomposition par un vecteur formé de variables Gaussiennes indépendantes ayant une moyenne 0 et la variance désirée. De plus on ajoute à la série une correction de la moyenne. Cette méthode, quoique attrayante pour d'autres modèles, est moins avantageuse que la méthode décrite plus haut parce qu'elle nécessite le calcul de la matrice des autocorré-

lations dont les valeurs théoriques sont données en terme de fonctions hypergéométriques (HOSKING, 1981). Par conséquent, le temps de calcul des auto-corrélations est beaucoup plus élevé que celui nécessaire pour passer le bruit FARMA (0,d,0) à travers les différents filtres. Toutefois, lorsque la FAC a été calculée et qu'on a obtenu la décomposition de Cholesky requise, cette méthode s'avère utile dans le cas où plusieurs réalisations indépendantes du processus doivent être simulées. Finalement, les deux méthodes sont équivalentes dans le cas du modèle FARMA (0,d,0).

Prévisions

L'utilisation des modèles ARMA dans un but de prévision, est généralement plus profitable lorsqu'on s'intéresse à des prévisions à un ou deux pas de temps suivants. Ceci est dû au fait que les fonctions de prévision produites par les modèles ARMA convergent rapidement et ce de façon exponentielle vers la moyenne de la série. Par conséquent, dans les modèles ARMA, les prévisions à long terme sont données par la moyenne μ ou par un de ses estimateurs. A partir de la définition même de la persistance, il est évident que ce n'est pas le cas lorsqu'un processus possède une « mémoire à long terme ». Pour un processus qui possède une « mémoire à long terme » les fonctions de prévision convergent elles aussi vers la moyenne μ mais dans ce cas la vitesse de convergence n'est pas exponentielle mais plus lente. Pour des séries persistantes la vitesse de décroissance de la fonction de prévision dépend du degré de persistance du processus.

Dans une perspective de prévision, la présence de persistance implique aussi que la variance de la fonction de prévision décroît vers la variance du processus σ_x^2 avec une vitesse qui pourrait être beaucoup plus lente qu'une fonction exponentielle selon le degré de persistance. Donc, les limites de confiance pour les prévisions à n pas de temps des processus persistants sont plus petites que celles des processus ayant une « mémoire à court terme », si n est plus grand que deux ou trois. On peut observer ceci en écrivant le modèle de la série sous la forme d'un processus linéaire (BOX et JENKINS, 1976) :

$$x_t = \sum_0^{\infty} \alpha_k a_{t-k} \quad (11)$$

Alors, la prévision à n pas de temps $x_t [n]$, est donnée par

$$x_t [n] = \sum_0^{\infty} \alpha_{k+n} a_{t-k} \quad (12)$$

mais, pour un modèle FARMA, $\alpha_k \sim k^{-1-d}$. Alors,

$$\text{var}\{x_t [n]\} \sim \sigma_x^2 \sim O(n^{-1+2d}) \quad (13)$$

L'équation 12 est surtout utile dans un cadre de prévision, si les estimations de a_t sont disponibles et si les coefficients α_k décroissent vers zéro assez

rapidement pour que la troncature provoquée par le calcul de $x_t [n]$ (voir équation 12) produise une erreur qui puisse être négligée. Pour les modèles FARMA, ces coefficients ne décroissent pas assez rapidement et c'est pourquoi il peut être utile de donner les expressions pour une fonction de prévision $x_t [n]$ qui n'implique pas d'approximations. La méthode proposée ici est basée sur la forme de la série donnée par l'équation 8. La fonction de prévision est :

$$x_t [n] = \phi_{1,t} (n) x_t + \dots + \phi_{t,t} (n) x_0 \quad (14)$$

où

$$\phi_{i,t} (n) = \phi_{n,t+n} + \sum_{j=1}^{n-1} \phi_{j,t+n-1} \phi_{i,t} (j) \quad (15)$$

Si on la compare avec l'équation 12, cette expression a l'avantage de n'impliquer aucune approximation que celle-ci provienne de la troncature d'une série infinie ou du calcul de valeurs résiduelles. De plus, en utilisant l'équation 14, il est possible de montrer que :

$$x_t [n] \approx \phi_{n,t+n} x_t + \dots + \phi_{t+n,t+n} x_0 \quad (16)$$

Comme on l'a mentionné précédemment, la fonction de prédiction décroît donc vers la moyenne théorique μ avec une vitesse plus lente qu'une fonction exponentielle. Par exemple, pour le modèle FARMA (0,d,0),

$$x_t (n) \approx \frac{n^{-d-1} x_t + \dots + (t+n)^{-d-1} x_0}{(-d-1)!} \quad (17)$$

ESTIMATIONS DES PARAMÈTRES

Comme on l'a déjà indiqué les coefficients des filtres $(1 - B)^d$ et $(1 + B)^d$ ne diffèrent que par leurs signes. Donc, étant donné que les résultats qui seront présentés au sujet de l'estimation sont valides pour les deux filtres, nous ne donnerons ici les résultats que pour le filtre $\nabla^d = (1 - B)^d$.

Il existe plusieurs méthodes d'estimations. Celles qui se rapportent au domaine des fréquences ne semblent pas aussi efficaces que les estimations se rapportant au domaine de la représentation temporelle. C'est pourquoi nous considérons ici seulement les méthodes relatives au domaine temporel.

Puisque la moyenne empirique ne converge que lentement vers la vraie valeur (propriété (e)), il est primordial de trouver un estimateur plus efficace que la moyenne. Le plus immédiat est l'estimateur de la moyenne obtenu par la méthode du maximum de vraisemblance (McLEOD et HIPEL, 1978a) qui est donné par :

$$\hat{\mu} = 1 \Sigma^{-1} X / 1 \Sigma^{-1} 1 \quad (18)$$

où Σ est la matrice d'autocovariance de la série, où $\mathbf{1}$ représente un vecteur colonne de 1 et \mathbf{X} représente le vecteur des observations. On peut montrer que la moyenne empirique est un estimateur efficace dans le cas où le paramètre de persistance d est non négatif mais qu'elle ne l'est pas lorsque le paramètre de persistance est négatif. Ceci concorde avec le fait déjà bien connu que la sur-différenciation peut conduire à des estimateurs inefficaces.

Bien qu'il soit difficile de donner un sens physique à l'anti-persistance, une valeur négative de d peut être pratique si on ne considère que la notion d'ajustement. En effet, il peut arriver que l'ajustement de modèles FARMA à des séries conduise à des valeurs de d négatives et dans ces cas il est important d'estimer la moyenne du processus en utilisant l'estimateur du maximum de vraisemblance tel que décrit par la relation 18. L'évaluation des termes de cette relation peut s'effectuer efficacement soit en utilisant la décomposition de Cholesky pour l'inverse de Σ donné par les coefficients de la régression linéaire partielle $\{\phi_{t,t}\}$ qui peuvent être calculés par l'algorithme de Levison-Durbin (DURBIN, 1960) ou en utilisant l'algorithme de Trench permettant d'inverser une matrice de Toeplitz (TRENCH, 1964).

