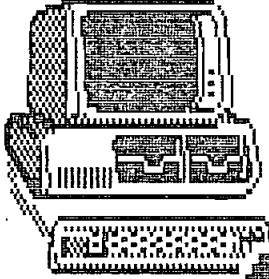




Projet de fin d'études

Ecole polytechnique de Thies

Département de Génie Civil

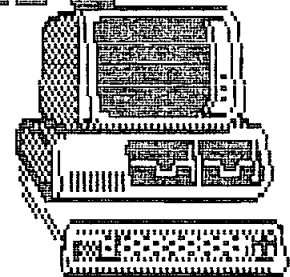


GC. 0296

Séries

Chronologiques

Méthode de
BOX-JENKINS



Abdou

ANZ

GUÉPE

Jun 1987

Proj
5.11

Analyse des séries chronologiques
par la méthode de **BOX-JENKINS**

A mes parents,

A tous mes frères et soeurs,

*A tous ceux qui ont contribué
à faire de moi ce que je suis.*

Fait par: **Abdou Aziz GUEWE**

Ecole Polytechnique de Thiès

Année Scolaire 1986-87

REMERCIEMENTS

Je tiens à remercier, ici, Mr. Jean Claude WARMAOES, professeur à l'Ecole Polytechnique qui, malgré ses lourdes charges à la tête du département de génie mécanique n'a jamais cessé de nous soutenir et de nous encourager tout au long du travail, pas toujours facile que nous avons eu à mener.

Je confonds dans le même hommage Mr Samba DIALLO, professeur de Route et transport au département de génie civil de l'Ecole Polytechnique de Thiès, dont l'expérience pratique dans le domaine nous a été d'un grand secours.

Que tout le personnel du centre de calcul trouve ici aussi l'expression de notre profonde gratitude pour le soutien actif qu'il n'a cessé de nous apporter.

Je remercie, enfin tout ceux qui, d'une façon ou d'une autre, ont fait de ce projet ce qu'il est aujourd'hui.

TABLE DES MATIERES

INTRODUCTION..... 1

CHAPITRE 1

PRESENTATION DE LA METHODE..... 4

1.2 Cheminement de la méthode..... 6

CHAPITRE 2

IDENTIFICATION 9

2.1 Processus Autorégressifs AR 10

2.2 Processus de moyenne mobile MA 11

2.3 Processus mixte ARMA 12

2.4 Méthodologie d'identification 13

2.5 Méthode 15

2.6 Exemples de processus 21

CHAPITRE 3

ESTIMATION 30

3.1 Estimé des paramètres..... 31

3.2	Qualité des coefficients.....	35
3.3	Redondance paramétrique.....	36

CHAPITRE 4

	DIAGNOSTIC CHECKING.....	37
4.1	Qualité des coefficients.....	37
4.1.1	Stationnarité.....	38
4.1.2	inversibilité.....	39
4.2	Autocorrélation des résidus.....	41
4.3	Test du portemanteau.....	42
4.4	Reformulation du modèle.....	45

CHAPITRE 5

	PREVISIONS.....	46
5.1	Intervalles de confiance.....	48
	CONCLUSION.....	49
	ANNEXES.....	51

INTRODUCTION

La statistique mathématique moderne peut se définir comme un ensemble de méthodes pour prendre des décisions raisonnables en présence d'incertitudes.

Cette réorientation de la statistique, opérationnelle depuis le début du siècle, sous la houlette des théoriciens anglo-saxons dont les chefs de file furent K. PEARSON et FISHER se veut totalement différente de l'idée qui consistait à considérer la statistique comme la science du dénombrement. Cette nouvelle approche, sous-tendue par le calcul des probabilités, est basée sur une méthode inductive, i.e., utiliser le raisonnement mathématique pour essayer de trouver le mécanisme générateur d'une série statistique donnée.

La démarche de la statistique, à partir de ce moment s'articule en trois phases qui peuvent être schématisées comme suit:

- La **description** qui consiste à effectuer une mise en ordre de la série considérée en vue de réduire le volume de données à une dimension plus maniable.

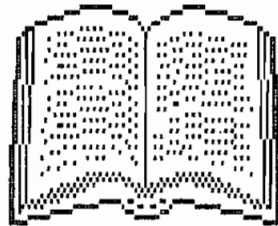
L'information contenue dans la série est alors exprimée sous forme condensée à l'aide de quelques graphiques ou valeurs types (moyenne, écart type, techniques multidimensionnelles...).

- L'analyse des résultats s'en suit. Il s'agit d'une tentative de formalisation de l'information contenue dans la série et exprimée lors de l'étape précédente par les graphiques ou valeurs types. Cette formalisation conduit à l'établissement d'une formule mathématique rendant compte du comportement passé de la série et permettant d'en faire une projection sur le futur.
- La **prévision** but ultime de la presque totalité des processus de modélisation consiste à projeter le modèle obtenu dans futur le pour organiser cet avenir et ainsi prendre des décisions.

Dans la grande majorité des modèles statistiques, il est couramment admis que les observations varient indépendamment les unes des autres. A tel point que, dans plusieurs applications, une interdépendance de ces observations, si petite soit elle, est considérée comme nuisible à la qualité du modèle. Mais une telle situation relève quand même de l'idéal car dans la réalité de tous les jours, les données, quelque soit leur nature ou leur processus d'acquisition, sont toujours liées entre elles. Essayer de quantifier cette dépendance temporelle existant à l'intérieur d'une série de données est le but que se fixe la méthode de BOX et JENKINS dont il est question dans ce projet. Plus même, elle fait de l'interdépendance existant à l'intérieur d'une série chronologique le soubassement de la démarche de construction de modèles stochastiques ainsi que de leur utilisation. Nous allons dans ce qui suit essayer de présenter succinctement cette méthode dans le but non seulement d'en révéler toute la puissance, mais aussi de montrer tout l'attrait qu'elle devrait avoir pour tous les corps de métiers en général et pour les ingénieurs de conception en particulier.

Chapitre

1



PRESENTATION

PRESENTATION

DE

LA METHODE

Avant l'apparition de la méthode de BOX-JENKINS, le type de prévision le plus souvent rencontré dans la pratique statistique consistait à admettre l'existence d'une loi fondamentale indépendante de la série, représentée par les données disponibles avec un certain caractère aléatoire. L'activité principale de la méthode consistait donc à isoler, le mieux possible, les caractéristiques statistiques de ladite loi afin de l'utiliser comme base des projections dans le futur.

Autrement dit, l'analyse statistique, jusqu'à l'apparition des méthodes telles que celle de BOX et JENKINS, se résumait à une tentative d'ajustement des observations à un modèle préétabli par le calcul de certains paramètres. Ces méthodes, applicables, quand série contenait peu de fluctuations, atteignaient

rapidement leurs limites quand la vraie loi de variation de la série ne suivait plus, pour différentes raisons, le cheminement qu'on voulait lui imposer. Quand on sait qu'une série est en général soumise à la combinaison de trois types de variation:

- la tendance de la série
- la composante cyclique
- les variations aléatoires

on se rend compte que jusqu'ici le travail de l'analyste des données n'était pas de tout repos.

La méthode de BOX-JENKINS a voulu s'affranchir de cette contrainte et se propose, au lieu de supposer que la série suit une loi de comportement préalablement établie, de déterminer qu'elle est, pour nous, la meilleure loi de comportement de la série.

Cette nouvelle approche de la prévision fait de la méthode développée par les professeurs BOX et JENKINS une véritable philosophie car, en plus de supprimer la nécessité de faire au départ l'hypothèse d'une loi de comportement de la série, elle offre l'opportunité de juger si une loi de variation est

satisfaisante et même ~~dans~~ quelle est la précision qu'on pourrait en attendre. Même si elle ne l'est pas, elle donne des indications supplémentaires permettant d'identifier la loi correcte.

1.2 Cheminement de la méthode

La démarche suivie par la méthode de BOX-JENKINS est celle en trois étapes de toutes les méthodes d'analyse prévisionnelle. A ce titre il peut être schématisé par le diagramme suivant

Etape 1 Identification

Choix d'un ou de plusieurs
modeles susceptibles d'être
mecanismes generateurs

Etape 2 Estimation

Estimer les parametres
des modeles choisis à
l'etape 1

Etape 3 Test
(Diagnostic chee-
king)

Verifier l'adequation
du modele

Faire des
precisions

Oui

Le modele
est-il satisfaisant

Non

Figure 1 : Organigramme de fonctionnement de
méthode de BOX/et JENKINS

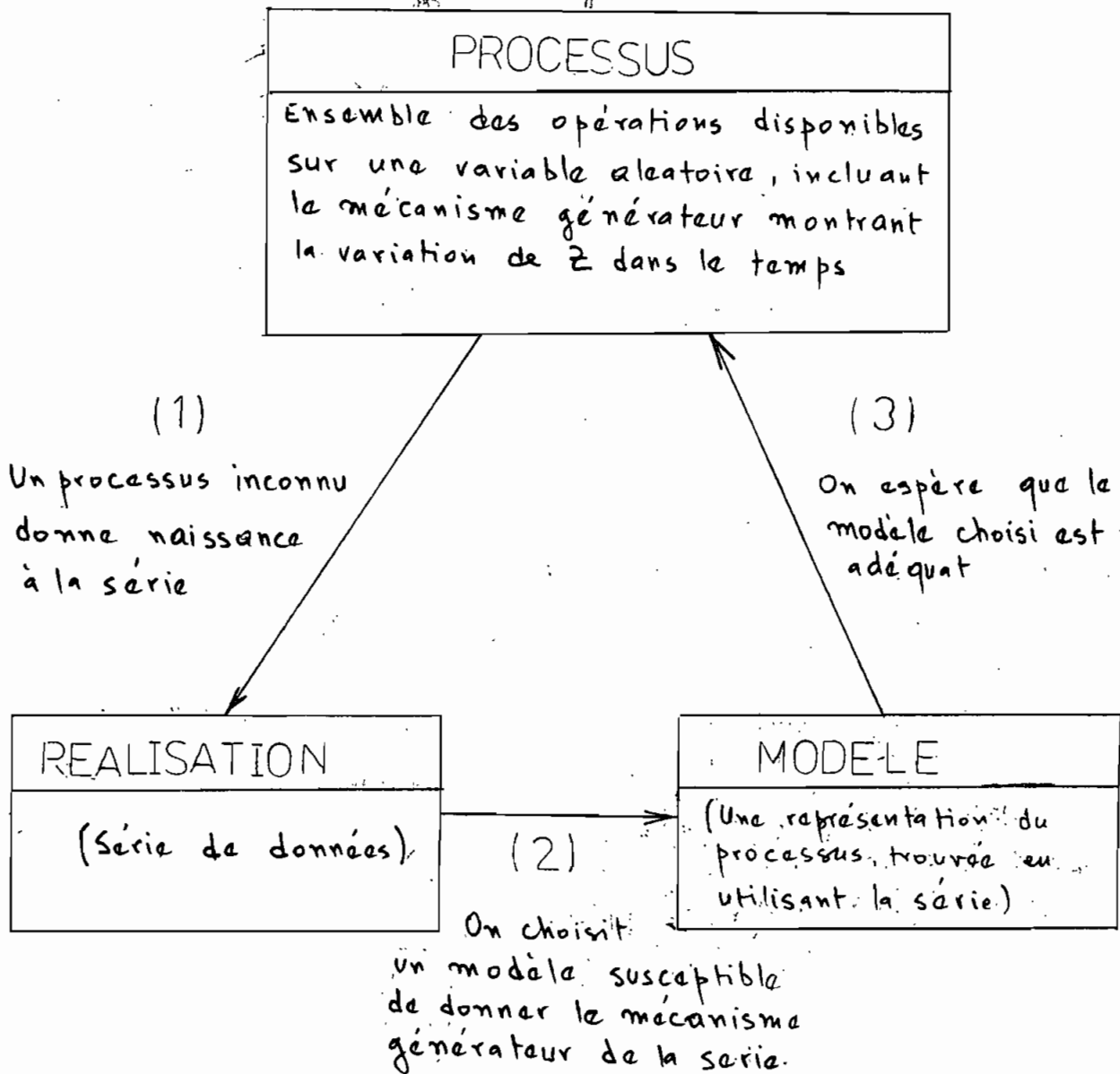
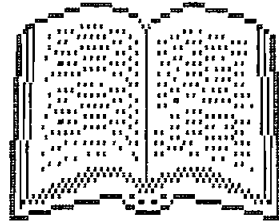


Figure 2 : Processus en trois étapes de la méthode de
BOX et JENKINS

Chapitre 2



INFORMATION

I D E N T I F I C A T I O N

L'identification est, par définition, le processus par le quel on reconnaît la classe d'appartenance d'un modèle. Elle repose sur le principe suivant:

Les données, dans une série chronologique, sont liées entre elles. La forme et la force de cette interdépendance sont indiquées par les deux caractéristiques statistiques que sont: l'autocorrélation que nous abrègerons par acf (auto-correlation function) et l'autocorrélation partielle ou pacf (partial autocorrelation function).

L'étude du comportement de ces deux paramètres permet de classer chaque série chronologique dans l'une des trois catégories suivantes.

2.1 Processus autorégressifs (AR)

Un processus est dit auto régressif quand la valeur de la variable aléatoire Z à un instant t donné est une combinaison linéaire des p valeurs antérieures de cette même variable aléatoire. On dit alors qu'on a un processus autorégressif d'ordre p et on le note $AR(p)$. L'équation générale d'un processus autorégressif d'ordre p s'écrit :

$$z_t = C + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \phi_3 z_{t-3} + \dots + \phi_p z_{t-p} + a_t$$

Cette équation se lira de la façon suivante:

La valeur de la variable aléatoire Z à l'instant t est significativement liée aux p valeurs précédentes de cette même variable aléatoire.

Exemples de processus AR()

$$z_t = C + \phi_1 z_{t-1} + a_t \text{-----} AR(1)$$

$$z_t = C + \phi_2 z_{t-2} + \phi_1 z_{t-1} + a_t \text{-----} AR(2)$$

$$z_t = C + \phi_3 z_{t-3} + \phi_2 z_{t-2} + \phi_1 z_{t-1} + a_t \text{-----} AR(3)$$

Note: On reconnaît dans le premier exemple un processus Markovien ce qui a fait dire à certains auteurs que la méthode de BOX et JENKINS est une nouvelle façon de considérer les processus stochastiques.

2.2 Les Processus de moyennes mobiles (MA)

Un processus est dit de moyenne mobile si la valeur de la variable aléatoire Z à un instant donné est une combinaison linéaire des valeurs des valeurs de l'erreur de régression jusqu'à l'ordre q . On dit que ce processus est une moyenne mobile d'ordre q et on le note $MA(q)$.

L'équation générale d'un processus de moyenne mobile peut alors s'écrire de la forme suivante:

$$z_t = C - \theta_1 z_{t-1} - \theta_2 z_{t-2} - \theta_3 z_{t-3} + \dots - \theta_q z_{t-q} + a_t$$

Exemples de processus AR()

$$z_t = C + \theta_1 a_{t-1} + a_t \text{----- MA(1)}$$

$$z_t = C + \theta_2 a_{t-2} + \theta_1 a_{t-1} + a_t \text{----- MA(2)}$$

$$z_t = C + \theta_3 a_{t-3} + \theta_2 a_{t-2} + \theta_1 a_{t-1} + a_t \text{----- MA(3)}$$

Note: Le signe (-) devant les paramètres θ est une convention d'écriture

- Le paramètre (C) est appelée tendance centrale de la série et une fonction de la moyenne de la série.
- L'appellation moyenne mobile est abusive car la somme des différents coefficients θ n'est pas égale à 1. Malgré cet inconvénient, elle est couramment utilisé dans la littérature.

2.3 Processus mixte

C'est, comme l'indique assez clairement son nom, un processus ayant les caractéristiques des deux pro-

cessus sus-citées. C'est le processus le plus général et le plus couramment rencontré en pratique, mais il est beaucoup moins utilisé que les précédents car demandant un effort de calcul nettement plus important donc moins économique. On le note ARMA(p,q) où p est l'ordre de la composante autorégressive du modèle et q est l'ordre de la composante de moyenne mobile.

2.4 Méthodologie de l'identification

Avant de présenter la démarche à suivre pour reconnaître un modèle, nous allons présenter quelques opérations sur les séries chronologiques:

* Le centrage

Il a pour but de supprimer le paramètre C des équations générale des modèles. Il est défini par l'équation suivante:

$$\tilde{z}_t = z_t - \mu$$

avec $\mu = (Ez_t)/n$ est l'espérance mathématique de la variable aléatoire Z

* La différenciation

Elle permet une expression plus concise, donc plus facile à manipuler, des modèles. Elle se présente sous la forme suivante:

$$Bz_t = z_{t-1}$$

$$B^k z_t = z_{t-k}$$

L'opérateur (B) appelé ainsi à cause de la terminologie anglo-saxonne (Backshift operator) se comporte comme l'opérateur différentiel (d).

