MARC MOORE
YVES LEMAY

Sur le principe de parcimonie dans l'ajustement des modèles ARMA

Statistique et analyse des données, tome 14, nº 2 (1989), p. 81-95
<http://www.numdam.org/item?id=SAD_1989__14_2_81_0>

Statistique et Analyse des Données
1989 - Vol. 14 n° 2 - p. 81 - 95

SUR LE PRINCIPE DE PARCIMONIE DANS L'AJUSTEMENT DES MODELES ARMA

Marc MOORE
Yves LEMAY

Département de mathématiques appliquées
Ecole Polytechnique de Montréal
C.P. 6079, Succ. A
Montréal, Qc, Canada, H3C 3A7
Bell Canada
Montréal, Qc

Résumé : *Lors du choix d'un modèle ARMA il importe de ne retenir que les paramètres essentiels. Il est alors d'usage de vérifier, à l'aide de tests statistiques, si l'estimation de chaque paramètre diffère d'une façon significative de zéro. Nous examinons l'influence du seuil d'un tel test sur la qualité de l'ajustement du modèle obtenu, et aussi relativement à la précision des prévisions calculées à partir de ce modèle. Des simulations et une étude de cas semblent indiquer que le seuil de 5%, souvent utilisé, peut parfois être trop petit.*

Abstract : *In the building of an ARMA model it is important to use only the parameters that are necessary. It is then usual to apply a statistical test to the estimation of each parameter and to retain only those that are significantly different of zero. We study the influence of the level of these tests on the adjustment of the model obtained and also on the precision of the forecasts computed from this model. Some simulations and a case study seem to indicate that the popular 5% level might be too low here.*

Mots clés : Modèle ARMA, parcimonie, prévision, niveau de signification.

Indices de classification STMA : 12-060, 12-070, 12-080.

Manuscrit reçu le 4 mai 198, révisé le 20 mars 1989

0 - INTRODUCTION

Les techniques dites "de Box-Jenkins" pour l'analyse des séries chronologiques jouissent d'une grande popularité auprès des utilisateurs. Cet intérêt est sans doute motivé par la valeur intrinsèque de l'outil, mais aussi par sa relative facilité d'application et l'existence de logiciels performants. Un des principes importants dans l'approche proposée par Box et Jenkins [1] est celui de la parcimonie dans l'introduction des paramètres : un modèle ne doit inclure que les paramètres nécessaires. La motivation de ce principe réside en particulier dans le désir d'obtenir de bonnes estimations pour ces paramètres. Lors du choix d'un modèle, l'application de ce principe requiert l'attention de l'utilisateur à deux stades :

a) La forme la plus "économique" d'un modèle doit être utilisée. Par exemple, un modèle de type moyenne mobile de degré un, MA(1), peut parfois remplacer un modèle autorégressif (AR) de degré plus élevé.

b) Une fois la forme du modèle choisie, on ne doit conserver que les paramètres qui sont significativement différents de zéro.

On trouve dans la littérature certains renseignements illustrant l'importance de la parcimonie au stade a). Par exemple, Ledolter et Abraham [4] ont montré que tout paramètre introduit inutilement contribue à augmenter de σ^2/n la variance de l'erreur inhérente à une prévision d'horizon un ; n est le nombre d'observations disponibles et σ^2 dénote la variance associée au processus de type bruit blanc sous-jacent au processus ARMA. Mentionnons aussi que Kunitomo et Yamamoto [3] ont étudié l'effet, sur l'estimation des paramètres et sur l'erreur quadratique moyenne associée aux prévisions, d'un mauvais choix du modèle ARMA. Par exemple, ils montrent que si le modèle exact est un AR(m) et que le modèle choisi est un AR(p), $p > m$, alors l'erreur quadratique moyenne associée aux prévisions obtenues avec le modèle choisi est toujours plus grande que celle associée aux prévisions obtenues avec le modèle exact ; toutefois ce n'est pas nécessairement le cas si $p < m$.