Dans le cas particulier du bruit différencié fractionnaire, HOSKING (1981) donne une expression approximative pour les coefficients de réflexion ou coefficients de régression approximative pour l'estimateur de la moyenne par la méthode du maximum de vraisemblance. Si dans le filtre $\nabla^d = (1 - \epsilon B)^d$ on a $\epsilon = -1$, cette expression approximative reste valable pour autant qu'on effectue le changement de signe approprié. Pour calculer l'estimateur du maximum de vraisemblance μ à partir des coefficients de la régression linéaire partielle on utilise l'expression :

$$\hat{\mu} = \frac{\sum_{t=0}^{N-1} (x_t - \phi_{1,t} x_{t-1} - \dots - \phi_{t,t} x_0)}{\sum_{t=0}^{N-1} (1 - \phi_{1,t} - \dots - \phi_{t,t})^2} \quad (19)$$

Dans cette étude on suppose que le paramètre de persistance d est non négatif, que la moyenne empirique est utilisée comme estimateur de la moyenne et que les observations ont été soustraites de la moyenne de l'échantillon.

Pour estimer les autres paramètres dans le domaine temporel, deux méthodes peuvent être utilisées. Il s'agit de l'estimation exacte par la méthode du maximum de vraisemblance ou d'une approximation du filtre ∇^d . La plupart des algorithmes d'estimations par la méthode du maximum de vraisemblance dépendent du calcul de l'erreur de prédiction au pas de temps suivant et qu'elles peuvent être exprimées en fonction des coefficients de la régression linéaire partielle. Ces coefficients peuvent être efficacement évalués par l'algorithme de Durbin-Levison. Enfin, connaissant ces valeurs e_t , les estimateurs des paramètres sont obtenus en minimisant la somme des carrés modifiée donnée par :

$$\log l = \sum (N - t + 1) \log(1 - \phi_{t,t}^2) + \sum e_t^2 \quad (20)$$

Du point de vue statistique l'estimation par le maximum de vraisemblance est intéressante mais l'ampleur des calculs nécessaires à la réalisation des

étapes décrites plus haut nous amène à considérer des algorithmes nécessitant un moins grand nombre de calculs comme des alternatives intéressantes.

L'algorithme proposé par LI et McLEOD (1986) est un algorithme « économique » au point de vue calcul. Il consiste d'abord à évaluer approximativement le filtre ∇^d par le filtre ∇_M^d . ∇_M^d est défini comme étant le nouveau filtre ne comprenant que les M premiers termes du filtre ∇^d , c'est-à-dire le filtre qui résulte de l'approximation du processus par un modèle autorégressif « long ». Par la suite la somme des carrés des résidus (où les résidus sont obtenus à la sortie des filtres ∇_M^d et du filtre ARMA) est minimisée. Pour calculer les résidus, on peut utiliser un algorithme comme celui donné par McLEOD et SALES (1983). La somme des carrés des résidus pourrait aussi être étendue en arrière dans le temps, comme l'ont recommandé BOX et JENKINS (1976), en utilisant la rétro-prévision. Il faut noter ici que l'approximation de ∇^d par ∇_M^d n'est pas optimale au sens des moindres carrés. Toutefois, puisque l'ordre de M est comparable à N cette approximation doit donner un résultat très proche de la valeur approximative optimale. JIMÉNEZ et McLEOD (1986) ont trouvé qu'il faut un ordre de $N^{1/2}$ pour obtenir des estimateurs qui soient convergents. Une règle couramment utilisée consiste à utiliser des séries temporelles qui contiennent au moins 50 observations pour effectuer l'ajustement. L'ordre de troncature M est choisi entre $N/4$ et $N^{1/2}$. Ce choix est effectué en essayant de trouver le meilleur compromis entre le degré d'approximation du filtre ∇^d par le filtre ∇_M^d et l'importance des calculs à effectuer. Pour des valeurs de N voisines de 50 on fixe M à la moitié du nombre d'observations. Le nombre de calculs à effectuer lorsqu'on utilise cet algorithme est beaucoup plus faible qu'en utilisant la méthode du maximum de vraisemblance. De plus, les estimateurs obtenus de cette façon sont asymptotiquement équivalents à ceux du maximum de vraisemblance et il semble que les estimateurs dans le cas d'échantillons finis sont en général assez voisins de ceux du maximum de vraisemblance. LI et McLEOD (1986) ont étudié les distributions asymptotiques des estimateurs dans le cas où la moyenne de la série est connue. Ils ont obtenu des expressions approximatives pour les variances des distributions asymptotiquement normales des estimateurs. On peut démontrer que l'estimation de la moyenne par la moyenne empirique n'affecte pas ces résultats asymptotiques. Toutefois, ces résultats ne sont probablement pas aussi valables lorsque la taille de l'échantillon est finie car dans ce cas il y a persistance à long terme et le paramètre d doit se trouver dans l'intervalle ouvert $(-0,5, 0,5)$. En pratique, cet intervalle est fermé et par simulation on peut observer que si le paramètre de persistance est voisin de 0,5 alors la probabilité que la valeur estimée de d soit égale à 0,5 est élevée. Un phénomène semblable a été observé par CRYER et LEDOLTER (1981) pour le modèle ARMA (0,1). Donc, la vitesse de convergence des estimateurs dépend des paramètres même pour des échantillons de taille relativement élevée, supérieure à 200. De plus il est important de noter que le modèle décrit plus haut ressemble beaucoup à l'ajustement d'un processus autorégressif d'ordre 1 si d n'est pas trop voisin de 0,5, par exemple, si d est inférieur à 0,3.

APPLICATION DU BOOTSTRAP SUR UN MODÈLE DE SÉRIE CHRONOLOGIQUE

Puisque le modèle FARMA est un modèle autorégressif infini et qu'il est non stationnaire lorsque $d \geq 0,5$, on s'attend à ce que les propriétés d'échantillonnage des estimateurs obtenus pour des échantillons ayant une taille finie soient différentes des approximations correspondant à des échantillons de grande taille. Par conséquent, il est intéressant d'obtenir davantage d'information au sujet des distributions d'échantillonnage relatives à ces échantillons finis. Pour augmenter nos connaissances au sujet de ces distributions, la technique du bootstrap proposée par COVER et UNNY (1986) semble une avenue intéressante à explorer.

Depuis que EFRON (1979) a proposé le bootstrap, plusieurs auteurs ont suggéré d'étendre la méthode originale à l'analyse des séries chronologiques. La majorité de ces travaux ont toutefois porté sur une généralisation directe de la procédure du bootstrap aux modèles de régression, procédure qui a été proposée par FREEMAN (1981). Toutefois, les processus de séries chronologiques dépendent des résidus passés et ce, de façon récursive. Donc, si on veut obtenir une bonne approximation de la loi de probabilité des estimateurs des paramètres en ré-échantillonnant les résidus ajustés et en les introduisant dans le modèle ajusté, il est primordial que la distribution de probabilité cumulée des résidus estimés donne une approximation précise de la vraie distribution de probabilité cumulée des résidus. Cette condition doit aussi être remplie dans le cas des modèles de régression ordinaire. Toutefois les erreurs introduites en utilisant les résidus ajustés entrent dans la série sous une forme nouvelle et peuvent donc avoir un effet cumulatif qui n'est pas présent dans un modèle régressif ordinaire pour lequel les erreurs entrent dans le modèle sous une forme additive. Il semble donc possible d'améliorer les procédures de bootstrap simples et c'est ce que COVER et UNNY (1986) ont fait en présentant un prolongement prometteur de la méthode du bootstrap.