En utilisant cette double notation nous pouvons réécrire les processus précédemment mentionnés de la manière suivante:

$$\text{AR} : \tilde{Q}(B) \tilde{z}_t = a_t$$

$$\text{MA} : \tilde{z}_t = \Theta(B) a_t$$

$$\text{ARMA} : \tilde{Q}(B) \tilde{z}_t = \Theta(B) a_t$$

où

$\tilde{Q}(B) = 1 - \tilde{Q}_1 B - \tilde{Q}_2 B^2 - \tilde{Q}_3 B^3 - \dots - \tilde{Q}_p B^p$ est l'opérateur autorégressif

$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \dots - \theta_p B^p$ est l'opérateur de moyenne mobile

2.5 Méthodologie

Comme mentionné au précédent chapitre, l'identification consiste essentiellement en l'étude de la loi de variation des deux caractéristiques statistiques que sont la fonction d'autocorrélation (acf) et la fonction d'autocorrélation partielle (pacf).

Comme toute autre type de corrélation, la fonction d'autocorrélation exprime la direction et la force d'un lien unissant deux variables aléatoires. L'autocorrélation ne se distingue de cette définition que par le fait qu'elle exprime la corrélation existant entre deux paramètres de la même variable aléatoire d'où le préfixe auto-. Comme fonction de corrélation ses valeurs possibles vont de -1 à 1 avec:

- 1 pour une corrélation parfaitement négative
- +1 pour une corrélation parfaitement positive
- 0 quand les deux variable aléatoire son totale-
ment indépendantes

La formule la plus utilisée pour calculer l'autocorrélation s'écrit:

$$r_k = \frac{\sum_{t=1}^{n-1} \tilde{z}_t \tilde{z}_{t+k}}{\sum_{t=1}^{n-k} (\tilde{z}_t)^2}$$

Cette équation est déduite d'un calcul de régression linéaire simple

Quant à l'autocorrélation partielle (pacf) elle exprime elle aussi la force et la direction de la relation d'interdépendance existant entre deux variables d'une même série chronologique mais, à la différence de son homologue (acf), elle permet en outre de tenir compte de l'effet des autres variables situées entre celles que l'on essaie de caractériser. C'est pourquoi elle est appelée autocorrélation avec mémoire. Pour calculer ses valeurs, il faudrait, idéalement procéder à un régression multiple. Mais YULE et WALKER ont développé des formules simples permettant, avec moins d'effort de programmation, d'en avoir des estimations très précises.

$$\bar{\phi}_{11} = r_1$$

$$r_k = \sum_{j=1}^{k-1} \bar{\phi}_{k-1, j} r_{k-j}$$

$$\bar{\phi}_{kk} = \frac{r_k - \sum_{j=1}^{k-1} \bar{\phi}_{k-1, j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \bar{\phi}_{k-1, j} r_j}$$

$$\bar{\phi}_{kj} = \bar{\phi}_{k-1, j} - \bar{\phi}_{kk} \bar{\phi}_{k-1, k-j}$$

Une fois ces deux caractéristiques calculées, on reconnaît les modèles grâce aux trois principes suivants:

- 1* Une loi de variation peut être approximée par un modèle autorégressif stationnaire si et seulement si sa fonction d'autocorrélation décroît, en valeur absolue, exponentiellement vers zéro tandis que sa fonction d'autocorrélation partielle est identiquement nulle au-delà du temps t-p. La valeur de p qui est aussi égale au nombre pics significatifs dans la représentation en histogramme des pacfs, est appelé ordre du processus. On note

alors ce processus $AR(p)$.

2* Une loi de variation peut être approximée par un modèle de moyenne mobile stationnaire si et seulement si sa fonction d'autocorrélation partielle décroît, en valeur absolue, exponentiellement vers zéro tandis que sa fonction d'autocorrélation est identiquement nulle au-delà du temps $t-q$. La valeur de q qui est aussi égale au nombre pics significatifs dans la représentation en histogramme des acfs, est appelé ordre du processus. On note alors ce processus $MA(q)$.

3* Une loi de variation peut être approximée par un modèle de processus mixte stationnaire si et seulement si fonction d'autocorrélation tout comme celle d'autocorrélation partielle décroissent rapidement vers zéro. L'ordre du processus est obtenu en prenant le temps (p) au-delà duquel les autocorrélations partielles peuvent être prises pour identiquement nulles et le temps (q) au-delà duquel les auto-

corrélations sont considérées comme toutes nulles.

Dans les trois propositions précédentes nous avons utilisée le terme "identiquement nulle". Ceci est à comprendre au sens statistique. Et pour savoir jusqu'à quel point une des valeurs des autocorrélations peut être prise pour non nulle; on procédera à un test de l'hypothèse de nullité H_0 .

Pour ce faire on calculera l'erreur commise sur l'estimation des autocorrélations. Pour un processus stationnaire dont les résidus sont normalement distribués BARTLETT a développé la formule suivante:

$$s(r_k) = (1 + 2 \sum_{j=1}^{k-1} r_j^2)^{1/2} * n^{-1/2}$$

Pour tester l'hypothèse H_0 nous utiliserons le T de STUDENT de la façon suivante:

$$t_{rk} = \frac{r_k}{s(r_k)}$$

La distribution du T de Student nous permet de

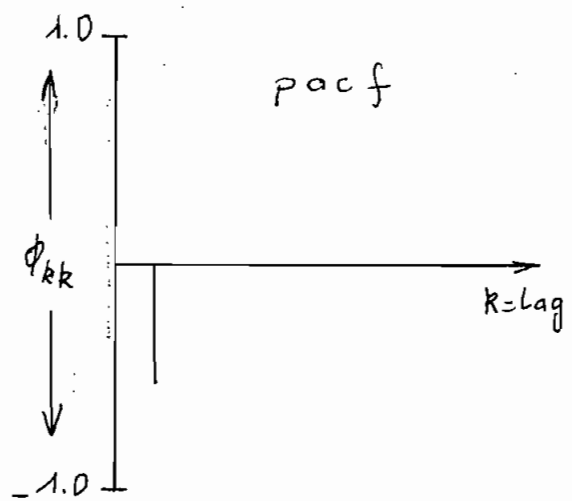
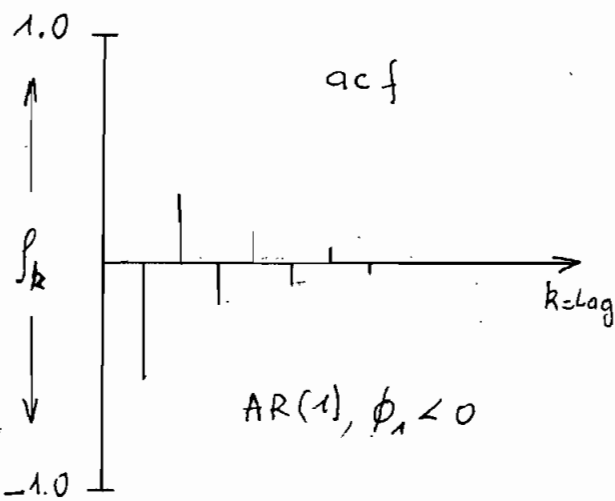
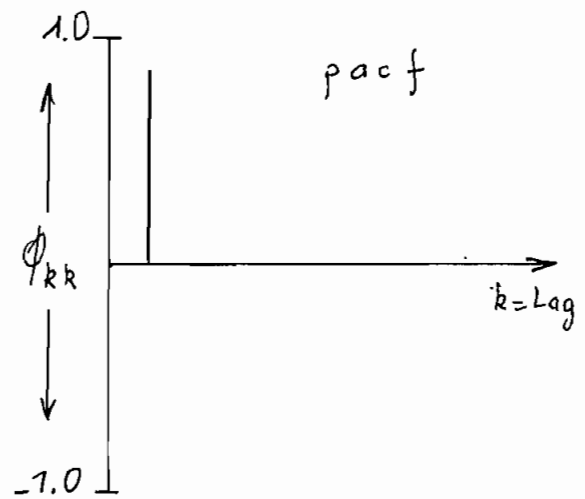
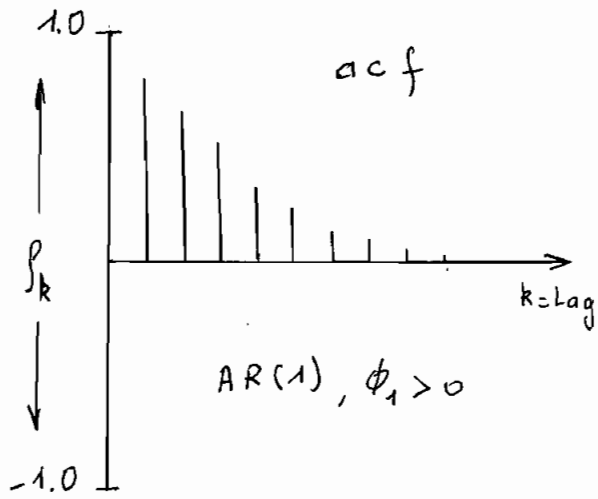
dire que:

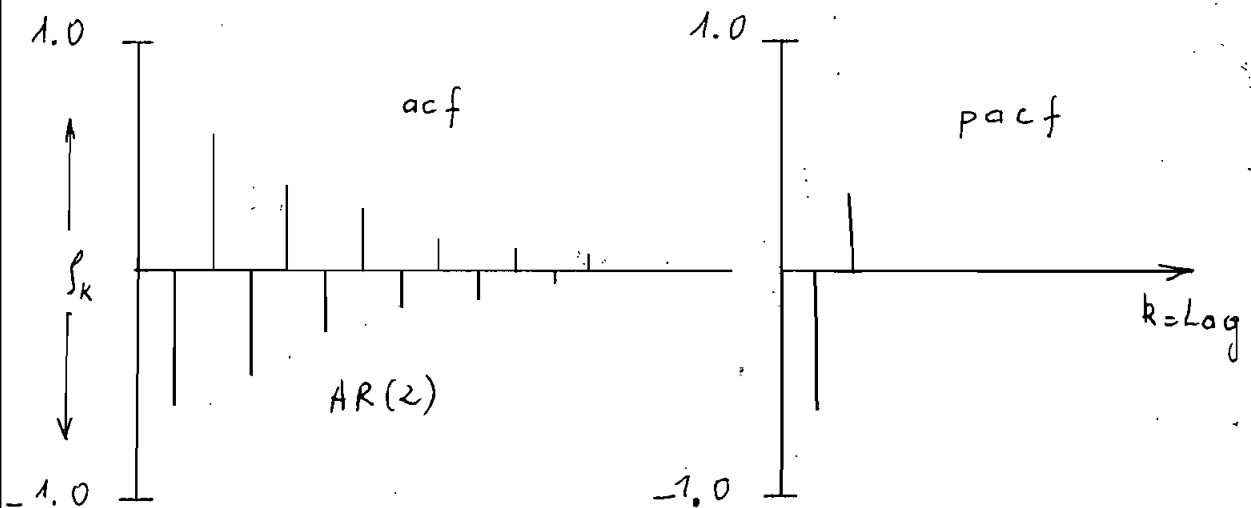
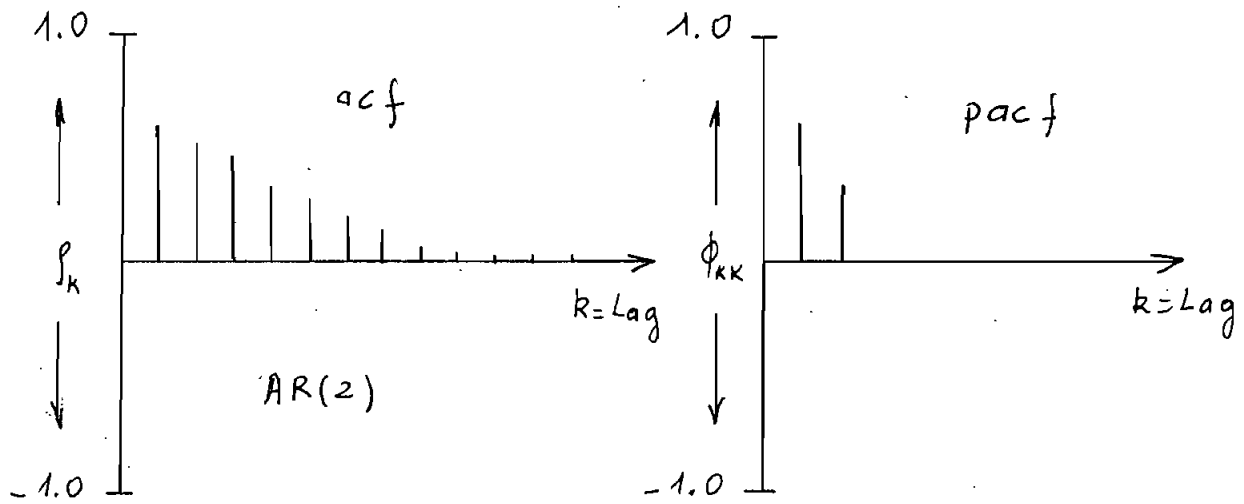
r_k est statistiquement différent avec 95% de
niveau de confiance si $t_{r_k} > 2.0$

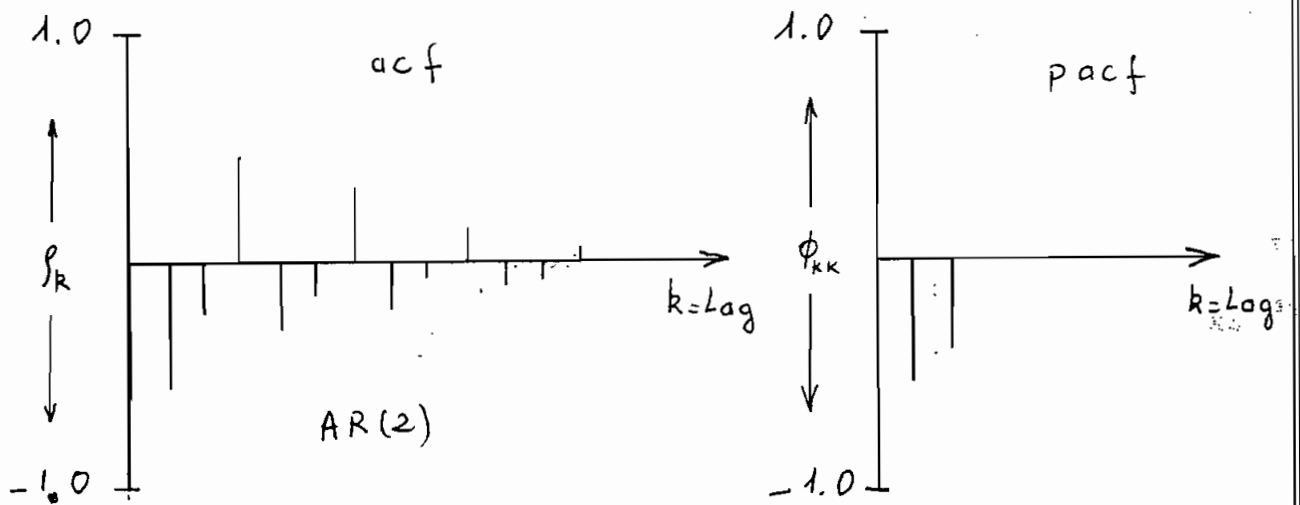
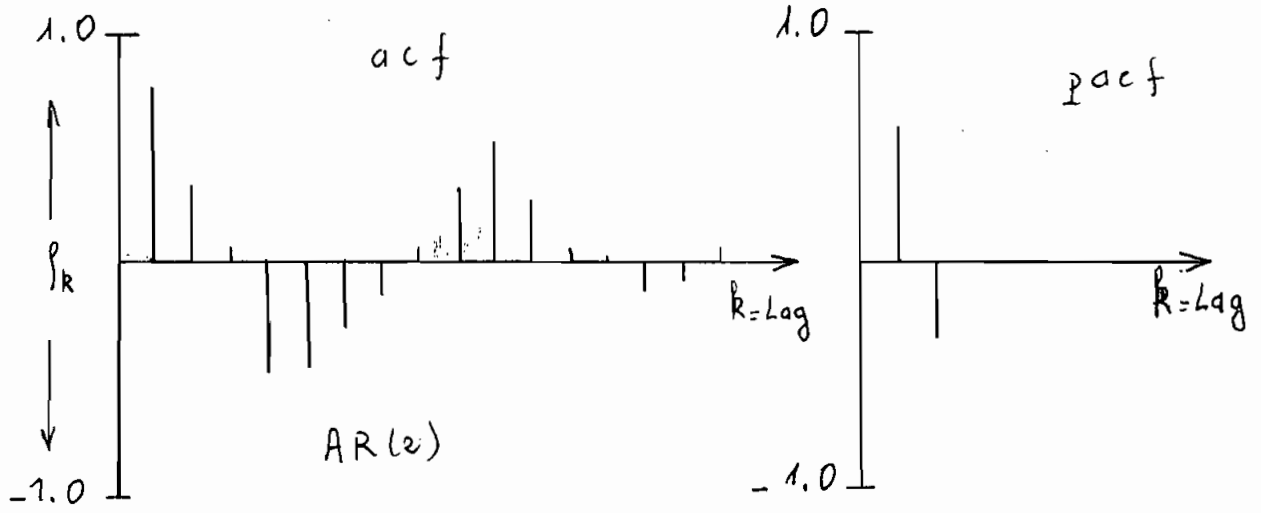
Pour les autocorrélations partielles BARTLETT
donne la formule suivante:

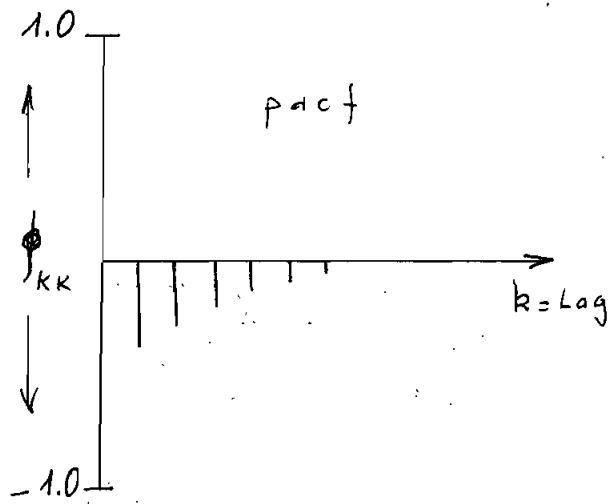
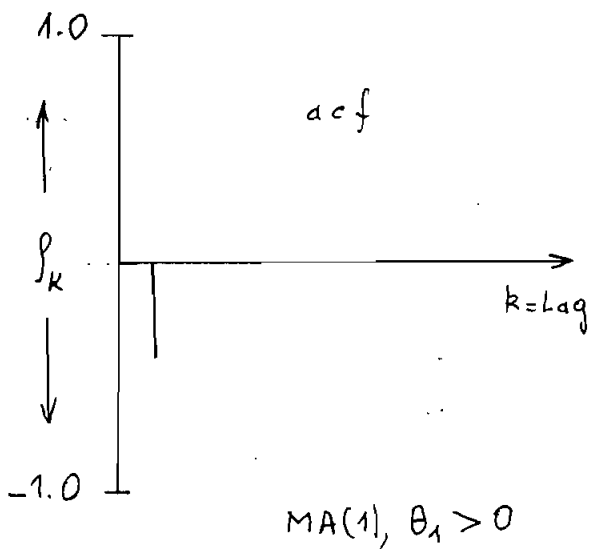
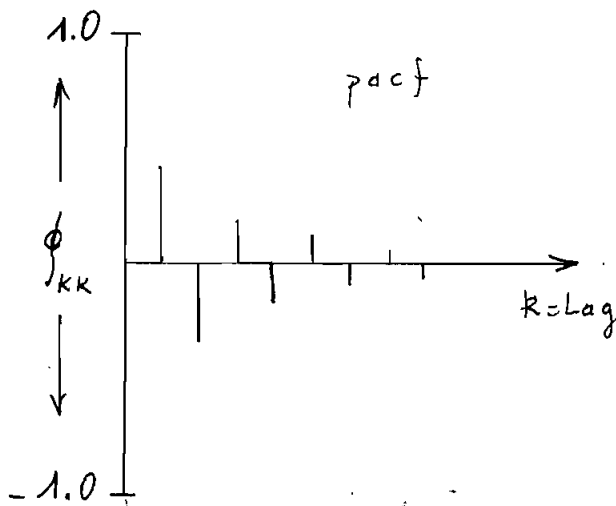
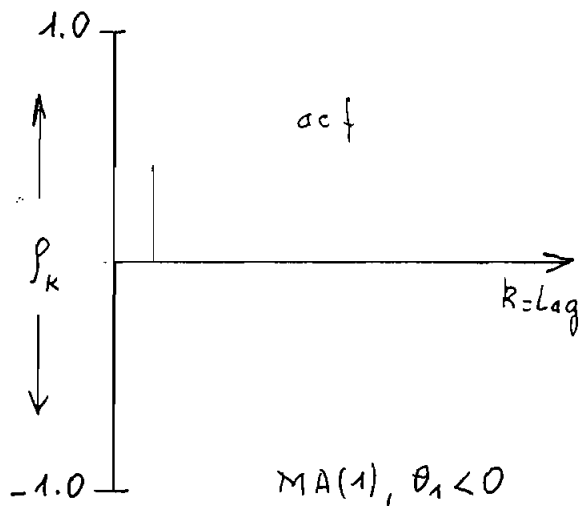
$$s(\hat{\rho}_{kk}) = n^{-1/2}$$

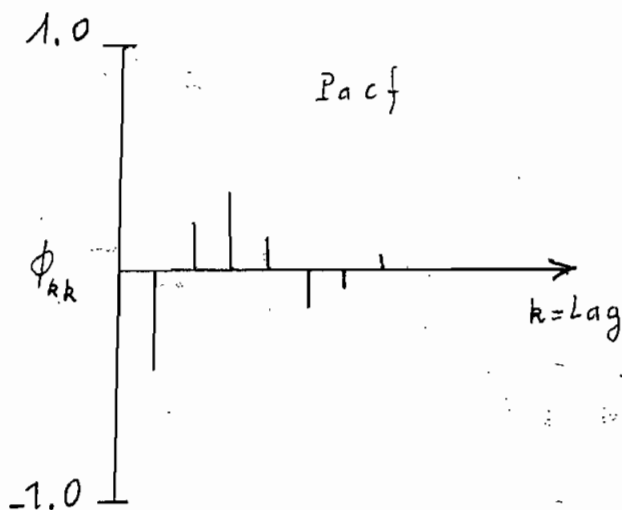
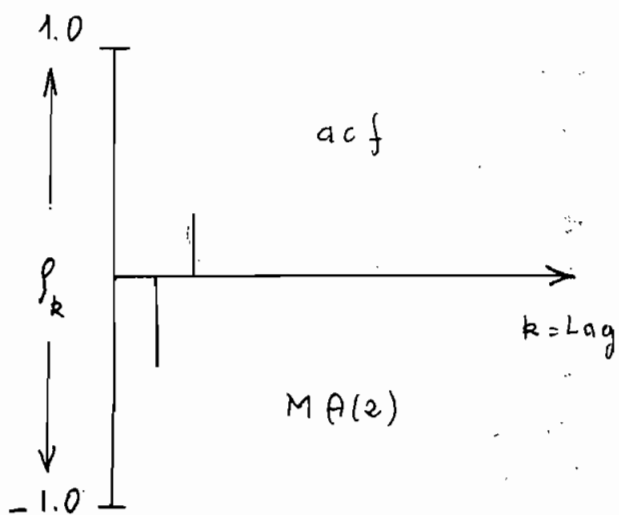
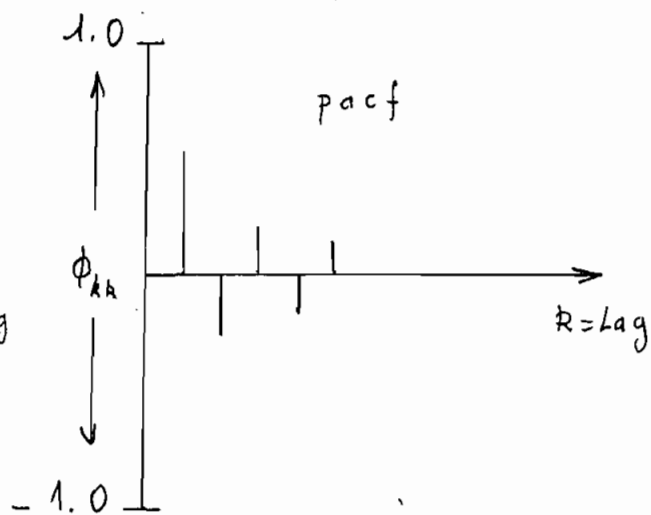
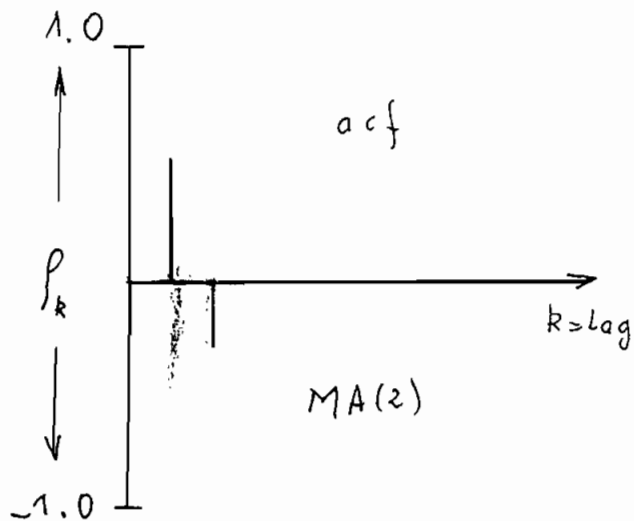
Le calcul de ces caractéristiques permet de
choisir un modèle d'essai après inspection visuelle
de la représentation graphique sous forme
d'histogramme des acfs et pacfs. Une fois un ou plu-
sieurs modèles retenus, on peut procéder à
l'estimation de ses paramètres.

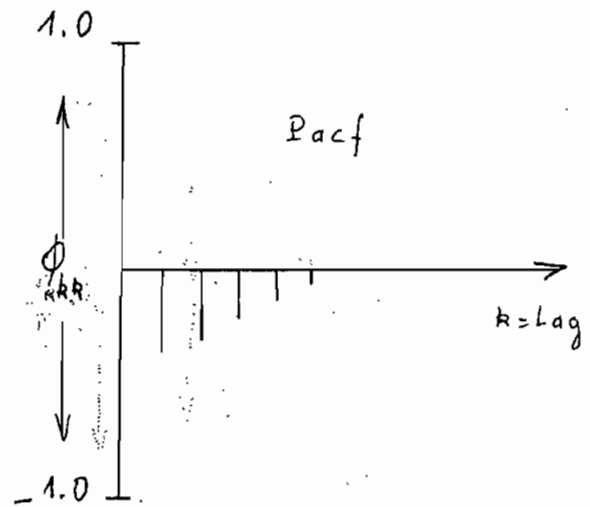
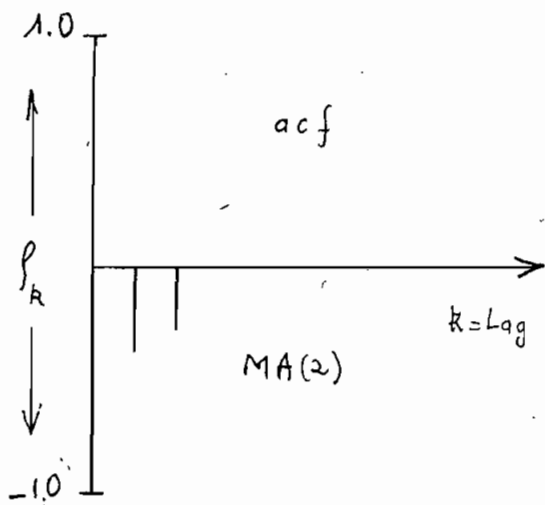
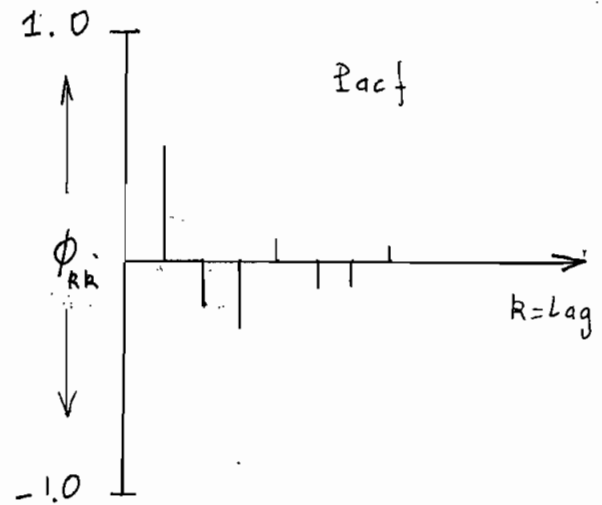
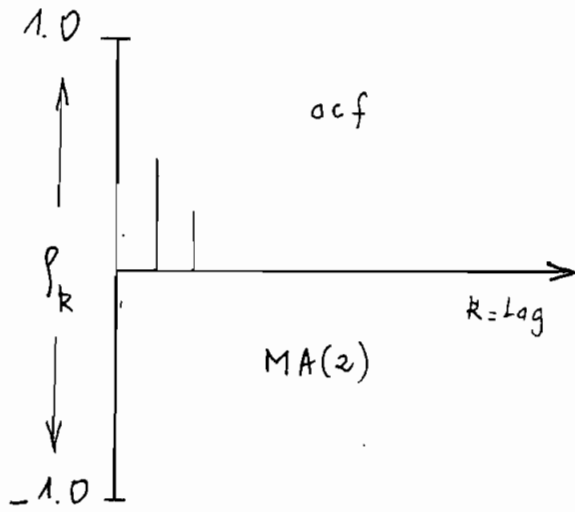