Le souci de parcimonie au stade b) se traduit en général par l'application de la procédure suivante. Une fois un modèle choisi et ses paramètres estimés, un test statistique est effectué pour chacun d'eux afin de juger s'ils sont significativement différents de zéro. En ne conservant dans le modèle que les paramètres ayant été jugés significatifs, on reprend l'estimation de ceux-ci. Ces nouvelles estimations sont à leur tour testées et la procédure se répète jusqu'à ce que toutes les estimations soient jugées significativement différentes de zéro. Dans ce contexte, considérons

α = P[conserver un paramètre lorsque celui-ci est égal à zéro]

β = P[ne pas rejeter l'hypothèse qu'un paramètre soit égal à zéro lorsque celui-ci n'est pas égal à zéro].

les erreurs de première et de deuxième espèce associées aux tests effectués à partir des estimations des paramètres. Le principe de parcimonie nous indique de choisir une petite valeur pour α (le seuil du test), mais la diminution de α entraîne l'augmentation de β. On sait déjà que la conservation d'un paramètre non requis peut avoir une influence sur la stabilité des prévisions. Par contre, négliger d'introduire un paramètre requis, en plus d'avoir un effet négatif sur la qualité de l'ajustement du modèle aux données, peut en principe générer des prévisions de moins bonne qualité. En effet, les prévisions calculées suivant la méthode Box-Jenkins sont sans biais et d'erreur quadratique moyenne minimale à la condition que le modèle utilisé soit exact, or en enlevant à tort un paramètre on risque de perdre ces propriétés. Quelle devrait alors être la valeur de α représentant un bon compromis? Souvent les auteurs d'ouvrages destinés aux utilisateurs, par exemple Pankratz [8], semblent recommander, au moins implicitement, qu'un paramètre ayant un rapport t (estimation divisée par son écart type) inférieur à deux doit être enlevé. Evidemment cette règle, très souvent suivie, correspond grosso modo à un test de seuil 5%. Nous croyons que cette règle provient en grande partie d'une habitude passablement ancrée chez les utilisateurs de méthodes statistiques. A ce sujet il est intéressant de lire des propos tenus par Lehmann [De Groot [2], pp. 255] ; "why do people use .05? - It's obviously a silly thing to do. You should take into account what power you can get. - ... they would do better to carry out the test at a somewhat higher significance level where they would have a better balance between the two kinds of error"...

Nous voulons nous adresser ici à l'application du principe de parcimonie au stade b), plus particulièrement à l'influence du choix de α sur la qualité de l'ajustement du modèle choisi et sur la qualité des prévisions. Des expériences de simulation furent conduites pour plusieurs modèles ARMA unidimensionnels, la description de celles-ci et des résultats types sont présentés à la section 1. A l'aide d'une banque de données à propos des trajectoires suivies par des icebergs, nous avons pu étudier empiriquement l'influence de α dans le cas de modèles ARMA bidimensionnels, certaines observations sont décrites à la section 2. De ces considérations il semble se dégager qu'un α de 5% peut parfois être trop petit et qu'il n'y a pas d'inconvénient à utiliser un α plus grand, on explicite cette constatation à la section 3.

1 - MODELES ARMA UNIDIMENSIONNELS

1.1 - Expérience de simulation

Considérons un modèle ARMA (p, q)

$$Z_t - \varphi_1 Z_{t-1} - \dots - \varphi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

où les paramètres $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$ sont connus et les chocs a_t sont indépendants et identiquement distribués suivant une loi gaussienne de moyenne 0 et variance 1. Etant donné un tel modèle, on génère 220 observations, $Z_1, \dots, Z_{220}$. Afin d'éliminer les effets initiaux possibles les 100 premières observations sont rejetées. Les 20 dernières observations sont mises de côté pour une étape ultérieure. On se place alors dans la situation où les degrés p et q sont connus mais où les paramètres $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$ sont inconnus, et on procède à leur estimation à partir des observations $Z_{101}, \dots, Z_{200}$ et en utilisant la méthode du maximum de vraisemblance (non conditionnelle) (Box et Jenkins [1]). Cette procédure permet d'obtenir, pour chaque paramètre τ, son estimation $\hat{\tau}$ et une estimation, $s(\hat{\tau})$, de l'écart-type de celle-ci. On décidera de conserver un paramètre τ si

$$|\hat{\tau} / s(\hat{\tau})| > r \qquad (1)$$

r étant une valeur pré-déterminée, par exemple $r = 1.96$ correspond grosso modo à un test de seuil 5%. Le test décrit en (1) sera dénoté par t_r. Une fois les paramètres à conserver déterminés, on effectue à nouveau l'estimation de ceux-ci en imposant la valeur 0 aux autres paramètres. Les nouvelles estimations sont à nouveau soumises au test (1) et on continue de cette façon jusqu'à ce que tous les paramètres soient conservés. Bien que dans la majorité des expériences effectuées, à chaque étape tous les paramètres non significativement différents de zéro furent enlevés, il serait théoriquement préférable de n'enlever qu'un paramètre à la fois (le moins significatif). Le praticien utilisant un logiciel fournissant les probabilités de dépassement (probabilités de signification) préférera une autre forme pour le critère (1), c'est-à-dire de conserver le paramètre si la probabilité de dépassement est inférieure à une probabilité prédéterminée $\alpha(r)$; les deux formes seront considérées dans la présentation des résultats.