L'idée de COVER et UNNY (1986) consiste à introduire un élément aléatoire dans la fonction de perte en ré-échantillonnant les positions des résidus et non pas les observations elles-mêmes. Ainsi les temps de décalage sont ré-échantillonnés avec remplacement de manière équiprobable. Cette idée de ré-échantillonner les temps de décalage est particulièrement intéressante en raison de la nature même des données qui dépend fortement du temps. Elle peut aussi être appliquée à n'importe quel modèle de séries chronologiques et non pas seulement à un modèle FARMA. De plus, cette idée reste valable pour d'autres processus stochastiques.

La technique à utiliser peut être décrite selon les étapes suivantes :

(a) Tirer un échantillon aléatoire de taille N , avec remplacement, à partir de la série des entiers compris entre 1 et N .

(b) Obtenir les valeurs estimées des paramètres en minimisant la somme des carrés des résidus et en utilisant des poids égaux au nombre de fois que le chiffre t apparaît dans l'échantillon aléatoire obtenu en (a).

(c) Répéter les étapes (a) et (b) un assez grand nombre de fois pour pouvoir obtenir une estimation fiable de la distribution des paramètres estimés.

Comme on le verra dans les applications, cette technique peut augmenter considérablement l'information que l'on possède sur les estimations des paramètres.

Toutefois, d'autres résultats théoriques devront être obtenus afin de confirmer théoriquement la validité de l'approche du bootstrap dans le cas des échantillons de taille finie.

APPLICATIONS PRATIQUES

Afin de démontrer les possibilités d'applications des modèles FARMA 14 séries présentées au tableau 1 sont analysées.

Tableau 1 Description des séries de débits.

Table 1 *Description of river flow records.*

Nom de la série	Situation géographique	Période	Longueur
Saugæen	Walkerton, Canada	1915-1976	62
Dal	Près de Norslund, Suède	1852-1922	70
Danube	Orshava, Roumanie	1837-1957	120
French Broad River	Asheville, Caroline du Nord, E.U.	1830-1900	70
Gota	Près de Sjotop-Vannersburg, Suède	1807-1957	150
Mckenzie	Mckenzie Bridge, Orégon, E.U.	1900-1956	55
Mississipi	St Louis, Missouri, E.U.	1861-1957	96
Neumunas	Smallininkai, Lituanie, U.R.S.S.	1811-1943	132
Rhin	Bâle, Suisse	1807-1957	150
Fleuve St-Laurent	Ogdensburg, New York, E.U.	1800-1930	131
Tamise	Teddington, Angleterre	1883-1954	71
Précipitations à Philadelphie	Philadelphie, E.U.	1800-1898	99
Précipitations à Fortaleza	Fortaleza, Brésil	1849-1979	131
Températures moyennes	Centre de l'Angleterre	1723-1970	248

Les données comprennent 11 séries de débits (en m³/s) de rivières réparties à travers le monde, 2 séries d'enregistrements de précipitations annuelles moyennes en mm, et finalement une série de températures annuelles mesurées en degrés Celsius.

Pour vérifier si on est en présence de processus à mémoire longue, il est difficile en pratique d'utiliser directement la définition de ce type de mémoire puisqu'elle s'exprime sous la forme $M = \infty$. C'est pourquoi on utilise le critère basé sur la persistance donné à l'équation 4. Donc, une FAC qui décroît lentement vers 0 peut indiquer que la série a une « mémoire à long terme ». Pour les séries qui ont une FAC empirique qui décroît vers zéro de manière hyperbolique, on examine ici la possibilité de les représenter par les modèles

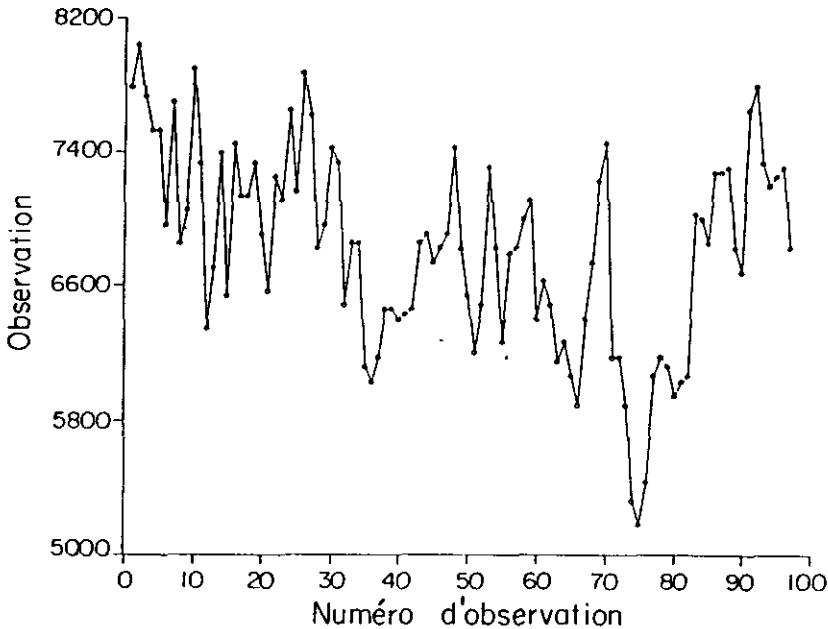


Figure 1 Fleuve St-Laurent à Ogdensburg, N.Y., 1860-1957, débit annuel.
St. Lawrence River at Ogdensburg, N.Y., 1860-1957, yearly flow.

FARMA. Parmi les 14 séries de données présentées dans le tableau 1, la série des débits du fleuve St-Laurent et la série de précipitations de Philadelphie présentent toutes deux une valeur estimée de FAC qui semble décroître vers 0 de manière hyperbolique.

Le comportement de ces deux séries suggère donc d'utiliser des modèles FARMA pour les ajuster. Les figures 1 et 2 représentent respectivement la série de débits du fleuve St-Laurent en fonction du temps et la FAC de la moyenne annuelle de ces débits. Les figures 3 et 4 montrent les mêmes graphiques pour les précipitations de Philadelphie. Les autres séries comme celle de la rivière Saugeen et celle des précipitations à Fortaleza, ont des valeurs estimées de la FAC qui n'indiquent pas de façon très prononcée la présence d'une persistance bien qu'elle puisse exister. La détection visuelle, à partir des auto-corrélations, de la présence d'une « mémoire à long terme » peut devenir difficile lorsque le paramètre de persistance d est voisin de 0 (entre 0 et 0,2). De plus, étant donné que la formule de Bartlett doit être multipliée par un facteur d'ordre N^{-1+4d} si $d \geq 0,25$, l'examen visuel des FAC des échantillons doit être utilisé avec prudence dans les cas de processus FARMA (p,d,q) susceptibles d'avoir une « mémoire à long terme ».