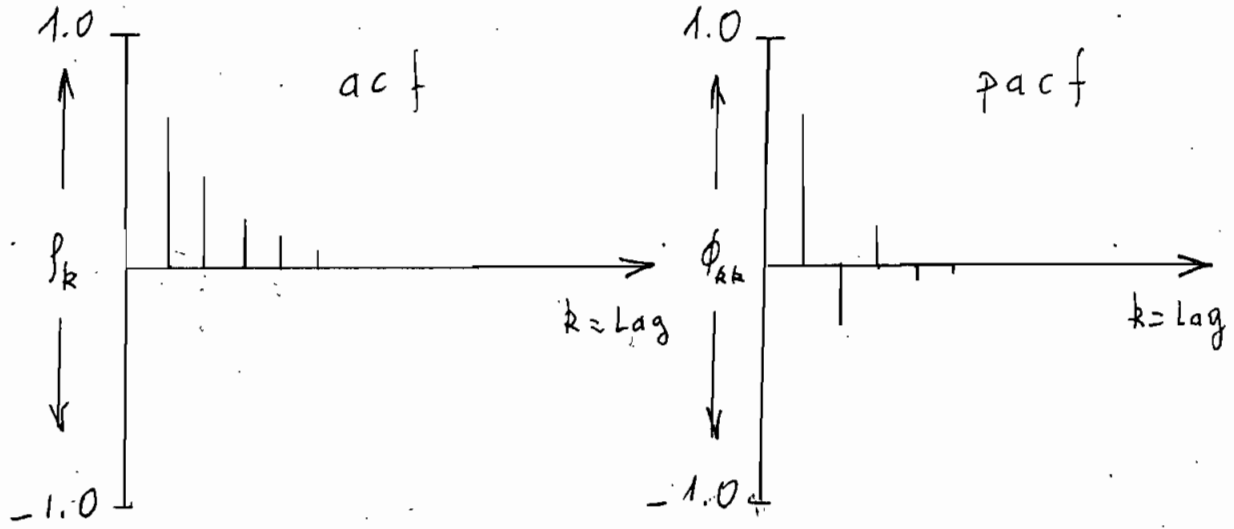




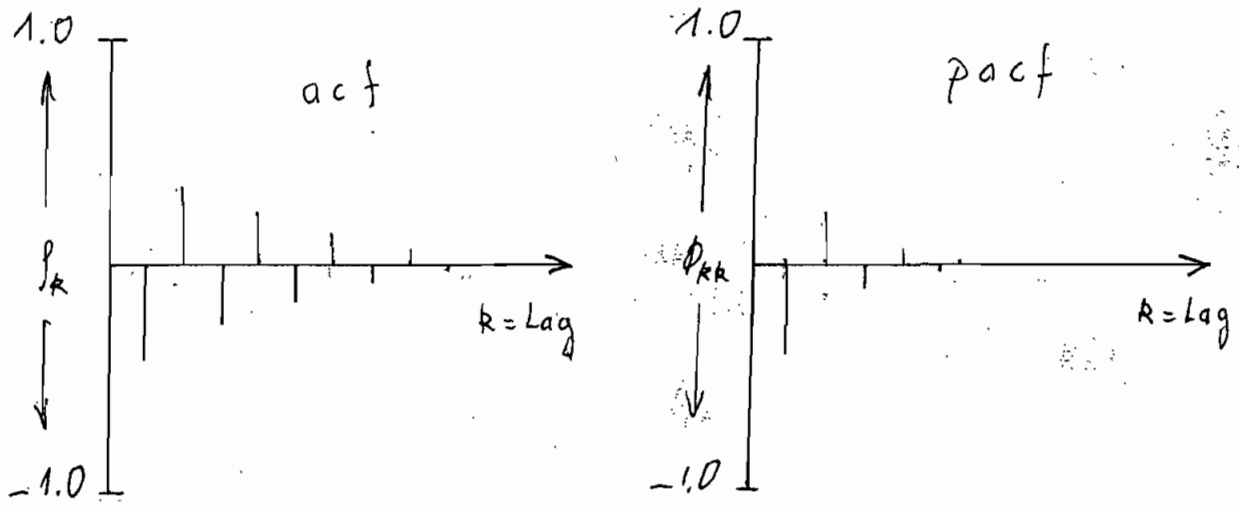




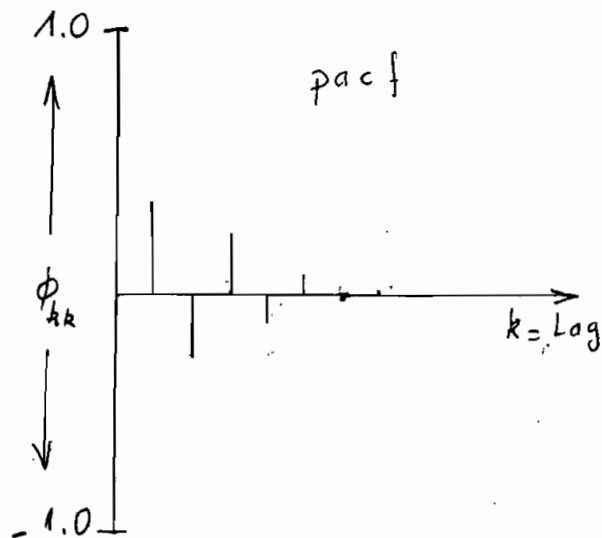
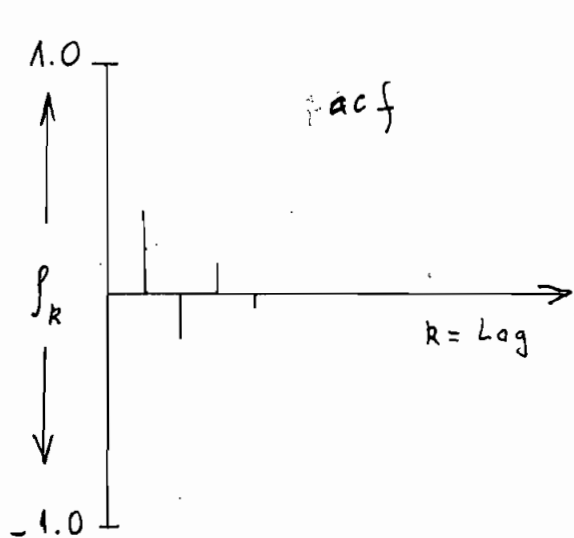




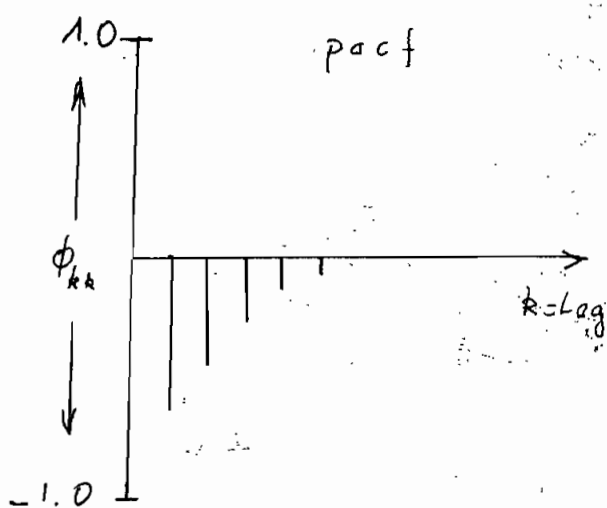
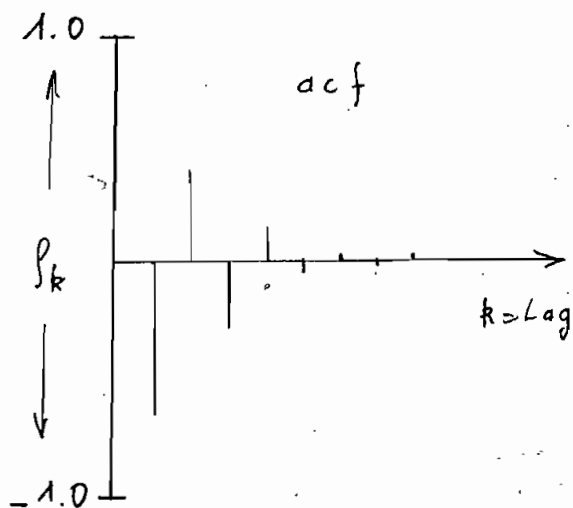
ARM(1,1)



ARMA(1,1)



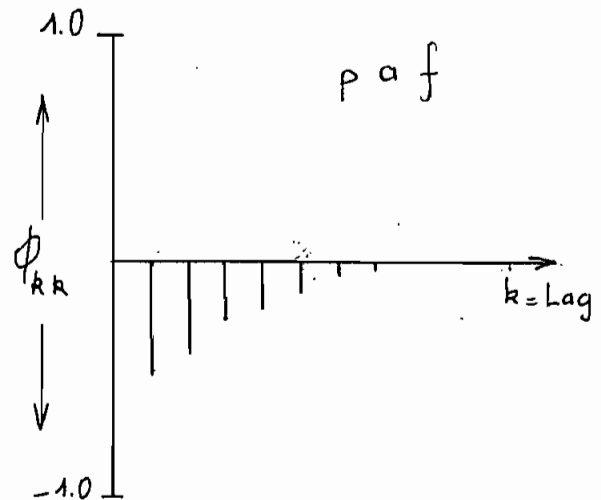
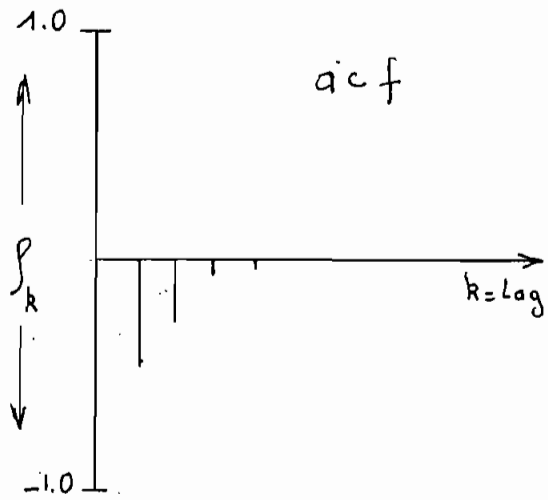
ARMA(1,1)



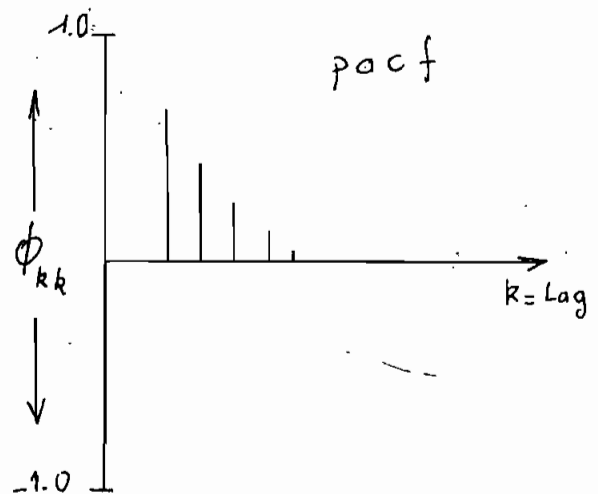
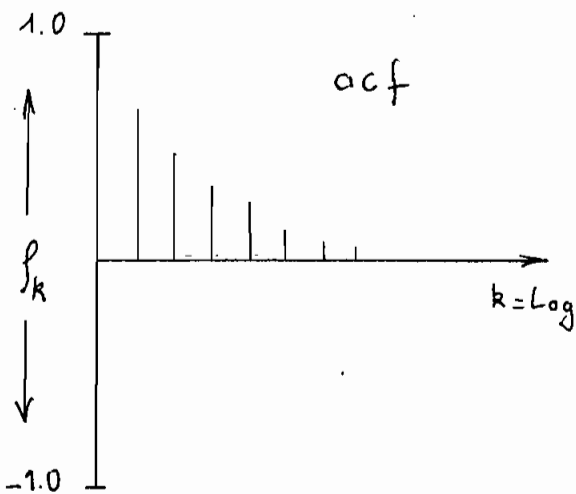
ARMA(1,1)

Proj
5.11

Analyse des séries chronologiques
par la méthode de BOX-JENKINS



ARMA(1,1)



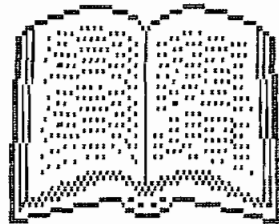
ARMA(1,1)

Fait par: ABDOUL ELIZ GUENIE
Ecole Polytechnique de Thiés

Année Scolaire 1986-87

Chapitre

3



ESTIMATION

ESTIMATION DES PARAMETRES

L'estimation des paramètres d'un modèle peut être vue comme étant un raffinement de l'analyse faite lors de l'étape d'identification. En effet, une fois les autocorrélations et autocorrélations partielles calculées et qu'un modèle de base (AR, MA, ou ARMA) a été choisi en suivant les principes exposés dans le chapitre précédent, l'estimation consiste à calculer les paramètres requis par le modèle en question et à discuter de leur qualité et de leur aptitude à modéliser la série donnée.

Il Existe plusieurs méthodes d'estimation de ces paramètres allant d'une simple analyse de regression (pour les processus autorégressifs purs) au lissage par la méthode des moindres carrés non linéaires.

La meilleure de ces méthodes à ce jour développée, est celle dite du compromis de MARQUARDT.

Compromis parce qu'alliant la résolution des systèmes équations non linéaires par la méthode de GAUSS-NEWTON au moindres carrés non linéaires. Dans sa démarche, le compromis de MARQUARDT part d'un ensemble de valeurs d'essais pour les coefficients tente d'en trouver de meilleurs, c'est à dire des coefficients minimisant la somme des carrés des résidus. Sa procédure systématique a pour résultat d'entraîner une convergence rapide vers les meilleures estimations possibles pour les paramètres recherchés.

Pour notre part nous avons préféré utiliser la méthode développée par YULE et WALKER pour estimer les paramètres du modèle pour des raisons d'ordre pratique, mais aussi guidé par un souci de cohérence entre toutes les étapes.

3.1 Estimer des paramètres

YULE et WALKER nous font remarquer que le calcul des paramètres d'un modèle général d'ARMA(p,q) peut

être fait en deux étapes successives:

- Déterminer les coefficients d'autorégression du modèle
- Réécrire le modèle sous la forme d'une moyenne mobile pure
- Déterminer les coefficients du modèle de moyenne mobile

En appliquant cette méthode on trouve pour les modèles d'ordre inférieur ou égal à 2 les valeurs de la page suivante.

La procédure de YULE-WALKER fournit en même temps les matrices de variance covariance qui permettent d'avoir, au moment de l'estimation, une idée sur non seulement la dispersion des paramètres estimées, mais aussi sur la degré de corrélation existant entre les paramètres que l'on vient d'estimer. Ceci nous permet, à ce stade, de procéder à une première vérification de l'adéquation entre le modèle que l'on suppose générateur de la série et la série .

Processus	Paramètres
AR(1)	$r_1 = \bar{\phi}_1$
MA(1)	$r_1 = \frac{-\theta_1}{1+\theta_1^2}$
AR(2)	$\bar{\phi}_1 = \frac{r_1(1-r_2)}{1-r_1^2}$ $\theta_2 = \frac{r_1 - r_2}{1-r_1^2}$
MA(2)	$r_1 = -\frac{\theta_1(1-\theta_2)}{1+\theta_1^2 + \theta_2^2}$ $r_2 = -\frac{\theta_2}{1+\theta_1^2 + \theta_2^2}$
ARMA(1,1)	$r_1 = \frac{(1-\theta_1\bar{\phi}_1)(\bar{\phi}_1 - \theta_1)}{1+\theta_1^2 - 2\bar{\phi}_1\theta_1}$ $r_2 = r_1\bar{\phi}_1$

Tableau 1: Valeurs des coefficients
pour les modèles les plus
courants

Les mêmes équations donnent aussi une estimation de l'erreur commise ainsi que de la corrélation entre les paramètres estimés. Ces résultats sont compilés dans le tableau suivant.

Modèles	Variance-corrélation
AR(1)	$V(\hat{\varrho}_1) = n^{-1} (1 - \hat{\varrho}_1^2)$
MA(1)	$V(\hat{\theta}_1) = n^{-1} (1 - \theta_1^2)$
MA(2)	$V(\hat{\theta}_1, \hat{\theta}_2) = n^{-1} (1 - \theta_2^2)$ $\alpha = -\frac{\theta_1}{1 - \theta_2}$
ARMA(1, 1)	$v(\hat{\varrho}) = \frac{(1 - \hat{\varrho}\theta)^2 (1 - \hat{\varrho}^2)}{(\hat{\varrho} - \theta^2)}$ $v(\hat{\theta}) = \frac{(1 - \hat{\varrho}\theta)^2 (1 - \theta^2)}{(\hat{\varrho} - \theta^2)}$ $\alpha = \frac{(1 - \hat{\varrho}\theta) (1 - \hat{\varrho}^2) (1 - \theta^2)}{(\hat{\varrho} - \theta^2)}$

Tableau 2: Variances et coefficients de corrélations des modèles courants

3.2 Qualité des coefficients

Tout comme lors de la précédente étape (IDENTIFICATION), la qualité statistique des coefficients estimés est à vérifier. Pour ce faire, on utilisera le même T-test de STUDENT pour évaluer l'hypothèse de nullité H_0 :

$$t_{\phi_k} = \frac{\hat{\phi}_k}{s(\hat{\phi}_k)}$$

$$t_{\theta_k} = \frac{\hat{\theta}_k}{s(\hat{\theta}_k)}$$

Avec $s(\hat{\phi}_k)$ et $s(\hat{\theta}_k)$ les racines carrés des variances données dans le tableau précédent.

Redondance paramétrique

On dit qu'on a une redondance paramétrique dans un modèle ARMA(p, q) s'écrivant :

$$\tilde{\phi}(B) \tilde{z}_t = \theta(B) a_t$$

où

$$\tilde{\phi}(B) = (1 - \alpha B) \phi'(B)$$

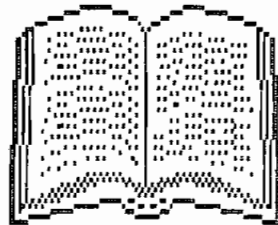
$$\theta(B) = (1 - \beta B) \theta'(B)$$

si $\alpha = \beta$.

En effet, après simplification, l'ordre du modèle diminue de 2 unités et ne pas relever cette particularité, serait d'essayer d'appliquer au modèle plus de paramètres qu'il n'en requiert, ce qui en plus de ne pas être économique conduit quelques fois à des singularités dans certaines matrices de travail. Ce serait le cas par exemple d'un processus ARMA(1, 1) où $\tilde{\phi} = \theta$. En effet, en ce moment, la variance du processus tend vers l'infini. Ce cas est rencontré dans la pratique lorsque les paramètres d'un modèle AR(1) ou MA(1) sont très petits et ont été assimilés à zéro lors d'un premier essai d'identification.

Chapitre

4



MANUSCRIPTIC

DIAGNOSTIC CHECKING

Une fois les paramètres du modèle expliqués dans le chapitre précédent, il s'agit dans celui-ci de procéder à la certification de ces paramètres pour se prononcer leur aptitude réelle à être utilisés pour établir des prévisions.

4.1 Qualité des coefficients

La première étape du diagnostic est la vérification de la qualité statistique des coefficients trouvés lors de l'étape d'identification du modèle. Pour ce faire, les paramètres trouvés doivent satisfaire plusieurs exigences qui ont pour nom :

- stationnarité
- inversibilité
- significativité

4.1.1 Stationnarité

L'objectif de ce critère est d'assurer une convergence de la procédure d'estimation de ces paramètres qui, est-il besoin de le rappeler, est une procédure itérative de résolution de systèmes d'équations non linéaires.

Pour vérifier si un processus est stationnaire ou non, nous procédons dans la pratique à :

- un inspection visuelle de la courbe représentative de la série considérée pour essayer de voir si on n'a pas une variation dans le temps de la moyenne auquel cas nous devons procéder à une différenciation de la série autant de fois que cela s'avérera nécessaire .
- un examen de la représentation des auto-corrélations pour voir si elles décroissent "rapidement" vers zéro. C'est à dire, pratiquement, si les valeurs absolues du T (T-val) sont inférieures à environ 1.6 pour les valeurs de k (time-lag) supérieures à 5.