Afin de juger de l'influence de r sur la qualité de l'ajustement du modèle obtenu nous considérons la statistique du portemanteau (Ljung et Box [6]),

$$Q = n\,(n+2) \sum_{k=1}^{K} \hat{r}_k^2 \,/\, (n-k)$$

où $\hat{r}_k$ est l'estimation de la corrélation à l'écart k dans les résidus ; ici $n = 100$ et nous considérons $K = 24$. Sous l'hypothèse de la validité du modèle obtenu, Q suit approximativement une loi khi-carré avec $\nu = K - b$ degrés de liberté, b étant le nombre de paramètres dans le modèle obtenu. La quantité Q étant connue, on calcule la probabilité qu'une variable aléatoire suivant une loi χ^2_ν prenne une valeur supérieure à Q ; dénotons par $d(r)$ cette valeur.

Le modèle obtenu peut être utilisé afin de calculer, à $t = 200$, les prévisions pour les variables $Z_{201}, \ldots, Z_{220}$. Puisque l'on connait les valeurs de ces variables, la qualité des prévisions et l'influence de r sur celles-ci peuvent être appréciées. Dénotons par $\hat{Z}^{(r)}_{200}(\ell)$ la valeur prédite à $t = 200$ pour ℓ unités de temps en avant, $\ell = 1, \ldots, 20$, lorsque le test t_r est utilisé. La précision de la prévision d'horizon ℓ sera exprimée par

$$v\,(r, \ell\,) = (\,\hat{Z}^{(r)}_{200}(\ell) - Z_{200+\ell}\,)^2.$$

La procédure décrite jusqu'ici est reprise N fois (d'une façon indépendante), toutes les quantités définies étant calculées chaque fois. Dénotons par $D(r)$ et $V(r,\ell)$ les moyennes des N valeurs de $d(r)$ et $v(r,\ell)$ respectivement. Dans une situation où le modèle est adéquat la valeur attendue de $D(r)$ est de 0.5. Lors de chaque simulation nous considérons aussi la procédure consistant à ne faire aucun test relativement aux estimations obtenues pour les paramètres. Dénotons dans ce cas par D et $V(\ell)$ les valeurs moyennes pour mesurer l'ajustement et la qualité des prévisions.

L'influence du test t_r sur l'erreur quadratique moyenne des prévisions d'horizon ℓ sera mesurée par

$$E(r,\ell) = 100\, \frac{[V\,(r,\ell) - (V\,(\ell)]}{V\,(\ell)}\ \%\,,\ \ell = 1, \ldots, 20.$$

Une valeur négative de $E(r,\ell)$ signifie qu'il y a un gain de précision en faisant le test t_r.

Les valeurs considérées pour r sont 1.96, 1.75, 1.5, 1.25 et 1, elles correspondent sous un modèle gaussien à des $\alpha(r)$ (seuils) de 5%, 8%, 13.4%, 21.1% et 31.7% respectivement. Lors de chaque répétition de l'expérience les valeurs des différentes quantités furent calculées pour chaque valeur de r. Dans chaque cas analysé 250 répétitions furent effectuées. Pour la génération des données et l'estimation des paramètres on utilisa le progiciel SAS [9]. La qualité de l'expérience fut évaluée en comparant les estimations des paramètres (lorsque l'on ne fait pas de test) aux valeurs des paramètres du modèle exact, et en utilisant la statistique du portemanteau afin de vérifier si la banque des 250 séries de 100 observations chacune était bien représentative du modèle simulé. Dans chacun des cas rapportés ici la qualité de l'expérience est très satisfaisante, c'est-à-dire que les estimations des paramètres sont toujours près des valeurs exactes de ceux-ci, et le test du portemanteau confirme fortement que les données correspondent bien au modèle considéré.