Dans le cas où le processus appartient à la famille des modèles FARMA, les autocorrélations partielles devraient décroître vers zéro de manière hyperbolique, la vitesse de décroissance étant indépendante du degré de

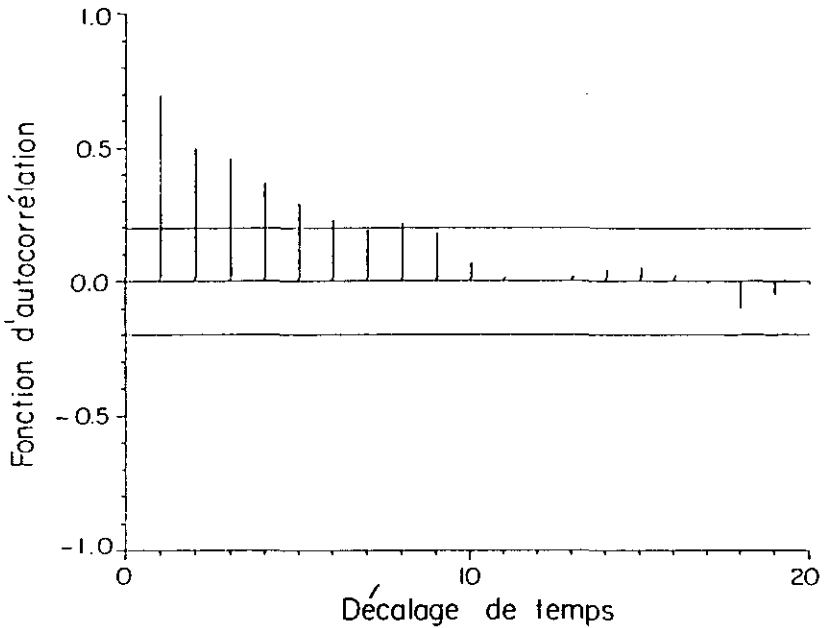


Figure 2 Fleuve St-Laurent à Ogdensburg, N.Y.
St. Lawrence (main stream) near Ogdensburg, N.Y.

persistance. Pour le cas particulier d'un modèle FARMA (0,d,0), une mémoire de « long terme » implique que toutes les auto-corrélations doivent être positives. Ce comportement de la fonction d'auto-corrélation partielle (FACP) du processus FARMA (0, d, 0) suggère que la détection de la persistance nécessite qu'on ne s'intéresse non seulement à la seule décroissance hyperbolique de la FACP mais aussi au comportement de ses signes. Ceci nous conduit à l'utilisation d'un test de signes non paramétrique qui permet de tester les signes de la FACP estimée. Parmi l'ensemble des estimations de la FACP pour les 14 séries de données considérées, aucune ne présente de façon significative une décroissance hyperbolique vers zéro. Toutefois, quelques-unes comme par exemple la série des débits du fleuve St-Laurent et celle des précipitations de Philadelphie, présentent des structures de la FACP qui sont généralement positives. L'application d'un test de signes sur ces FACP permet d'appuyer l'hypothèse que ces deux séries présentent des signes de persistance.

Une autre caractéristique associée à une série peut indiquer la présence d'une persistance. Il s'agit du comportement des moyennes partielles de l'échantillon, \bar{x}_k , du processus qui sont définis par :

$$\bar{x}_k = \frac{(x_{k-1} + \dots + x_0)}{k} \quad (21)$$

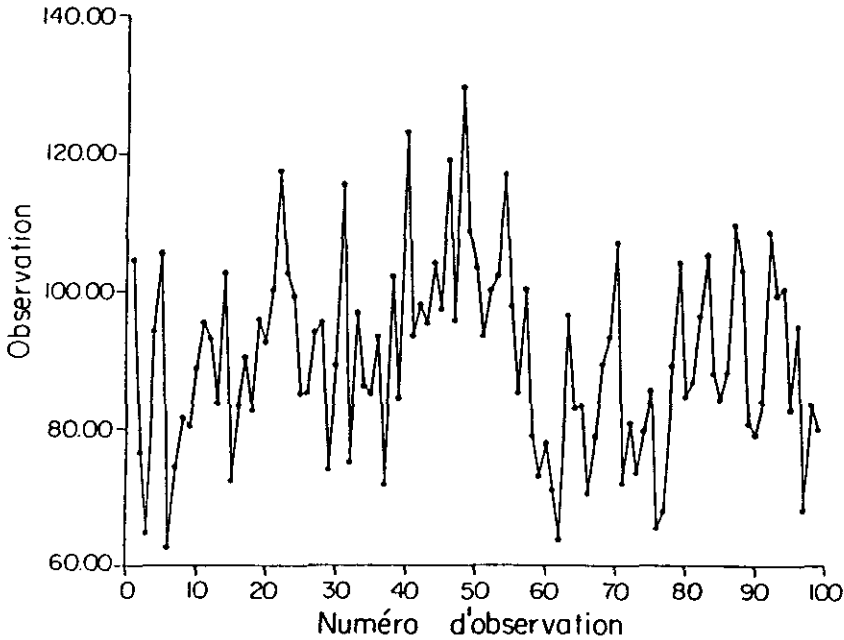


Figure 3 Moyenne annuelle des précipitations à Philadelphie, 1820-1950.
Average annual rainfall at Philadelphia, 1820-1950.

Dans le cas d'une série avec mémoire de « court terme », le tracé des valeurs de \bar{x}_k en fonction de k devrait montrer une grande variabilité stochastique pour les premières valeurs de k , mais, lorsque k atteint une certaine valeur, le tracé devrait montrer une décroissance vers une valeur à peu près constante et présenter une faible variabilité stochastique. Toutefois, dans le cas d'une série avec mémoire de « long terme » le même type de tracé devrait montrer une forte variabilité stochastique pour les quelques premières valeurs de k et ensuite, à partir d'une certaine valeur de k , une faible tendance qui oscillerait autour d'une valeur constante lorsque k augmente. Lorsque k devient très élevé (cette valeur de k dépend du degré de persistance), \bar{x}_k devrait atteindre une valeur constante. Etant donné que les valeurs actuelles de la série sont corrélées avec les valeurs passées, les valeurs actuelles et les valeurs passées de \bar{x}_k sont fortement corrélées. Par conséquent, le tracé de \bar{x}_k en fonction de k peut montrer des tendances locales. Afin de détecter une persistance il est intéressant de considérer la vitesse de décroissance de ces tendances locales vers une constante et aussi d'examiner la présence d'une légère tendance globale. Toutefois, la présence de tendances locales dans le tracé de \bar{x}_k en fonction de k ne constitue pas en elle-même une indication de la présence d'une persistance. Parmi les 14 séries considérées, les débits du St-Laurent montrent une tendance générale décroissante.