Ces nombres ne sont donnés qu'à titre indicatif.

- examiner chaque coefficient $\hat{\theta}$ pour voir s'il satisfait les conditions limites établies. Ces coefficients seront donnés plus bas, dans le tableau 3.

4.1.2 Inversibilité

Ce critère s'applique aux coefficients θ des processus de moyenne mobile (MA). Il sert à vérifier que l'influence des autres valeurs de la série diminue à mesure que l'on s'éloigne dans le passé (ce qui est presque naturel). Cela peut se faire en s'assurant que les conditions du tableau 3 sont vérifiées.

Le tableau 3 est donné à la page suivante

Ordre	Stationnarité	Inversibilité
1	$-1 < \phi_1 < 1$	$-1 < \theta_1 < 1$
2	$-1 < \phi_2 < 1$ $\phi_2 + \phi_1 < 1$ $\phi_2 - \phi_1 < 1$	$-1 < \theta_2 < 1$ $\theta_2 + \theta_1 < 1$ $\theta_2 - \theta_1 < 1$
1, 1	$-1 < \phi_1 < 1$	$-1 < \theta_1 < 1$

Tableau 3 Limites de validité des coefficients

Une fois les paramètres vérifiés, on doit procéder à la vérification globale de tout le modèle. Pour ce faire, on se rappelle qu'on avait posé comme hypothèse que les erreurs résiduelles suivaient une distribution normale de tendance centrale nulle. C'est cette caractéristique que l'on essaie de vérifier en se penchant sur l'autocorrélation des résidus.

Autocorrélation des résidus

Soit r_{acf} , l'autocorrélation des résidus (a_t) de la série. Le calcul de ce paramètre se fait dans le but d'avoir une idée, comme lors de l'identification, sur la relation d'interdépendance existant entre les résidus (a_t). Si cette caractéristique est statistiquement différente de zéro, on peut alors conclure qu'il y a un ou plusieurs paramètres qui n'ont pas été pris en compte dans l'expression mathématique de la série, dépendamment du nombre de racfs qui ont été non nulles.

L'autocorrélation des résidus se calcule de la même manière que l'autocorrélation de la série. La série s'écrit alors:

$$r_k(\tilde{a}_t) = \frac{\sum_{t=1}^n \tilde{a}_t \tilde{a}_{t+k}}{\sum_{t=1}^n (\tilde{a}_t)}$$

$$s(r_k) = (1 + 2 \sum_{t=1}^n r_j^2)^{1/2} n^{-1/2}$$

s étant la variance des racfs (BARTLETT)

Cette dispersion des racfs nous permet de vérifier l'hypothèse de nullité en utilisant le test du T de STUDENT.

$$t(r_k(a_t)) = \frac{r_k(a_t)}{s(r_k)}$$

Les valeurs des autocorrélations résiduelles sont statistiquement différentes de zéro si:

$$k \leq 3 \quad | \text{T-val} | \geq 1.25$$

$$k > 3 \quad | \text{T-val} | \geq 1.50$$

Test du Portemanteau

Il est utilisé pour tester l'hypothèse de nullité H_0 mais avec la condition de simultanéité en plus.

$$H_0: r_1=r_2=r_3=\dots=r_{15}=0$$

Pour ce faire, LJUNG et BOX proposent de calculer la caractéristique Q suivante:

$$Q = n(n+2) \sum_{k=1}^m (n-k)^{-1} r_k^2(a_t) \quad \begin{matrix} 5 \\ 1 \end{matrix}$$

LJUNG et BOX ont montré que Q suivait une distribution en KHI-CARRE avec 15-m degrés de liberté où m est le nombre de paramètres nécessaires pour un modèle donné:

m=2 pour les processus AR(1) et Ma(1)

m=3 pour les processus AR(2), MA(2) et ARMA(1, 1)

Les valeurs de Q obtenues dans la formule sont à comparées avec les valeurs critiques données dans le tableau suivant:

Processus	Degré de Liberté	Niveau de confiance		
		75 %	90 %	95 %
AR(1)	13	16.0	19.8	22.4
AR(2)	12	14.8	18.5	21.0
MA(1)	13	16.0	19.8	22.4
MA(2)	12	14.8	18.5	21.0
ARMA(1, 1)	12	14.8	18.5	21.0

Tableau 4: Valeurs critiques de Q

Si le paramètre Q calculé est inférieur à une valeur du tableau précédent, alors on dit que l'hypothèse qui veut que les coefficients d'autocorrélations des résidus soient tous nuls, est satisfaite à un niveau de confiance correspondant à la colonne qui a été utilisée pour établir cette comparaison.

Pour un processus $AR(1)$ par exemple, pour les autocorrélations des résidus soient significativement différentes de zéro avec 95 % de niveau de confiance, il faut que $Q > 22.4$.

Il existe aussi d'autres caractéristiques permettant de se prononcer sur le choix d'un ou de plusieurs modèles stochastiques comme mécanismes générateurs d'une série donnée. Nous en avons choisie deux: la RMSE (Root Mean Square Error) et la MAPE (Mean Absolute Percent Error). La première est une autre expression de la dispersion des résidus tandis que la seconde mesure le degré de précision, en valeur relative, qu'on devra attendre du modèle. Elles sont données par les formules suivantes:

$$RMSE = \frac{1}{n} \sum_{t=1}^n a_t^2$$

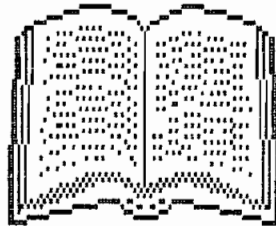
$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{a_t}{z_t} \right|$$

Reformulation du modèle

La reformulation du modèle a lieu quand on trouve une certaine autocorrélation des résidus. Comme guide de la reformulation on utilisera le diagramme des autocorrélations des résidus qui donne le nombre de paramètres manquant à la formulation par le nombre d'autocorrélations résiduelles non nulles.

Chapitre

5



REVISION

PREVISIONS

Cette ultime étape de la méthode de BOX-JENKINS est la finalité de tout modèle d'analyse prévisionnelle. La prévision se fait par l'utilisation de l'équation développée et vérifiée lors des précédentes étapes:

$$z_t = C + \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

avec

$$C = \mu \left(1 - \sum_{i=1}^n \phi_i \right)$$

$$\mu = \frac{1}{n} \sum_{t=1}^n z_t$$

A l'aide de l'équation précédente, pour n'importe quelle valeur de t , z_t peut être calculé si les valeurs antécédentes (z_{t-1}, z_{t-2}, \dots) sont connues. A partir d'une origine définie, toutes les valeurs subséquentes de z_t peuvent être générés en faisant les hypothèses suivantes:

- A un instant t postérieur à la dernière date de la série, les valeurs de z_{t-1} et de a_{t-1} sont égales à leurs valeurs les plus probables

- * pour z_t , sa valeur la plus probable est celle obtenue à partir de l'équation précédente.
- * pour a_t , sa valeur la plus probable est égale à zéro.

Rappelons que, tel que mentionné à l'introduction, la méthode de BOX et JENKINS s'appliquent à la prévision à court terme. La raison en est que la supposition faite précédemment conduit, pour une série stationnaire, à une convergence des prévisions vers la moyenne arithmétique de la série car, si la valeur la plus probable du terme tenant compte de l'erreur résiduelle est nulle, on remarquera, en regardant l'expression générale des séries, qu'à la limite nous n'aurons que des termes autorégressifs qui, comme on l'a déjà montré au chapitre précédent tendent vers zéro, à mesure que l'ordre du processus augmente. On peut même dire que pour un processus de type $MA(q)$, à un instant $t_0 \geq t+q$, t étant l'origine de la prévision, toutes les valeurs prévues de z_{t_0} sont égales à la moyenne de la série considérée.

Si on a par exemple un processus MA(1), toutes les valeurs de prévisions au-delà de l'ordre 1 sont égales la moyenne de la série.

Notons que si la série n'est pas originellement stationnaire et qu'il a donc fallu lui appliquer la méthode de BOX et JENKINS après l'avoir rendue stationnaire par différenciation successive, ce qui vient d'être dit plus haut ne lui est pas applicable car ne fluctuant pas autour d'une valeur fixe (critère de stationnarité), on ne doit pas s'attendre à ce que ses valeurs les plus probables (prévisions) convergent vers la moyenne de la série.

Intervalles de confiance des prévisions

Soit $e_t(l)$ l'erreur de prévision commise sur une valeur située à une distance l de l'origine des prévisions. On la note:

$$e_t(l) = z_t - z_{t+l}$$

La valeur de la dispersion des erreurs résiduelles de prévision est donnée par leur variance.

$$\sigma^2 [e_t(1)]$$

Si les erreurs résiduelles sont normalement distribuées (hypothèse que nous avons déjà formulée), et si la taille de l'échantillon est assez grande, alors les prévisions suivent elles aussi une distribution qui peut être prise pour normale. On peut donc, autour de chaque point de prévision, définir un intervalle de confiance en utilisant une table de probabilités pour une distribution normale.

C O N C L U S I O N

Nous dirons que l'outil qui a été présenté dans les pages précédentes n'est et ne sera jamais qu'un outil, puissant certes, mais il n'est ni plus ni moins qu'un instrument de travail dont la pleine mesure n'est atteinte que par un utilisateur ayant un solide background en théorie des processus stochastiques. Mais il n'en demeure pas moins utilisable (dans sa version programme d'ordinateur) par n'importe quelle personne sachant manipuler un tant soit peu des données statistiques.

Sa facilité d'abord fait, de jour en jour, grossir le nombre de ses adeptes. Mais il faudrait peut être dire qu' aucune méthode statistique ne peut fournir des résultats de prévision plus précis que les données elles-mêmes. C'est pourquoi une attention soutenue devrait être accordée aux méthodes d'acquisitions des données en vue de leur traitement statistique.

La méthode qui vient d'être présentée est générale et s'applique à tous les corps de métiers. Cette universalité est due au fait qu'elle ne présuppose pas l'existence d'une loi fondamentale sous-jacente à une série statistique donnée mais propose plutôt une démarche rigoureuse applicable quelque soit l'origine et la dimension de la série.

Nous avons en ce qui nous concerne essayer de déblayer le terrain et de proposer à l'analyste des données, un outil qui, bien que paraissant simplifié, n'en permet pas moins de résoudre la quasi-totalité des problèmes qui se posent à lui.

Les cinq modèles stochastiques présentés tout au long de ce travail doivent nécessairement satisfaire ses besoins. Si tel n'était pas le cas, il devrait se poser des questions sur les jugements qu'il a fait tout au long du processus de modélisation.

Le programme joint en annexe devrait l'aider dans cette tâche souvent ignorée qu'est l'analyse prévisionnelle.

ANNEXES

ANNEXES

Dans les pages qui vont suivre nous donnons le listing du programme qui a été écrit comme support pratique de l'analyse des données par la méthode de BOX et JENKINS.

Le programme écrit en FORTRAN 77 se veut modulaire et chacun des chapitres précédents fait l'objet de sous programmes et sont ainsi aisément repérables.

La technique de choix à l'aide de menus a été utilisé tout au long du développement de ce logiciel, ceci dans le but d'offrir à l'utilisateur une certaine forme d'interactivité lui permettant de revenir à n'importe quel moment de la modélisation sur les décisions qu'il a prises ou alors de modifier l'ordre des étapes pour satisfaire des besoins propres. On peut par exemple, tout de suite après avoir estimé les coefficients du modèle, procéder à une prévision dans le but, par exemple, d'avoir une idée sur le

comportement de la série dont on dispose à un certain ordre.

Il faut quand même noter que cette version est préliminaire car nous n'avons pas pu procédé procéder, au moment de la remise du projet à un test complet du logiciel. Néanmoins une version compilée devrait être disponible au centre de calcul de l'Ecole Polytechnique de Thiès.

BIBLIOGRAPHIE

GEP BOX et G.M JENKINS "Time series
Analysis—Forecasting and control"

A. PANKRATZ "Forecasting with univariate Box
and JENKINS models"

WHEEL WEIGHT et MAKRILAKIS "Choix et
valeur des methodes de precision"

D. DURAND "Hydrologie statistique approfondie"

UNESCO / WMO "Hydrological forecasting"

HINES et MONTGOMERY "Probability and
statistics in Ingeneering and management science"

```
Programm Arima
Implicit real*8(a-h, o-z)
character*12 entree, sortie
Dimension donnees(2000), acf(15, 3), pacf(15, 3)
```

```
c
C
1  Format('1', 'Entrer le nom du fichier de données.....', a12, $)
2  Format(' ', 'Entrer le nom du fichier de résultats.....', a12, $)
   Read(*, 1) Entree
   read(*, 2) Sortie
c   Open(unit=5, file=Entree)
c   Open(unit=6, file=sortie)
c
   do 10 i=1, 10000
10  read(5, *, ERR=20) donnees(i)
20  n=i
c
   format('1', 20x, '* * * M E N U   G E N E R A L   * * *')
4   Format(4(/))
c
6   format(' ', 10x, '[1] Identifier la série', /, 10x)
7   format('[2] Estimer les paramètres du modèle', /, 10x)
8   format('[3] Tester les paramètres ', /, 10x)
9   format('[4] Faire des prévisions', //, 10x)
5   format('[0] Quitter le programme', ///, 30x, $)
C
C
30  write(*, 3)
   write(*, 4)
   Write(*, 6)
   write(*, 7)
   write(*, 8)
   write(*, 9)
   write(*, 5)
   read(*, *) in
C
   if(in.eq.0) goto 100
   if(in.eq.1) goto 200
   if(in.eq.2) goto 300
   if(in.eq.3) goto 400
   if(in.eq.4) goto 500
c
   goto 30
c
200  call identifier
   goto 30
300  call Estimer
   goto 30
400  call Tester
   goto 30
500  call Prevoir
   goto 30
c
C
100  stop
END
```

```

Subroutine trace(acf, pacf)
Implicit real*8(a-h, o-z)
character code*2, cligne*40
dimension xligne(40), acf(15, 3), pacf(15, 3)

```

```

codeFF=char(12)
codeLF=char(10)

```

```

xligne(1)=-1.
Do 100 i=2, 40
xligne(i)=xligne(i-1)+.05

```

```

1 Format(' ', ' ', 58('='), '||')
2 Format(' ', ' ', a58, '||')
3 Format(' ', ' ', 3('===='), 40('='), '||')
4 Format(' ', ' ', Lag || Coef. || T-Val || ', 20x, ', 19x, '||')
5 Format(' ', ' ', 3('===='), 40('='), '||')
6 Format(' ', ' ', 15, ' ', f5.2, ' ', f5.2, ' ', a40, '||')
7 Format(' ', ' ', 3('===='), 40('='), '||')

```

```

write(6, *) codeFF
write(6, 1)

```

```

write(6, 2) '
Write(6, 2) '          A U T O C O R R E L A T I O N S
write(6, 2) '
write(6, 3)
write(6, 4)
write(6, 5)

```

```

do 10 i=1, 15
ac1=acf(i, 1)
ac2=acf(i, 2)
call ligne(ac1, ac2)
write(6, 6) i, ac1, acf(i, 3), codeligne
continue
write(6, 7)
write(6, *) codeLF
write(*, *) 'Appuyer sur < pour continuer'
read(*, *) code

```

```

write(6, 1)
write(6, 2) '
Write(6, 2) '          Autocorrélations partielles
write(6, 2) '
write(6, 3)
write(6, 4)
write(6, 5)

```

```
c
do 20 i=1,15
pac1=pacf(i,1)
pac2=pacf(i,2)
call ligne(pac1,pac2,xligne,cligne)
write(6,6) i,pac1,pacf(i,3),cligne
20 continue
write(6,7)
write(6,*) codeLF
write(*,*) 'Appuyer sur ↵ pour continuer'
read(*,*) code
```

```
c
c
Return
End
```

```
c
c
```

```
Subroutine Identifier(don, nx, acf, pacf)
Implicit real*8(a-h, o-z)
dimension don(nx), acf(15, 3), pacf(15, 3)
character*2 code
```

```
CodeCS=char(12)
```

```
Format(' ', '*** MENU D' IDENTIFICATION ***', ///)
Format(/, 10x, '[1] Calculer les coefficients de corrélation')
Format(/, 10x, '[2] Différenciere la série')
Format(/, 10x, '[0] Retourner au Menu Général')
Format(////, 20x, 'Vous désirez ?', i2, $)
```

```
write(*, *) codeCS
write(*, 1)
write(*, 2)
write(*, 3)
write(*, 4)
read(*, 5) in
```

```
if(in.eq.0) return
if(in.eq.1) then
do 20 i=1, nx-1
don(i)=don(i+1)-don(i)
nx=nx-1
else
if(in.eq.2) then
call centrer(don, nx, acf, pacf)
endif
endif
goto10
```

```
END
```

```
Subroutine centrer(dat, nx, acf, pacf)
Implicit real*8(a-h, o-z)
dimension dat(nx), d(2000), acf(15, 3), pacf(15, 3)
```

```
c
c
  Do 10 I=1, nx
d(i)=dat(i)
som=som+d(i)
10  som2=som2+d(i)**2
    moy=moy/nx
```

```
c
c
  do 20 i=1, nx
20  d(i)=d(i)-moy
```

```
c
    call acorr(d, nx, acf)
    call pacorr(acf, pacf)
    call trace(acf, pacf)
```

```
c
    return
    end
```

```

subroutine ligne(a,b,c,xligne)
implicit real*8(a-h,o-t)

character, car(40)
dimension xligne(40)
do 100 i=1,40
100 car(i)=' '
do 10 i=1,40
aa=a-xligne(i)
if(aa) 920,20,10
10 CONTINUE
20 if(i-20) 30,40,50
30 do 35 j=i,19
40 car(i)='«'
go to 40
50 do 55 j=21,i
55 car(i)='»'
40 car(20)='1'
do 60 i=20,40
bb=b-xligne(i)
if(bb) 70,70,60
60 continue
70 car(i)=']'
car(20-i)='['
write(6,*)(car(i),i=1,40),'||'
return
END
SUBROUTINE ESTIMER(acf,pacf,p,t)
implicit real *8 (a-h,o-z)
dimension acf(15,3),pacf(15,3),p(15),t(15)
character code*2
LOGICAL EVAL
code rb=char(07)
1 FORMAT('o')'*** ESTIMATION DES PARAMETRES DU MODELE
2 FORMAT( || ||,10x,'[1] ar(1): auto regressif d'ordre 1')
3 FORMAT (/,10x,'[2] ar(2): auto regressif d'ordre 2')
4 FORMAT (/,10x,'[3] MA(1): moyenne mobile d'ordre 1')
5 FORMAT (/,10x,'[4] ma(2) : moyenne mobile d'ordre 2')
6 FORMAT (/,10x,'[5] arma(1,1): modele mixte')
7 FORMAT (/,10x,'[6] CAS GENERAL ')
8 FORMAT ( || ||,10x,'[0] retour au menu appelant ')
9 FORMAT ( || ||,20x,'vous desirez ?',i2,$ )
DO 12 i=1,15
P(i)=0
12 T(i)=0
11 DO 10.i=1, 8
10 WRITE (*,i)
read (*,9) in
C IF(in.lt.O.OR.IN.GT.6) go to 11
IF (in.eq.1) then v1=acf(1,1)
v2=acf(1,2)
v3=acf(1,3)
if(in.eq.1) p(1)=R1
If(in.eq.3) then
t1=-1/2/R1+(1/(2*R1)**2-1)**.5

```

```

t2=-1/2/R1-(1/(2*R1)**2-1)**.5
T(1)=donax(t1,t2)
endif
if(in.eq.2) then
p(1)=(1-R2)*r1/(1-R1**2)
p(2)=(R2-R1**2)/(1-R1**R2)
if (dab>(p(1)).lt.1
1.and.(p(1)+P(2)).lt.1
2.and.(p(2)-p(1)).lt.1)
eval=dab>p(1).lt.1.and.(p(1)+p(2)).lt.1.and.(p(2)-p(1)).lt.1
IF(.not.eval) .call message
endif
if(in.eq.5)then
p(1)=r2/R(1)
t1=0.
P1=p(1)
do 20,i=1,100
T1=((1-P1*T1)*(p1-t1)(r1-1-T1**2)/2/P1
T(1)=t1
IF EVAL =dabs(p1).lt.AND.DABS(t1).LT.1
IF(.NOT.EVAL) call message
endif
if(in.eq.4) then
T1=0
T2=0
DO 30 i=1 , 100
T1=(t2-t1)/R1/1+T1**2+T2**2)
30 T2=-1/R2/(1+T1**2+T2**2)
eval=dabs (t1).LT.1.AND.(T2+T1).LT.1.AND.(t2-T1).LT.1.AND.DABS(T2).LT.1
IF(.NOT.EVAL) call message
endif
if(IN.EQ.6) call car gen (acf,pacf,p,t)
if(IN.EQ.0) return
write(*,*) ' appuyer sur ↵ pour continuer '
read (*,*) code
go to ||
end
SUBROUTINE MESSAGE
CHARACTER CODE
C
WRITE (*,*) 'LES PARAMETRES DE CE MODELE SONT EN DEHORS DES LIMITES PERMISES
WRITE (*,'///',20x) 'appuyer sur ↵ pour continuer'
read (*,*) code
C
return
end
SUBROUTINE CAR GEN(acf,pacf,p,t)
implicit real*8(a-h,o-z)
character code *2
dimension acf(15,3) , pacf(15,3),p(15),t(15)
C
write (*,'10x') , 'non encore implementee'
write (*,10x), 'appuyer sur ↵ pour continuer,'\
READ (*,*) code
end

```



```

SUBROUTINE TESTER (n,mx,res,dat,model)
IMPLICIT REAL*8(a-h,o-z)
DIMENSION RES(-1:nx),dat(-1:nx),t(15),p(15)
  dat(-1)=0
  dat(0)=0
  res(-1)=0
  res(0)=0
  s1=0
  s2=0
  do 10 i=1,nx
  z=dat(i)
  z1=dat(i-1)
  z2=dat(i-2)
  a1=res(i-1)
  a2=res(i-2)
  A=Z-(P1*Z1+P2*Z2-T1*A1-T2*A2)
  res(i)=a
  S1=s1+dabs(a/z)
  s2=s2+A**2
10 CONTINUE
  m=0
  DO 20 I=1,15
  IF (T(i).NEQ.0) m=m+1
  IF (p(i).NEQ.0) m=m+1
20 CONTINUE
  M=m+1
  ESMR =1/(n-m)*s2
  epam =1/N*s1
C
  if (modele.eq.1) then
  S(1,1)=p1
  S(1,2)=(1-p1**2)/n
  S(1,4)=1
  ENDIF
  if (modele.eq.2) then
  s(1,1)=p1
  s(2,1)=p2
  s(1,2)=(1-p2**2)/n
  s(2,2)=s(1,2)
  s(1,4)=-p1(1-p2)
  s(2,4)=s(1,4)
  ENDIF
  IF (modele.eq.3) then
  s(1,1)=t1
  S(1,2)=(1-T1**2)/N
  S(1,4)=1
  ENDIF
  IF(modele,eq,4) then
  s(1,1)=t1
  s(2,1)=t2
  s(1,2)=(1-t2**2)/n
  s(2,2)=s(1,2)
  s(1,4)=-t1/(1-t2)
  s(2,4)=s(1,4)
  ENDIF

```

```
IF (modele.eq.5) then
s(1,1)=p1
s(2,1)=t1
s(1,2)=(1-p1*t1)**2*(1-p1**2)/(p1-t1**2)
s(2,2)=(1-p1*t1)**2*(1-t1**2)/(p1-t1**2)
s(1,4)=(1-p1**2)*(1-p1**2)(1-t1**2)/(p1-t1**2)
s(2,4)=s(1,4)
ENDIF
m=m-1
do 30 i=1,m
30  s(i,3)=dabs(s(i,1)/s(i,2))
SUBROUTINE DIGNOSTIC (res,nx)
IMPLICIT REAL*8 (a-h,o-z)
DIMENSION racf(15,3),res(nx)
CALL CENTER(res,nx,racf)
```

```

SUBROUTINE TESTER
IMPLICITE REAL*8(a-h,o-z)
DIMENSION racf(15,3),rpacf(15,3),coef(4)
DATA /COEF/22.4,21.0,22.4,21.0,21.0,21.0/
CALL CENTER (ics,nx,racf,rpacf)
s1=0
do 10 k=1,15
s1=s1+racf(k,1)**2/(nx-k)
xq=nx*(nx+2)*s1
IF(xq.LT.COEF(modele)) then
write(*,*) 'le modele choisi est adequat pour la precision'
else
write (*,20) xq
write (*,*) 'elle ne satisfait pas a l'hypothese de nulliteavec 95%
DE NIVEAU DE CONFIANCE '
ENDIF
WRITE (*,*) 'appuyer sur <J pour retourner au menu '
read (*,*) code
20 format (' ',_ _ _ la valeur Q du test de LJUNG-BOX vaut',
f5.1 )

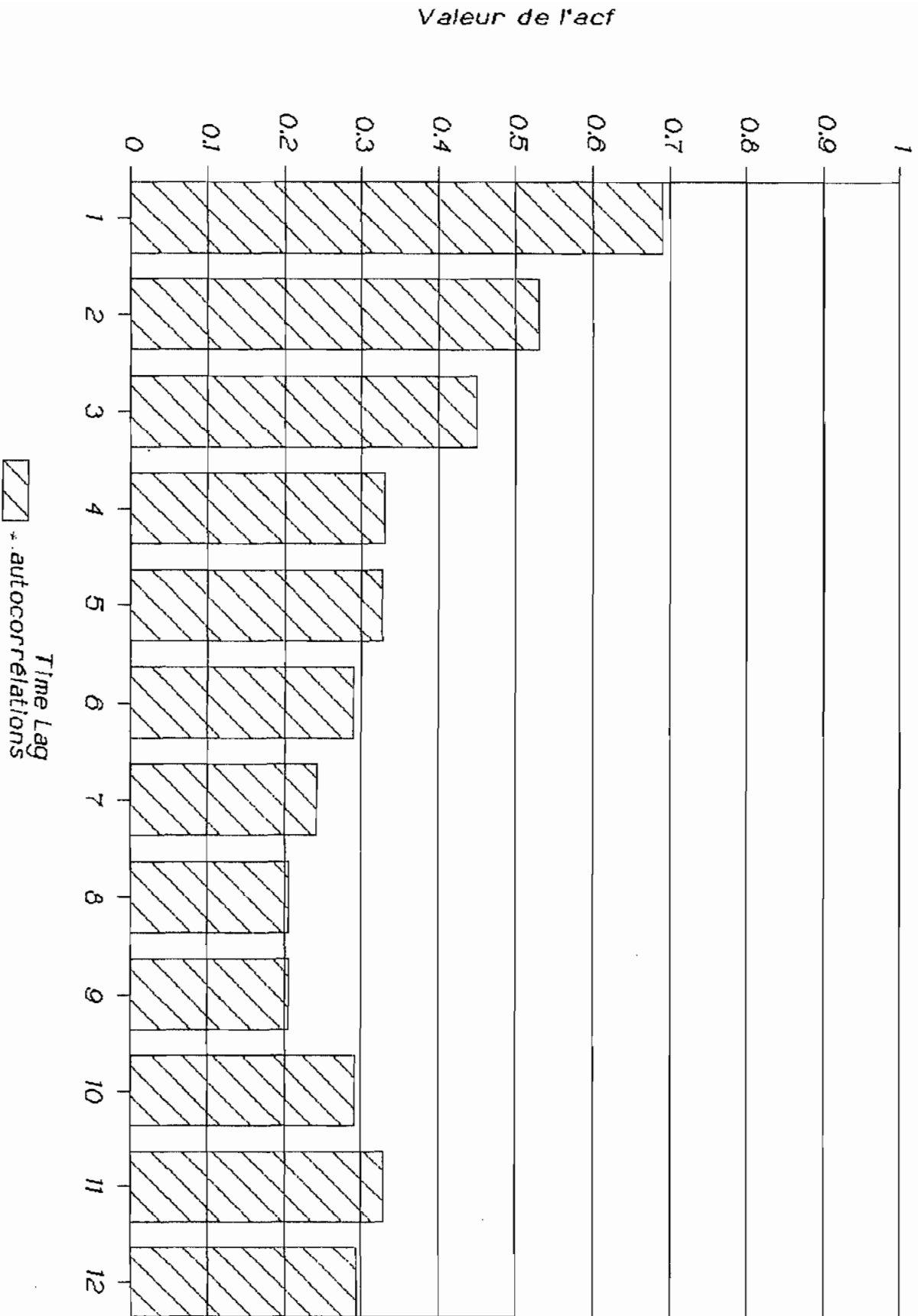
return
END
SUBROUTINE PREVOIR
IMPLICIT REAL *8(a-h,o-z)
CHARACTER CODE*2
DIMENSION DONNEES(2000),dat(2000),res(2000)
w(i)= p1*dat(i-1)*p2*pdat(i-2)*res(i)
-t1*res(i-1)-t2*res(i-2)
t1=t(1)
p1=p(1)
t2=t(2)
p2=p(2)
pdat(-1)=0
FDAT(0)=0
DO 10 i=1,nx
10 pdat(i)=w(i)
20 FORMAT (' ','ENTRER LES BORNES DE L'INTERVALLE DE PRECISION ',
$)
READ (*,20) 10,11
IF (10.GT.NX) then
DO 39 i=nx,10
30 pdat(i)=w(i)
ENDIF

RETURN
END