Cette tendance est assez modérée pour supposer qu'elle peut être due à la présence d'une persistance dans la série et non pas à une non-stationnarité. Le tracé de \bar{x}_k en fonction de k pour les débits du fleuve

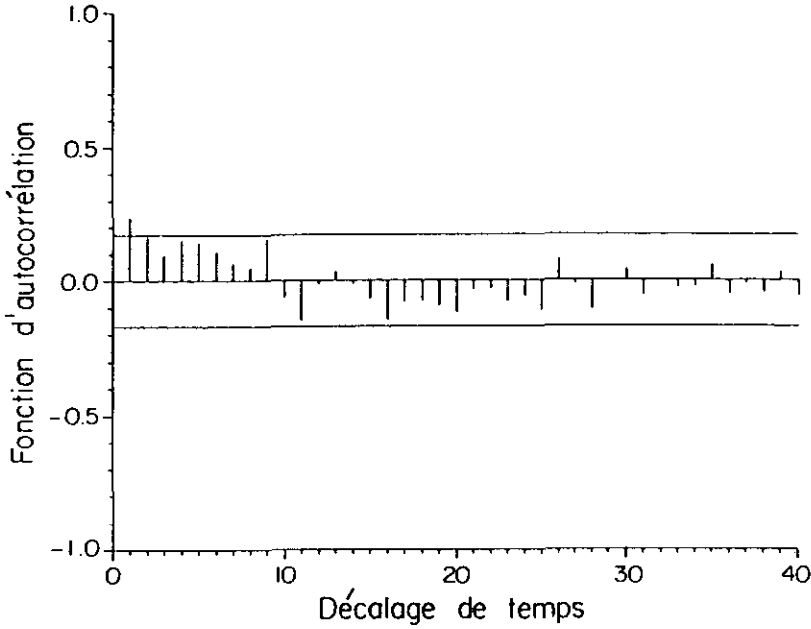


Figure 4 Précipitations mensuelles (en mm) à Philadelphie, 1820-1950.
Monthly precipitation, mm, Philadelphia, 1820-1950.

St-Laurent est présenté à la figure 5. Pour quelques-unes des autres séries, des tendances locales en x_k semblent exister même à la fin de la série. Enfin, pour la majorité des autres séries, le comportement des moyennes partielles est cohérent avec celui des séries à mémoire « court terme », c'est-à-dire, une décroissance rapide du tracé vers une valeur constante.

Tous les modèles FARMA examinés pour l'ajustement des séries présentées au tableau 1, sont des sous-ensembles du modèle FARMA (2,d,1) donné par :

$$(1 - \phi_1 B - \phi_2 B^2) \nabla^d (B) (x_t - \mu) = (1 - \phi_1 B) a_t \quad (22)$$

où ϕ_i est le paramètre d'autorégression d'ordre i et ϕ_1 est le premier paramètre de moyenne mobile. Dans le cas de la série des débits du fleuve St-Laurent le modèle AR(3), donné par :

$$(1 - \phi_1 B - \phi_3 B^3) (x_t - \mu) = a_t \quad (23)$$

a aussi été examiné parce qu'il s'agit du modèle sélectionné par Mc LEOD et al. (1977) parmi la classe des modèles ARMA. Pour chacune des séries de données, le modèle FARMA tiré de l'équation 22 qui était le plus adéquat a été sélectionné en utilisant le critère d'information minimum d'Akaike (CIA) (AKAIKE, 1974). Il s'agissait de ne considérer que les modèles pour lesquels l'hypothèse de bruit blanc des résidus ajustés était acceptée. Le tableau 2 présente pour chaque série, les valeurs d'estimation des paramètres du

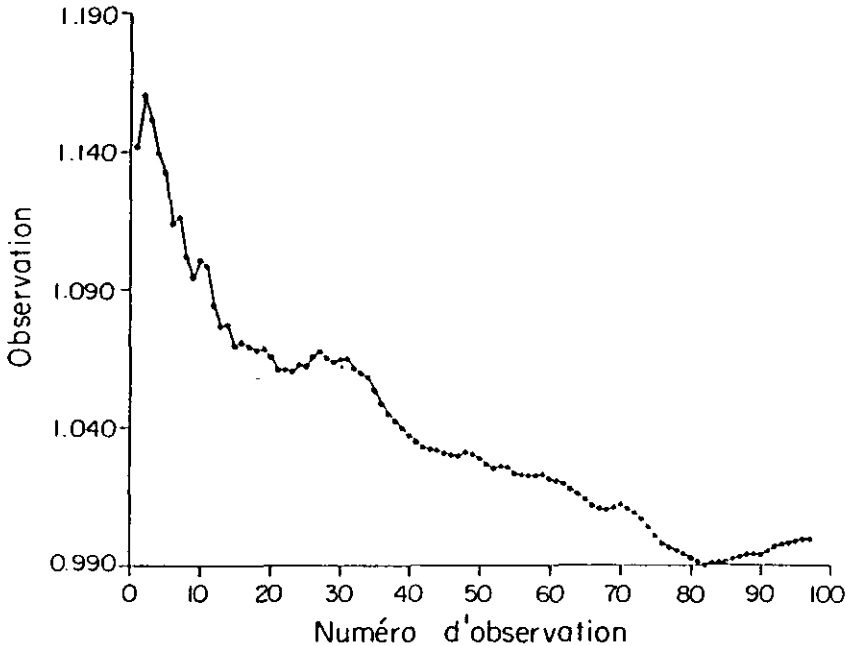


Figure 5 Sommes partielles pour le fleuve St-Laurent, New York, 1860-1957.
Partial sums of the St. Lawrence, New York, 1860-1957.

modèle obtenues par la méthode du maximum de vraisemblance ainsi que leurs écarts types indiqués entre parenthèses. Les séries pour lesquelles la valeur d'estimation de d fournie par le tableau 2 est positive, présentent un comportement persistant. De plus, puisque le degré de persistance dépend de la valeur du paramètre d , les séries ayant de plus fortes valeurs pour d ont un degré de persistance plus élevé que celui des séries ayant une valeur de d faible. Par exemple, le modèle du fleuve St-Laurent a été estimé comme étant un modèle FARMA $(0,d,0)$ avec $d = 0,4999$ ce qui indique que les débits du fleuve St-Laurent sont très persistants, c'est-à-dire que la mémoire M est infinie et que par conséquent le passé éloigné influence le présent. La faible vitesse de convergence de la moyenne empirique vers la vraie moyenne est une conséquence de cette influence. Ce taux de décroissance est de l'ordre de $O(N^{-0,0001})$, où N est le nombre d'observations. Cet ordre de grandeur de la convergence est aussi valable pour la fonction de prédiction et pour la valeur estimée de la FAC dans le cas des débits du fleuve St-Laurent. Le fleuve St-Laurent présente la particularité d'avoir de grandes masses d'eau ce qui pourrait suggérer un modèle ayant un effet de réservoir pour lequel le pas de temps est plus grand que le pas de temps utilisé pour mesurer les valeurs de la série. Tous les modèles du tableau 2 qui présentent une persistance, et en particulier ceux relatifs au fleuve Mckenzie, à la Tamise et aux précipitations de Philadelphie, sont des modèles FARMA $(0,d,0)$. Il y a d'autres séries de données pour lesquelles le CIA a conduit au choix d'un modèle ARMA mais la différence entre le CIA minimum pour les modèles ARMA et pour les

modèles FARMA(0,d,0) était très faible. Enfin, notons que quelques rivières ne présentent aucun signe d'autocorrélation de deuxième ordre alors que le modèle optimal selon le CIA était simplement la moyenne. Il s'agit des séries du fleuve Dal, du Danube et du Rhin. Notons que la plupart des résultats obtenus concordent avec les modèles suggérés par la FAC, par la FACP et par le comportement des moyennes partielles.

Tableau 2 Les meilleurs modèles FARMA.

Table 2 Best FARMA models

Nom de la série	ϕ_1	ϕ_2	d	θ_1
Saugeen	-	-	-	-
Dal	-	-	-	-
Danube	-	-	-	-
French Broad R.	-0,234 (0,12)	-	-	-
Gota	0,587 (0,08)	-0,27 (0,08)	-	-
Mckenzie	-	-	(0,1)	-
Mississippi	0,29 (0,1)	-	-	-
Neumunas	-	-	-	-0,19 (0,08)
Rhin	-	-	-	-
Fleuve St-Laurent	-	-	0,499 (0,08)	-
Tamise	-	-	0,12 (0,1)	-
Précipitations à Philadelphie	-	-	0,23 (0,08)	-
Précipitations à Fortaleza	0,240 (0,08)	-	-	-
Températures moyennes	0,120 (0,06)	0,20 (0,06)	-	-

Les erreurs-types des valeurs d'estimation sont indiquées entre parenthèses en dessous de la valeur estimée.

Standard deviations of the estimates are given in parenthesis under the estimated value.

La technique du bootstrap proposée par Cover et Unny a été utilisée afin d'augmenter l'information de l'échantillon de taille finie pour l'estimation du paramètre de persistance d. Pour quelques-unes des séries de données, le CIA ne permet pas de séparer clairement les modèles ayant un paramètre de persistance d de ceux qui n'en ont pas. On remarque aussi que la plupart du temps le meilleur modèle FARMA ayant un paramètre de persistance était le modèle FARMA (0,d,0). Ces deux remarques nous amènent à nous intéresser à la fiabilité des estimations de d et de leur erreur type. C'est pour cela que l'on a utilisé la technique du bootstrap proposée par Cover et Unny, en considérant le modèle FARMA (0,d,0) pour toutes les séries de données. Même si ce modèle n'est pas adéquat pour quelques-unes des séries, il n'en reste pas

moins que la connaissance que l'on obtient du comportement des estimateurs de d reste précieuse. L'utilisation du bootstrap permet d'obtenir deux types d'information. Premièrement, pour chaque série on a calculé, par la technique du bootstrap, la moyenne et l'écart type de l'échantillon des valeurs estimées de d . Deuxièmement, la distribution des estimations de d pour chaque série de données a été évaluée en utilisant une estimation non paramétrique de la densité (FRYER, 1977). En utilisant ces deux types d'information, on peut décider si la série présente ou non une évidence de persistance. Les moyennes, les écarts types et les estimations de d obtenus en utilisant la méthode approchée du maximum de vraisemblance sont présentées dans le tableau 3. Les figures 6 et 7 présentent respectivement la fonction densité de probabilité des estimations de d pour les débits du fleuve St-Laurent et pour les précipitations observées à Philadelphie. Plusieurs conclusions peuvent être tirées de l'examen de ces figures : (a) Les approximations valables pour des grands échantillons ne le sont pas nécessairement pour des échantillons ayant une taille finie. Par exemple, la distribution de l'estimation du paramètre d pour le fleuve St-Laurent ne semble pas avoir de queue ; la densité semble être concentrée dans l'intervalle 0,492 – 0,4999, avec un mode voisin de 0,496 alors que l'ajustement initial peut nous laisser supposer la possibilité d'un comportement non stationnaire ; le bootstrap semble indiquer que même si le comportement statistique de la série est pratiquement non stationnaire, un ajustement adéquat peut être obtenu par le modèle FARMA (0,0,496,0). De plus, la fonction densité semble légèrement asymétrique. (b) La moyenne des estimations du paramètre d obtenues par ré-échantillonnage ainsi que l'écart type de ces estimations ne représentent pas nécessairement les données. (c) Il semble que les écarts types asymptotiques sont plus petits que les estimations

Tableau 3 Estimation du paramètre d obtenue en utilisant le bootstrap.

Table 3 Estimation of the parameter d using bootstrapping.

Nom de la série	Moyenne de d	Ecart-type	\hat{d}	Ecart type
Saugeen	0,110	0,212	0,108	0,100
Dal	0,028	0,177	0,024	0,093
Darube	0,069	0,157	0,059	0,072
French	0,148	0,168	0,134	0,093
Gota	0,365	0,245	0,388	0,064
Mckenzie	0,284	0,142	0,274	0,105
Neumunas	0,105	0,137	0,103	0,068
Précipitations Phil.	0,210	0,110	0,229	0,078
St.-Laurent	0,495	0,055	0,499	0,792
Tamise	0,139	0,149	0,120	0,093
Températures	0,153	0,079	0,151	0,050

La valeur du paramètre de persistance d du modèle $\nabla^d(B) x_t = a_t$ a été estimée en utilisant la technique du bootstrap telle que décrite dans notre étude. Cette valeur a été estimée par la valeur moyenne obtenue par bootstrap et elle est présente dans la deuxième colonne. La troisième colonne contient l'erreur type de l'estimation obtenue par bootstrap. La quatrième colonne donne la valeur de d qui est la valeur estimée de d obtenue par la méthode du maximum de

vraisemblance. La dernière colonne présente l'écart type asymptotique de d (qui est égal à $\sqrt{6/\pi^2 N}$).

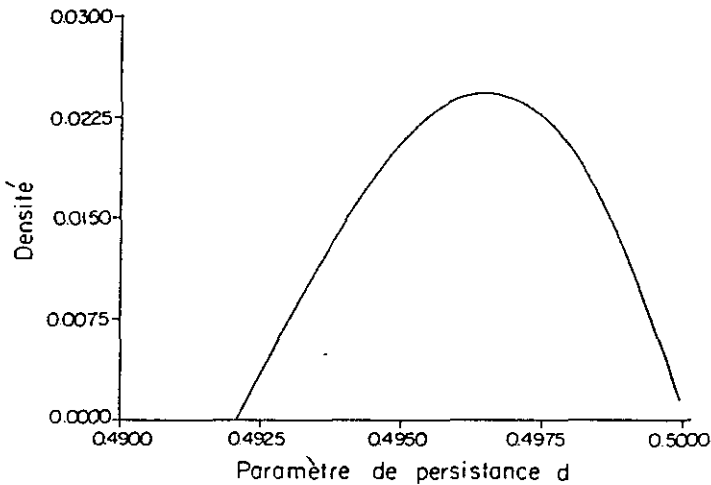


Figure 6 Densité de probabilité des estimations du paramètre de persistance d obtenues par la technique du bootstrap pour le fleuve St-Laurent.
Probability density of the persistence parameter d by bootstrapping for the St-Lawrence River.

obtenues par bootstrap pour des valeurs de d qui ne sont pas très près de 0,5. Ce comportement est inversé pour des valeurs de d voisines de 0,5.