```

FONCTION D'AUTOCORRÉLATIONS

Cas #1



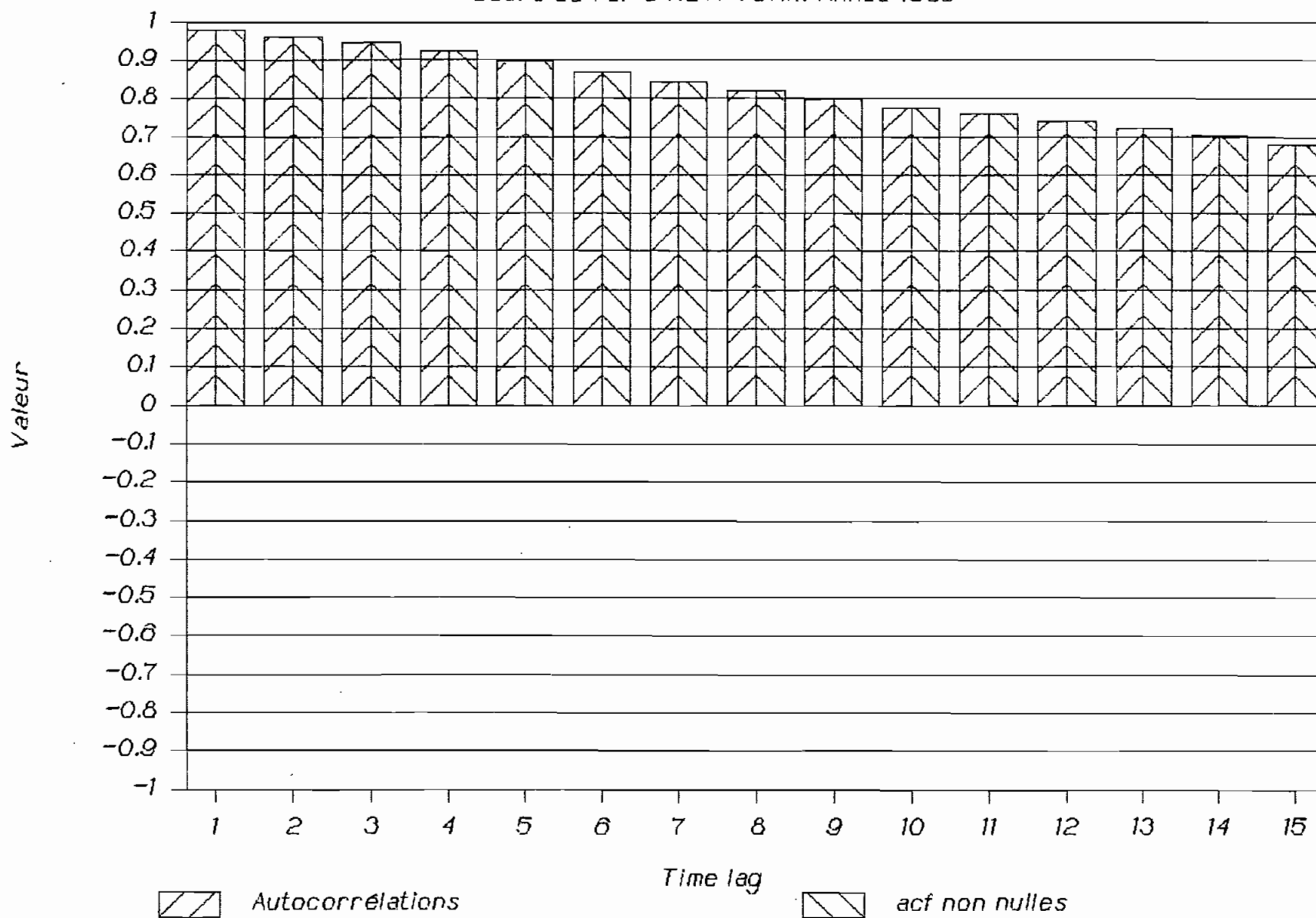
SERIES CHRONOLOGIQUES

Exemples traités

Leg	acf	s(r)	Tval	pacf	s(r)	Tval	acfcorr	pacfcorr
Casi								
1	0.69	0.129	5.343	0.69	0.129	5.343	0.69	0.69
2	0.527	0.18	2.924	0.099	0.129	0.763	0.527	0
3	0.447	0.204	2.187	0.099	0.129	0.764	0.447	0
4	0.33	0.22	1.499	-0.061	0.129	0.471	0	0
5	0.33	0.228	1.444	0.153	0.129	1.184	0	0
6	0.291	0.236	1.234	-0.007	0.129	0.051	0	0
7	0.244	0.242	1.008	0.006	0.129	0.046	0	0
8	0.211	0.246	0.859	-0.015	0.129	0.115	0	0
9	0.211	0.249	0.848	0.083	0.129	0.645	0	0
10	0.295	0.252	1.171	0.197	0.129	1.527	0	0
11	0.339	0.258	1.314	0.087	0.129	0.671	0	0
12	0.307	0.265	1.16	-0.045	0.129	0.352	0	0
13	0.268	0.271	0.99	-0.036	0.129	0.282	0	0
14	0.142	0.275	0.515	-0.179	0.129	1.386	0	0
15	0.055	0.276	0.2	-0.093	0.129	0.721	0	0
Année 1982								
1	0.975	0.063	15.568	0.975	0.063	15.568	0.975	0.975
2	0.951	0.107	8.92	0.018	0.063	0.294	0.951	0
3	0.922	0.136	6.781	-0.137	0.063	2.18	0.922	-0.137
4	0.894	0.159	5.639	0.021	0.063	0.339	0.894	0
5	0.868	0.177	4.9	0.044	0.063	0.696	0.868	0
6	0.843	0.193	4.363	-0.021	0.063	0.336	0.843	0
7	0.815	0.207	3.938	-0.06	0.063	0.96	0.815	0
8	0.792	0.219	3.609	0.06	0.063	0.964	0.792	0
9	0.77	0.23	3.343	0.043	0.063	0.681	0.77	0
10	0.748	0.24	3.114	-0.033	0.063	0.519	0.748	0
11	0.726	0.249	2.916	-0.017	0.063	0.268	0.726	0
12	0.703	0.257	2.731	-0.042	0.063	0.678	0.703	0
13	0.676	0.265	2.554	-0.072	0.063	1.148	0.676	0
14	0.649	0.271	2.393	-0.021	0.063	0.33	0.649	0
15	0.624	0.277	2.249	0.027	0.063	0.436	0.624	0
Année 1983								
1	0.979	0.064	15.362	0.979	0.064	15.362	0.979	0.979
2	0.962	0.109	8.832	0.066	0.064	1.039	0.962	0
3	0.944	0.139	6.779	-0.025	0.064	0.384	0.944	0
4	0.923	0.163	5.657	-0.07	0.064	1.104	0.923	0
5	0.896	0.183	4.891	-0.18	0.064	2.83	0.896	-0.18
6	0.869	0.2	4.342	-0.032	0.064	0.498	0.869	0
7	0.844	0.215	3.925	0.034	0.064	0.53	0.844	0
8	0.82	0.228	3.596	0.05	0.064	0.778	0.82	0
9	0.797	0.24	3.324	0.041	0.064	0.647	0.797	0
10	0.776	0.25	3.101	0.049	0.064	0.762	0.776	0
11	0.758	0.26	2.915	0.037	0.064	0.587	0.758	0
12	0.741	0.269	2.758	0.025	0.064	0.397	0.741	0
13	0.722	0.277	2.609	-0.084	0.064	1.311	0.722	0
14	0.702	0.284	2.466	-0.108	0.064	1.687	0.702	0
15	0.681	0.291	2.337	-0.036	0.064	0.569	0.681	0
Année 1984								
1	0.979	0.063	15.477	0.979	0.063	15.477	0.979	0.979
2	0.959	0.108	8.876	0.012	0.063	0.197	0.959	0
3	0.942	0.138	6.833	0.081	0.063	1.288	0.942	0
4	0.925	0.162	5.723	-0.03	0.063	0.468	0.925	0
5	0.904	0.182	4.982	-0.075	0.063	1.179	0.904	0
6	0.888	0.199	4.469	0.083	0.063	1.307	0.888	0
7	0.872	0.214	4.074	-0.012	0.063	0.197	0.872	0
8	0.857	0.228	3.761	0.031	0.063	0.492	0.857	0
9	0.842	0.24	3.504	0.01	0.063	0.162	0.842	0
10	0.83	0.252	3.295	0.042	0.063	0.662	0.83	0
11	0.818	0.263	3.115	0.016	0.063	0.256	0.818	0
12	0.805	0.273	2.954	-0.027	0.063	0.429	0.805	0
13	0.792	0.282	2.807	-0.026	0.063	0.411	0.792	0
14	0.781	0.291	2.686	0.06	0.063	0.947	0.781	0
15	0.769	0.299	2.574	-0.016	0.063	0.246	0.769	0
Année 1985								
1	0.074	0.075	0.993	0.074	0.075	0.993	0	0
2	-0.822	0.075	10.903	-0.832	0.075	11.098	-0.822	-0.832
3	0.093	0.115	0.809	0.872	0.075	11.634	0	0.872
4	0.97	0.116	8.391	0.594	0.075	7.927	0.97	0.594
5	0.064	0.155	0.413	0.383	0.075	5.106	0	0.383
6	-0.796	0.155	5.138	0.225	0.075	3.008	-0.796	0.225
7	0.099	0.176	0.563	0.102	0.075	1.353	0	0
8	0.94	0.177	5.32	0.019	0.075	0.26	0.94	0
9	0.054	0.203	0.265	-0.012	0.075	0.154	0	0
10	-0.77	0.203	3.793	-0.008	0.075	0.109	-0.77	0
11	0.106	0.219	0.483	0	0.075	0.004	0	0
12	0.91	0.219	4.155	-0.009	0.075	0.122	0.91	0
13	0.043	0.239	0.181	-0.01	0.075	0.129	0	0
14	-0.744	0.239	3.108	0	0.075	0.005	-0.744	0
15	0.112	0.252	0.444	0.003	0.075	0.033	0	0