CONCLUSION

Comme les applications le montrent, les récents développements relatifs aux modèles FARMA (p,d,q) permettent de représenter et d'estimer convenablement la persistance. La non-normalité du comportement asymptotique de l'auto-corrélation donne tout son intérêt à l'utilisation des modèles FARMA pour représenter la persistance. Le cas de la formule de Bartlett qui n'est plus valide pour les modèles à mémoire longue en est un exemple. Ceci implique que les limites de confiance utilisées habituellement ne sont pas valides et ne devraient donc être utilisées qu'à titre indicatif seulement. L'algorithme d'estimation selon la méthode approchée du maximum de vraisemblance proposé par LI et McLEOD (1986) a été généralisé au cas où la moyenne est inconnue. Un prolongement intéressant de cet algorithme consiste à utiliser une approximation optimale du filtre ∇^d par une auto-régression ce qui s'avère précieux lorsque la longueur de la série est courte, c'est-à-dire entre 50 et 100 observations. Cette approximation optimale est donnée par les coefficients de la régression linéaire partielle (équation 7).

Quelques généralisations du concept de différenciation sont aussi exposées dans cette étude. En raison de sa simplicité, le filtre $(1 + B)^d$ peut plus particulièrement être utilisé conjointement avec un filtre ARMA, ceci afin de

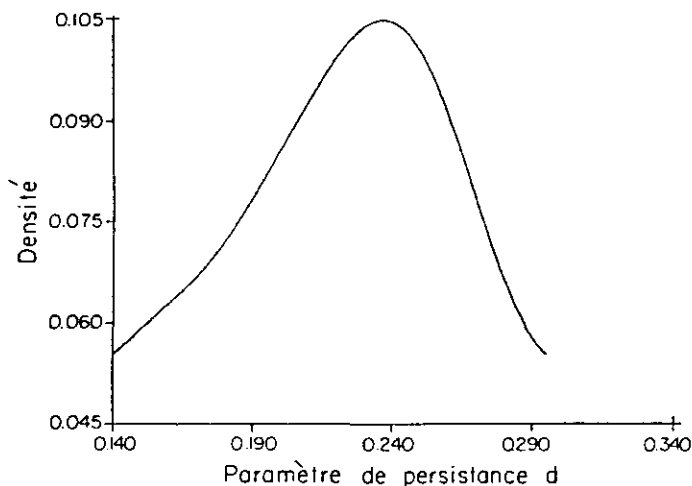


Figure 7 Densité de probabilité des estimations du paramètre de persistance d obtenues par la technique du bootstrap pour les précipitations à Philadelphie.

Probability density of the persistence parameter d by bootstrapping for the Philadelphia rainfall.

représenter la persistance. Ce filtre peut être utile en pratique si la FAC empirique présente une alternance de signes mais que sa valeur absolue décroît lentement vers zéro.

On a aussi présenté une procédure de simulation pour laquelle il n'est pas nécessaire de calculer la matrice d'auto-corrélation. Dans les cas où seulement quelques simulations sont nécessaires, cette procédure s'avère plus rapide que celles basées sur la décomposition de Cholesky.

Lorsque plusieurs simulations sont nécessaires on devrait considérer le coût d'obtention de la FAC par rapport au grand nombre de calculs impliqués pour la méthode proposée dans cette étude.

Finalement, la technique de bootstrap proposée par COVER et UNNY (1986) peut être utilisée pour augmenter la connaissance qu'on a de la distribution d'échantillonnage des paramètres dans le cas d'échantillons de taille finie. Ceci est surtout utile pour des modèles FARMA (p,d,q) lorsqu'il est difficile de décider si la série présente ou non un comportement persistant.

REMERCIEMENTS

Les auteurs désirent remercier le comité de lecture de la revue pour ses précieux commentaires, ainsi que toutes les personnes qui ont collaboré à la traduction de leur article en français.

APPENDICE

On utilise un algorithme de rétro-prévision semblable à celui utilisé par McLEOD et HOLANDA SALES (1983) afin de calculer la somme inconditionnelle des carrés des résidus obtenue lorsqu'on suppose que le modèle peut être représenté par une approximation de longue auto-régression

$$\phi(B) \nabla_M^d(B) x_t = \theta(B) a_t \quad (24)$$

Un algorithme de prévision en arrière semblable à celui de McLEOD et HOLANDA SALES (1983) permet de calculer la somme inconditionnelle des carrés des résidus pour un modèle ARMA saisonnier. La somme inconditionnelle des carrés des résidus

$$S = \sum_{t=-\infty}^N [a_t]^2 \quad (25)$$

(où $[.]$ représente l'espérance par rapport aux observations) est approximativement obtenue par :

$$S = \sum_{t=1-Q}^N [a_t]^2 \quad (26)$$

où Q est un point de troncature relativement élevé. La forme conditionnelle de l'équation 24 est donnée par :

$$\phi(B) \nabla_M^d(B) [x_t] = \theta(B) [a_t] \quad (27)$$

où $[a_t] = 0$, $t > N$. On peut aussi l'exprimer sous la forme d'un modèle à deux étapes:

$$\nabla_M^d(B) [x_t] = [c_t] \quad (28)$$

et

$$\phi(B) [c_t] = \theta(B) [a_t] \quad (29)$$

Pour l'approche selon la rétro-prévision de Box-Jenkins on a aussi besoin de la version en avant de l'équation 24

$$\phi(F) \nabla_M^d(F) x_t = \theta(F) e_t \quad (30)$$

où F est l'opérateur de décalage de temps en avant et e_t est une suite de variables aléatoires indépendantes et normales avec une moyenne nulle. La méthode utilise la forme conditionnelle de l'équation 30

$$\nabla_M^d(F) [x_t] = [B_t] \quad (31)$$

et

$$\phi(F) [B_t] = \theta(F) [e_t] \quad (32)$$

où $e_t = 0$, $t < 1$. La somme inconditionnelle des carrés peut être obtenue par l'exécution successive des étapes suivantes :

Etape 0 : Choisir Q et M.

Etape 1 : Calculer les coefficients d'auto-régression de ∇_M^d .

Etape 2 : Calculer $[b_t]$ en utilisant l'équation 32 pour $t = N + Q, \dots, 1$. Au départ, $[b_t] = 0$.

Etape 3 : Exécuter la rétro-prévision de la série $[b_t]$ en utilisant l'équation 32. Ceci peut être fait en utilisant l'algorithme SARMAS de McLEOD et HOLANDA SALES (1983).

Etape 4 : Exécuter la rétro-prévision de la série des $[x_t]$ en utilisant l'équation 31.

Etape 5 : Calculer la valeur de $[c_t]$ pour les séries $t = 1 - Q, \dots, N$ en utilisant l'équation 28.

Etape 6 : Calculer la valeur de $[a_t]$ pour les séries $t = 1 - Q, \dots, N$ en utilisant l'équation 29.

Etape 7 : Calculer S en utilisant l'équation 25.

Les étapes 1 à 7 peuvent être répétées jusqu'à ce qu'on ait atteint une certaine limite de tolérance qui aura été établie au début. Les paramètres sont obtenus en minimisant S, donné à l'équation 25. On a trouvé que l'algorithme de minimisation présenté par POWELL (1964) peut être utile pour minimiser S.

RÉFÉRENCES BIBLIOGRAPHIQUES

- AKAIKE H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control.*, AC-19 : 716-723.
- BARTLETT M.S., 1946. On the theoretical specification of sampling properties of auto-correlated time series. *J. Roy. Statist. Soc. Suppl.*, 8 : 27-41.
- BALLERINI R. et BOES D.C., 1985. Hurst behavior of shifting level processes. *Water Resour. Res.*, 21 : 1642-1648.
- BOES D.C. et SALAS J.D., 1978. Nonstationarity of the mean and the Hurst phenomenon. *Water Resour. Res.*, 14 : 135-145.
- BOX G.E.P. et JENKINS G.M., 1976. *Time Series Analysis : Forecasting and Control*. Holden-Day, San Francisco, California (Second Edition).
- BRASS R. et RODIRGUEZ-ITURBE I., 1984. *Random Functions and Hydrology*. Addison-Wesley, pp. 550.
- BRILLINGER D., 1976. *Time series : Data Analysis and Theory*. Holden-Day. San Francisco, California (Second Edition).
- COVER K.A. et UNNY T.E., (1986). Allocation of computer intensive statistics to parameter uncertainty in streamflow synthesis. *Water Resour. Bull.*, 22.
- COX D.R., (1984). Long range dependence : a review. In *Statistics : An Appraisal*, H.A. David et H.T. David (editors). *Proceedings of the 50th Anniversary Conference*, Iowa State Statistical Laboratory, Iowa State University Press, Ames, Iowa, pp. 55-74.
- CRYER J.D. et LEDOLTER J., (1981). Small-sample properties of the maximum likelihood estimator in the first-order moving average model. *Biometrika*, 68 : 691-694.
- DURBIN J., 1960. The fitting of time-series models. *Rev. Inst. Int. Stat.*, 28 : 233-244.
- EFRON B., 1979. Bootstrap methods : another look at the jackknife. *Ann. Stat.*, 7 : 1-26.
- FREEMAN D., 1981. Bootstrapping regression models. *Ann. Stat.*, 9 : 1218-1229.
- GRANGER C.W.J., 1980. Long-memory relationships and the aggregation of dynamic models. *J. Econ.* 14 : 227-238.
- GRANGER C.W.J. et JOYEUX R., 1980. An introduction to long memory time series models and fractional differencing. *J. Time Series Analysis* 1 : 15-29.
- FRYER M.J., 1977. A review of some non-parametric methods of density estimation. *J. Inst. Maths. Applics.*, 20 : 335-354.
- HANNAN E.J., 1970. *Multiple Time Series*. John Wiley, New York.
- HIPEL K.W., 1981. Geophysical model discrimination using the Akaike information criterion. *IEEE Trans. Autom. Control* AC-26(2) : 358-378.
- HIPEL K.W. et McLEOD A.I., 1990. *Time Series Modelling for Water Resources and Environmental Engineers*. Elsevier, Amsterdam (in press).
- HOSKING J.R.M., 1981. Fractional differencing. *Biometrika* 68 : 165-176.
- HOSKING J.R.M., 1984. Modeling persistence in hydrologic time series using fractional differencing. *Water Resour. Res.* 20 : 1898-1908.
- HOSKING J.R.M., 1985. Fractional differencing modeling in hydrology. *Water Resour. Res.* 21 : 677-682.
- HURST H.E., 1951. Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.*, 116 : 770-808.
- KLEMES V., 1974. The Hurst phenomenon - A puzzle ? *Water Resour. Res.* 10 : 675-688.
- KOTTEGODA N.T., 1980. *Stochastic water resources technology*. MacMillan Press, London, England.
- LI W.K. et McLEOD A.I., 1986. Fractional time series modelling. *Biometrika* 73 : 165-176.
- MANDELBROT B.B. et WALLIS J.R., 1969a. Computer experiments with fractional Gaussian noises, 1, Averages and variances. *Water Resour. Res.* 5 : 228-241.
- MANDELBROT B.B. et WALLIS J.R., 1969b. Computer experiments with fractional Gaussian noises, 2, Rescaled ranges and spectra. *Water Resour. Res.*, 5 : 242-259.
- MANDELBROT B.B. et WALLIS J.R., 1969c. Computer experiments with fractional Gaussian noises, 2, Mathematical appendix. *Water Resour. Res.*, 5 : 260-267.
- McLEOD A.I. et HIPEL K.W., 1978a. Preservation of the rescaled adjusted range. *Water Resour. Res.* 14 : 491-508.

- McLEOD A.I. et HIPEL K.W., 1978b. Simulation procedures for Box-Jenkins models. *Water Resour. Research* 14 : 969-975.
- McLEOD A.I., HIPEL K.W. et LENNOX W.C., 1977. Advances in Box-Jenkins modeling, 2, applications. *Water Resour. Res.* 13 : 511-586.
- McLEOD A.I. et SALES P.R., 1983. An algorithm for the approximate likelihood calculations of ARMA and seasonal ARMA models. *Appl. Statistics* 32 : 211-223.
- MEJIA J.M., RODRIGUEZ-ITURBE I. et DAWDY D.R., 1972. Streamflow simulation, 2, The broken line process as a potential model for hydrologic simulation, *Water Resour. Res.* 8 : 931-941.
- PARZEN E., 1982. ARMA models for time series analysis and forecasting. *J. Forecasting* 1 : 67-82.
- POWELL M.J.D., 1964. An efficient algorithm for finding the minimum of a function of several variables without calculating derivatives. *Computer Journal* 7 : 155-162.
- ROSENBLATT M., 1961. Independence and dependence. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, 431-443.
- ROSENBLATT M., 1979. Some limit theorems for partial sums of quadratic forms in stationary Gaussian variables. *Z. Wahrscheinlichkeitstheorie und verwandte Gebiete* 49 : 125-132.
- ROSENBLATT M., 1981. Limit theorems for Fourier transforms of functional of Gaussian sequences. *Z. Wahrscheinlichkeitstheorie und verwandte Gebiete* 55 : 123-132.
- SALAS J.D., BOES D.C., YEVJEVICH V. et PEGRAM G.G.S., 1979. Hurst phenomenon as a pre-asymptotic behaviour. *J. Hydrol.* 44 : 1-15.
- SALAS J.D., DELLEUR J.W., YEVJEVICH V. et LANE W.L., 1980. *Applied Modelling of Hydrologic Time Series*. Water Resources Publications, Littleton, Colorado.
- TRENCH W.F., 1964. An algorithm for the inversion of finite Toeplitz matrices. *S.I.A.M. J. Appl. Math.* 12 : 515-521.
- TAQQU M., 1975. Weak convergence to fractional Brownian motion and the Rosenblatt process. *Z. Wahrscheinlichkeitstheorie und verwandte Gebiete* 31 : 287-302.