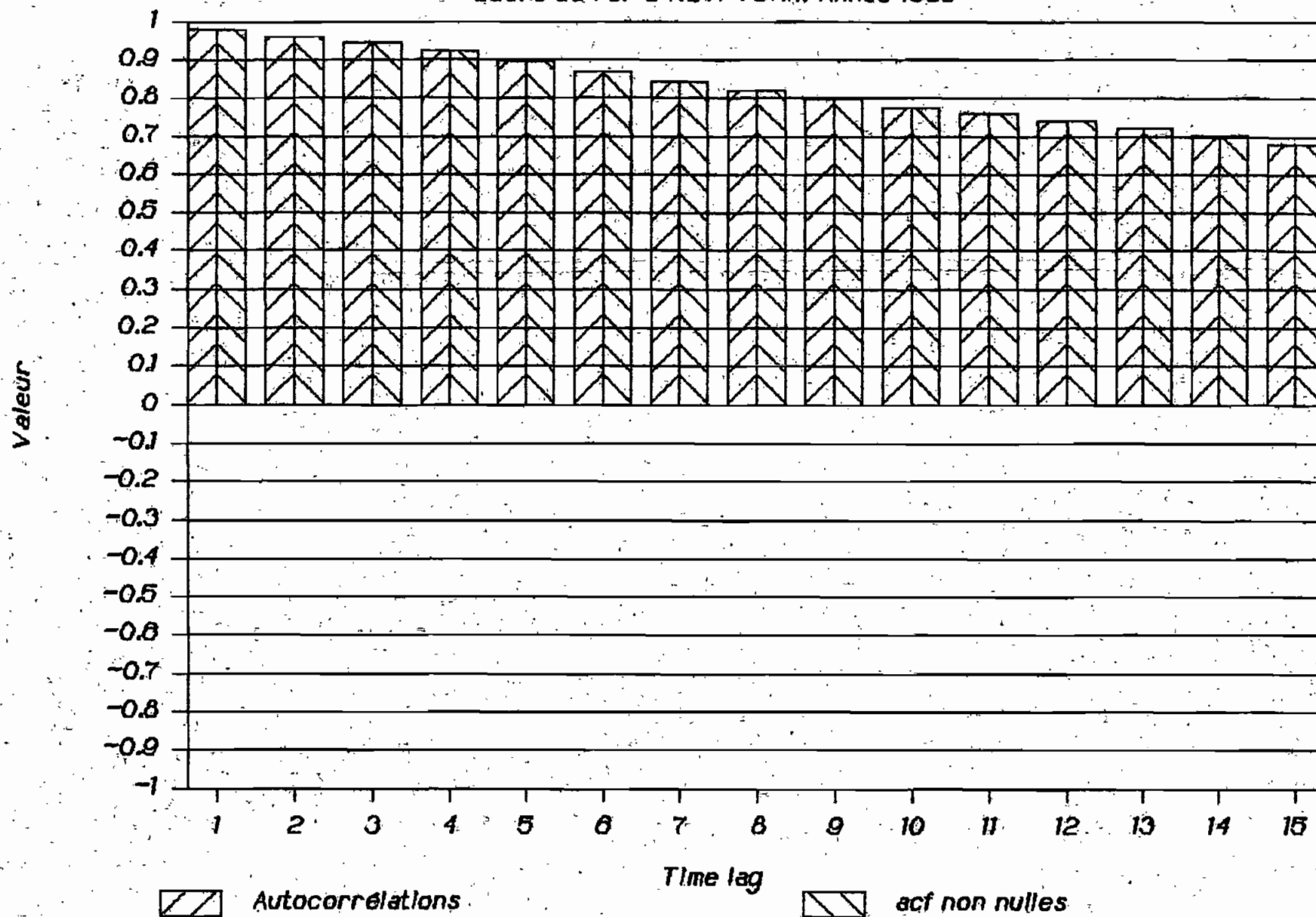
Fonctions d'autocorrélations

Cours de l'or à NEW YORK. Année 1983



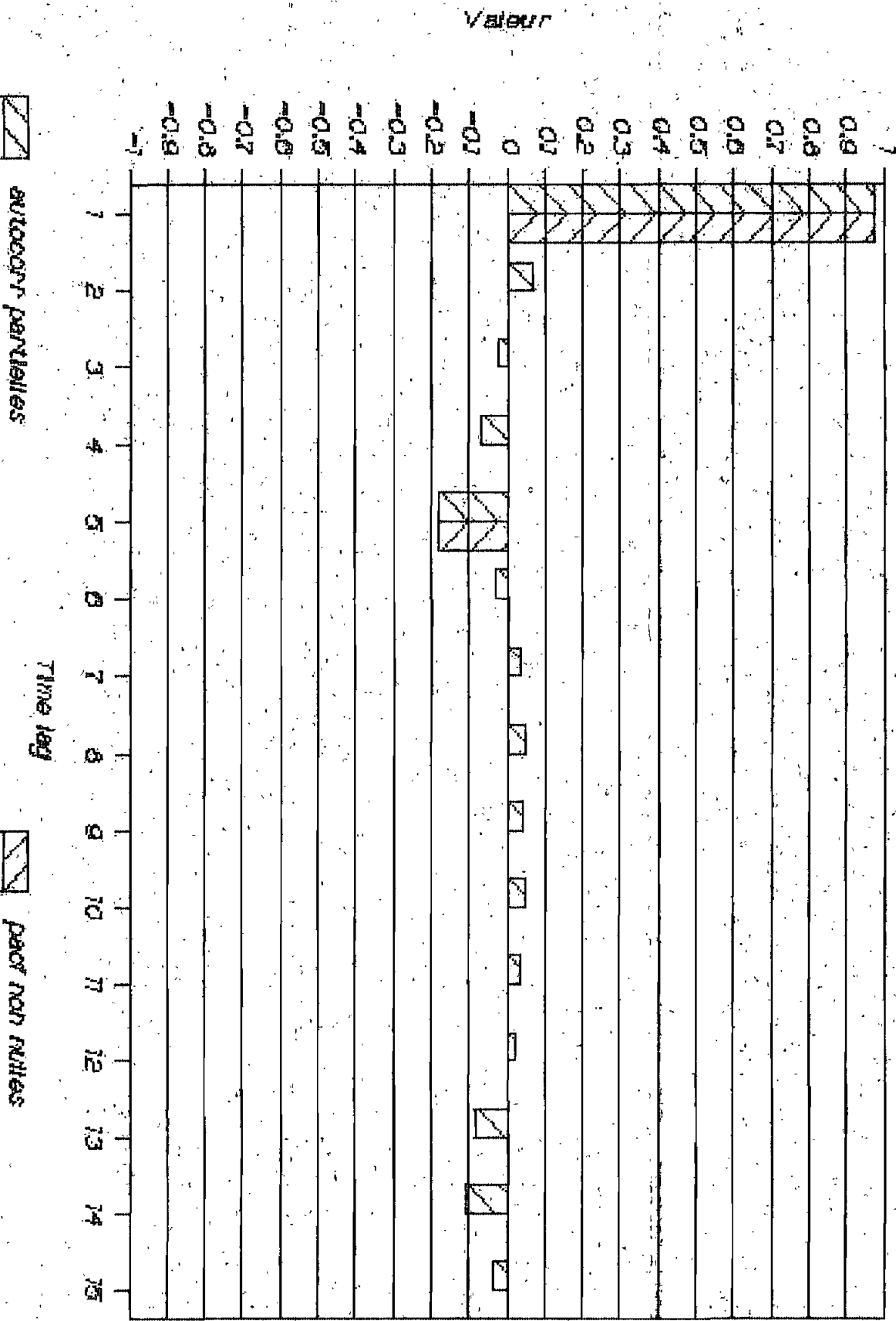
Fonctions d'autocorrélations

Cours de l'or à NEW YORK, Année 1983



Fonctions d'autocorrélations partielles

COURS DE T OR A NEW YORK, Année 1983



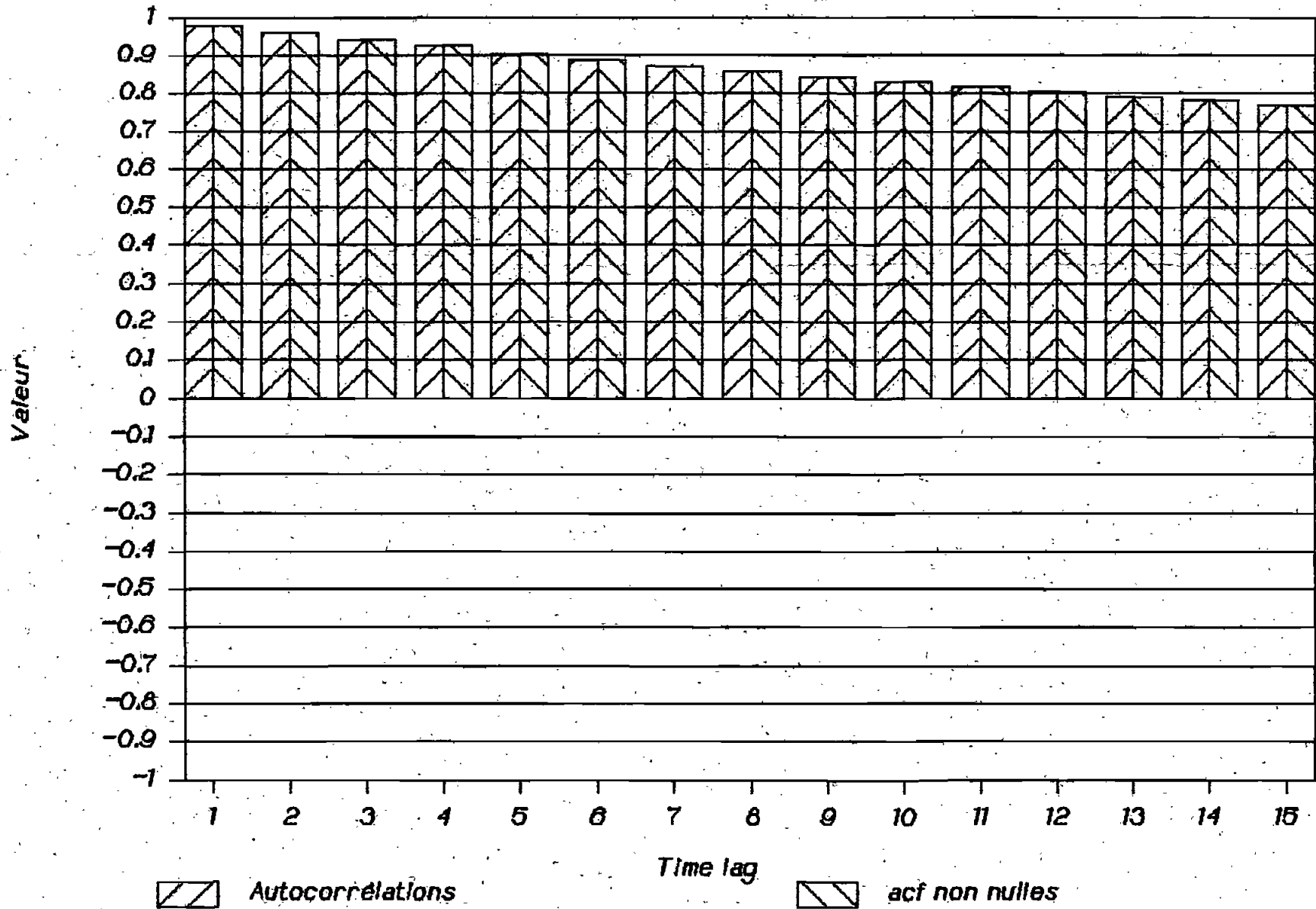
autocorr partielles



pas de zéros

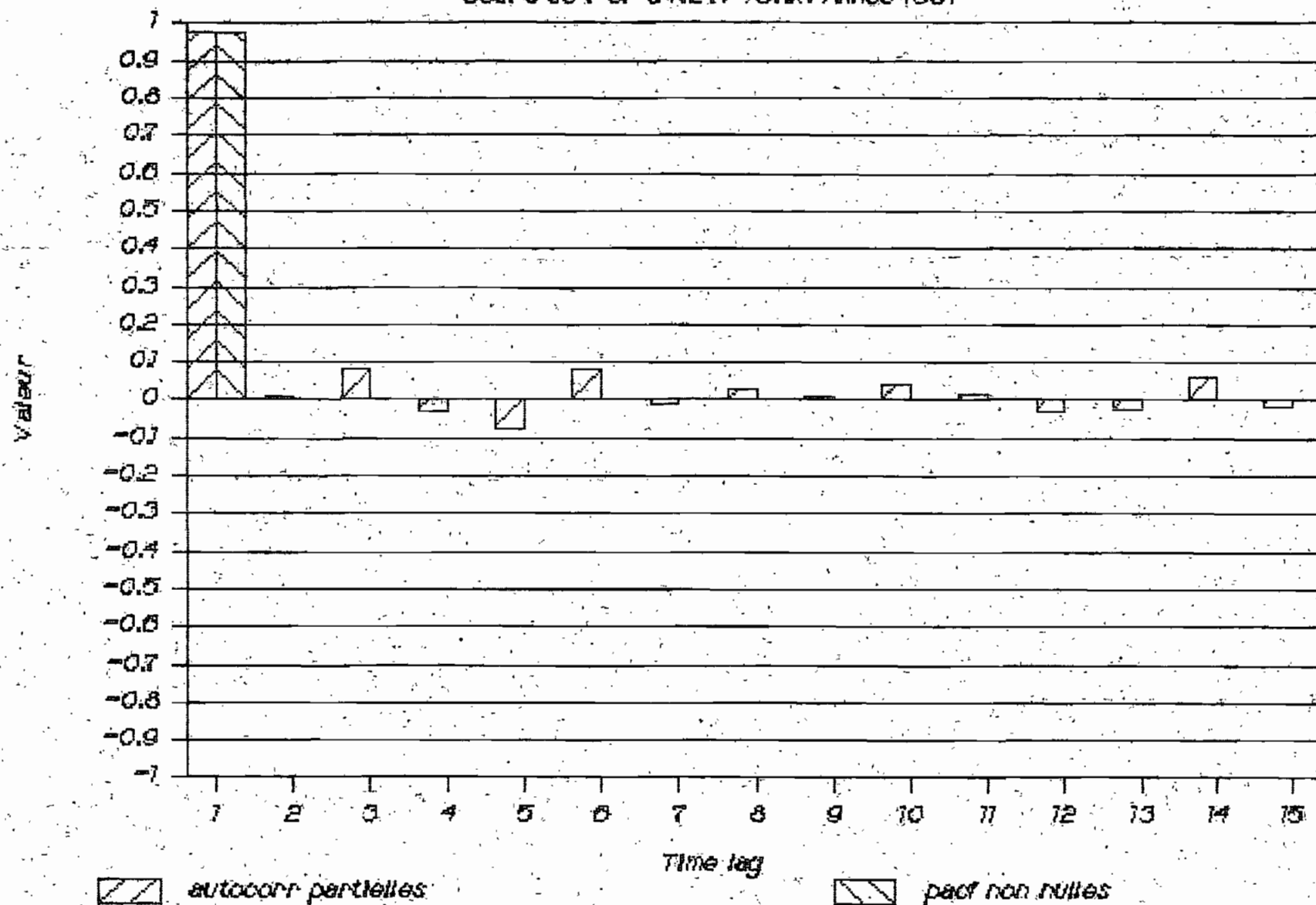
Fonctions d'autocorrélations

Cours de l'or à NEW YORK. Année 1984



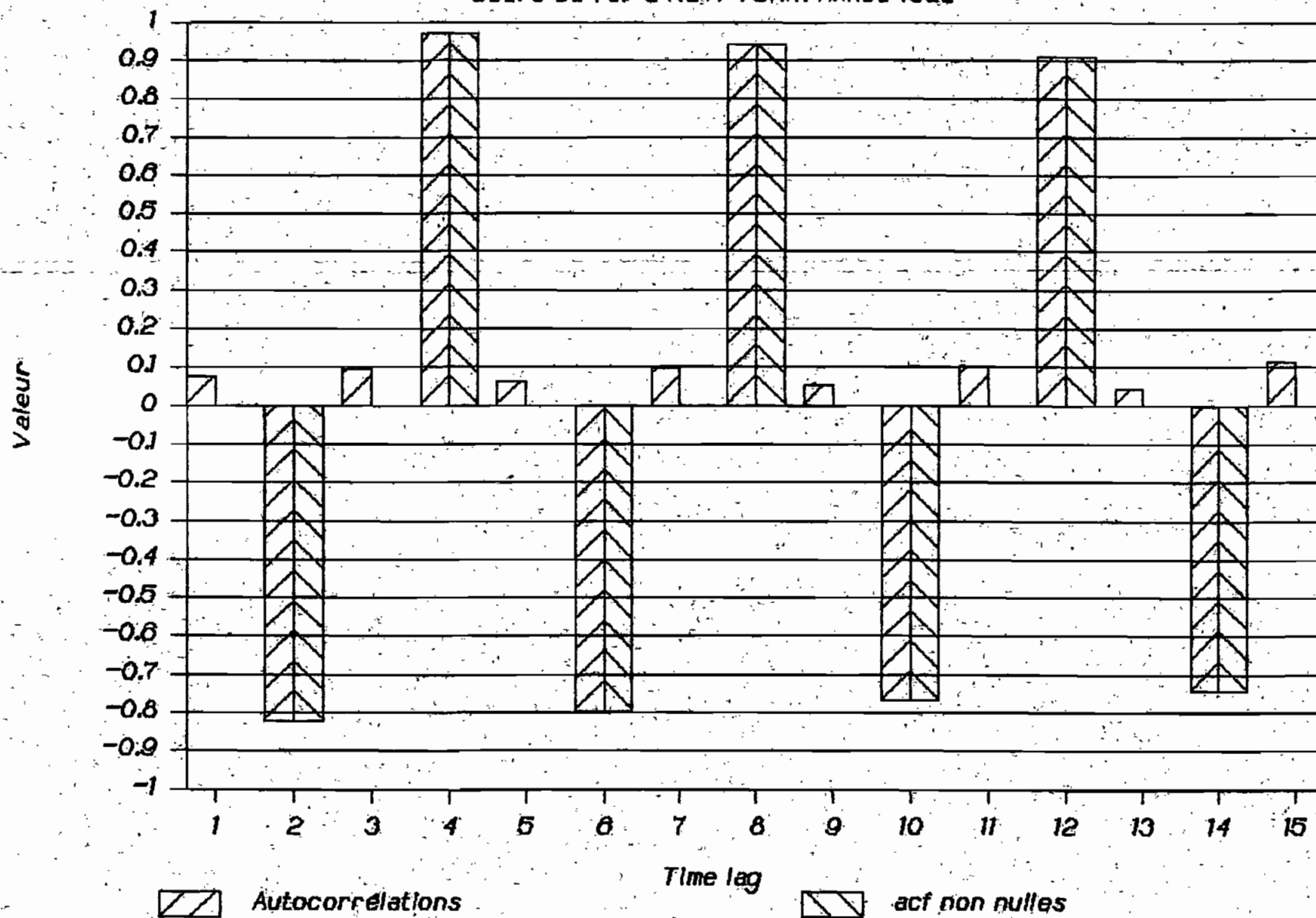
Fonctions d'autocorrélations partielles

Cours de l'or à NEW YORK. Année 1984



Fonctions d'autocorrélations

Cours de l'or à NEW YORK. Année 1985



Fonctions d'autocorrelations partielles

COURS DE STATISTIQUE A NEW YORK, ANNÉE 1967

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
0.9																
0.8																
0.7																
0.6																
0.5																
0.4																
0.3																
0.2																
0.1																
0																
-0.1																
-0.2																
-0.3																
-0.4																
-0.5																
-0.6																
-0.7																
-0.8																
-0.9																



autocorr. partielles

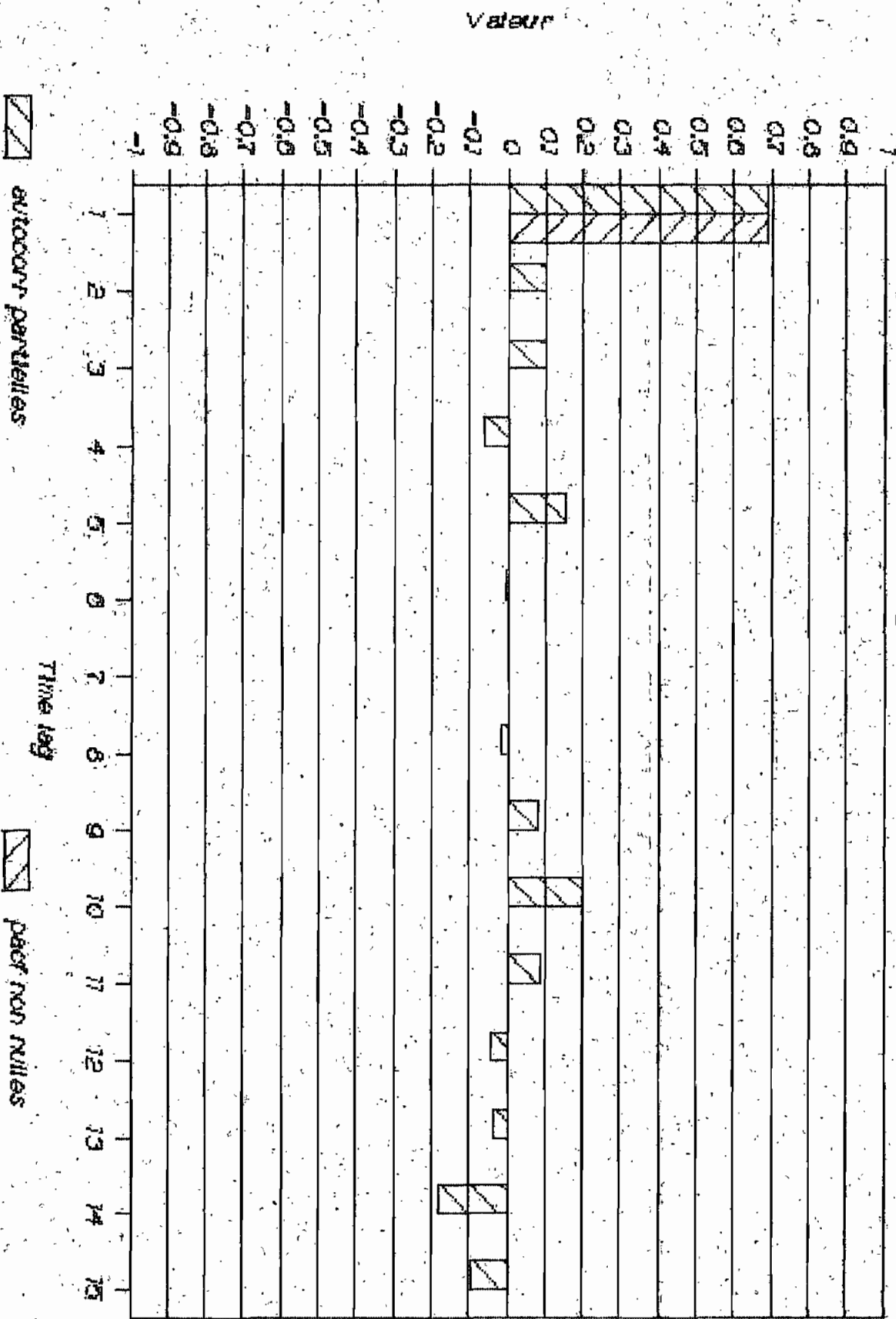
Time lag



pecc non nulles

Fonctions d'autocorrélations partielles

Casi : Change in business inventories



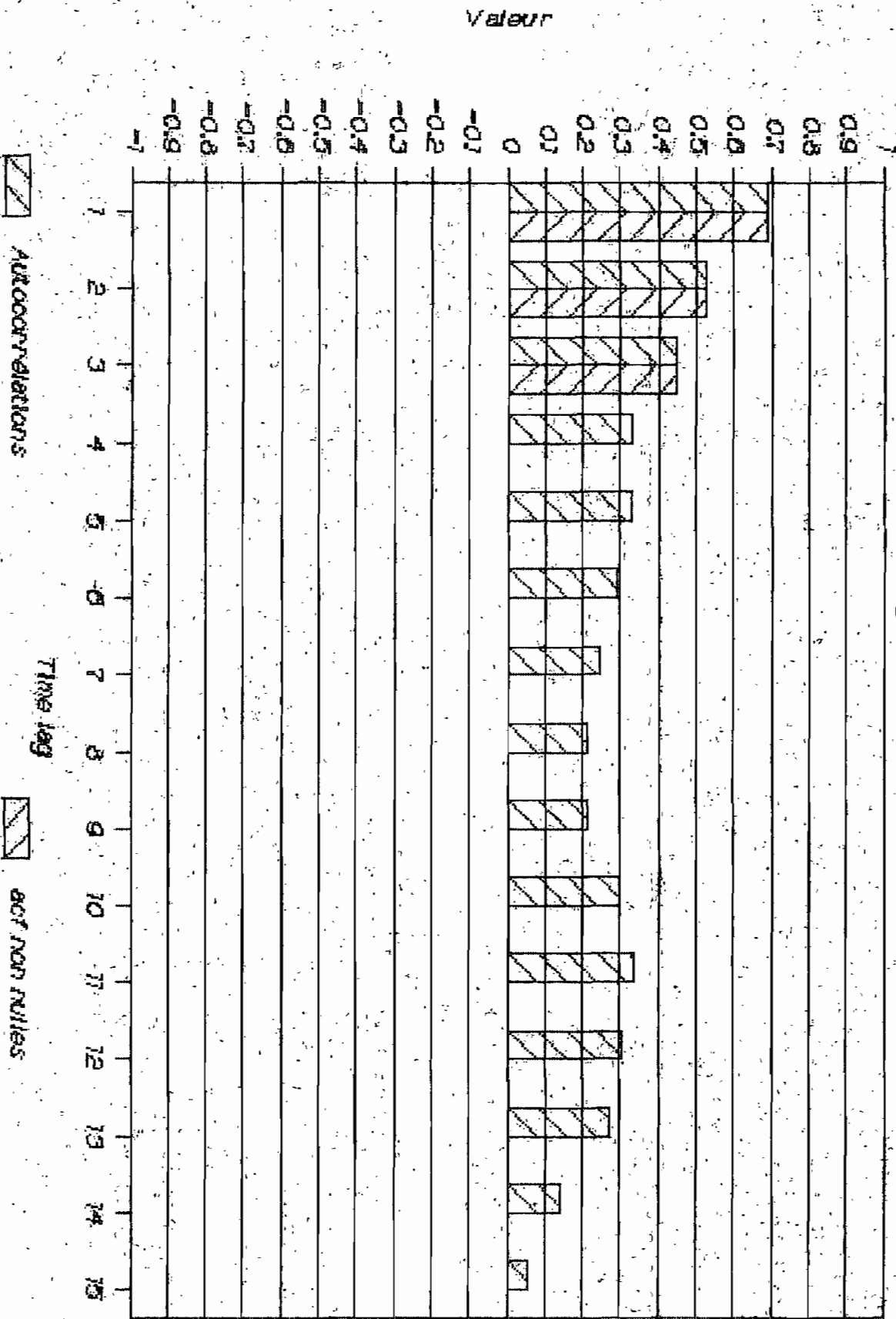
autocorr partielles



pas de non nulles

Fonctions d'autocorrélations

Case: Change in business inventories



Fonctions d'autocorrélations

Cours de l'or à NEW YORK. Année 1982