1.2 Exemples

Considérons d'abord des modèles de type ARMA (1, 1),

$$Z_t - \varphi_1 Z_{t-1} = a_t - \theta_1 a_{t-1},$$

les valeurs considérées pour les paramètres φ_1 et θ_1 apparaissant au Tableau I. La troisième colonne de ce tableau présente les valeurs de D et des $D(r)$.

Puisque tous les paramètres considérés sont différents de zéro, l'effet d'un test t_r peut être de conserver le bon type de modèle, un ARMA (1,1), ou d'engendrer un mauvais modèle qui peut être un MA(1), un AR(1) ou un bruit blanc. Les quatre dernières colonnes du Tableau I donnent le pourcentage de chaque type de modèle obtenu après l'application des tests. On constate qu'un test au seuil de 5% peut avoir un effet relativement important sur le degré d'ajustement aux observations, principalement si les paramètres sont petits en valeur absolue. Avec le test t_1, on retrouve presque le degré d'ajustement correspondant à la situation où aucun test n'est effectué. Lorsque les paramètres sont petits et que r est grand la procédure a trop tendance à sélectionner un mauvais modèle.

Le tableau II permet de constater l'effet de r sur la qualité des prévisions. Le test t_1 n'a jamais d'effet important. Il peut arriver qu'un test avec $r = 1.96$ ou 1.75 ait une influence négative sur l'erreur quadratique moyenne des prévisions à très court terme. Cette influence ne sera perceptible que pour des paramètres de grandeur moyenne.

TABLEAU I
Résultats relatifs a l'ajustement pour des modèles ARMA (1,1)

	TEST r (α(r))	D ou D (r)	% de ARMA (1,1)	% de MA (1)	% de AR (1)	% de ARMA (0,0)
φ = .1 θ = − .1	1.96 (.05)	0.42	12.8	7.6	3.6	76.0
	1.75 (.08)	0.43	14.4	8.8	6.0	70.8
	1.5 (.13)	0.44	18.0	8.0	6.4	67.6
	1.25 (.21)	0.47	21.2	8.4	9.6	60.8
	1 (.32)	0.49	25.6	11.2	14.0	49.2
	aucun	0.55	—	—	—	—
φ = .2 θ = − .2	1.96 (.05)	0.25	3.2	17.6	16.0	63.2
	1.75 (.08)	0.29	6.0	18.8	20.0	55.2
	1.5 (.13)	0.34	9.2	19.6	25.2	46.0
	1.25 (.21)	0.42	13.2	24.4	31.2	31.2
	1 (.32)	0.47	18.4	29.2	33.2	19.2
	aucun	0.51	—	—	—	—
φ = .3 θ = − .3	1.96 (.05)	0.39	7.6	34.4	34.0	24.0
	1.75 (.08)	0.46	15.2	36.4	35.6	12.8
	1.5 (.13)	0.50	26.0	34.0	34.0	6.0
	1.25 (.21)	0.52	35.2	31.2	31.2	2.4
	1 (.32)	0.52	48.0	26.0	24.8	1.2
	aucun	0.52	—	—	—	—
φ = .5 θ = − .3	1.96 (.05)	0.49	56.0	6.0	38.0	0
	1.75 (.08)	0.51	64.8	4.4	30.8	0
	1.5 (.13)	0.51	70.8	4.0	25.2	0
	1.25 (.21)	0.52	78.0	3.6	18.4	0
	1 (.32)	0.52	86.8	1.2	12.0	0
	aucun	0.52	—	—	—	—

TABLEAU II
Résultats relatifs aux prévisions pour les modèles ARMA (1,1)

		E (r,ℓ) ; ℓ = 1, ..., 5				
	r (α(r))	1	2	3	4	5
$\varphi = .1$ $\theta = -.1$	1.96 (.05)	-0.16	0.74	-0.41	-0.04	0.15
	1.75 (.08)	0.42	0.86	-0.60	-0.04	0.10
	1.5 (.13)	1.48	0.45	-0.49	0.07	-0.03
	1.25 (.21)	0.70	-0.41	-0.30	-0.05	0.01
	1 (.32)	-0.68	-0.29	-0.19	-0.25	-0.08
$\varphi = .2$ $\theta = -.2$	1.96 (.05)	6.21	-0.85	0.20	-0.30	-0.15
	1.75 (.08)	6.34	-0.58	0.45	-0.26	-0.23
	1.5 (.13)	8.41	0.71	0.59	-0.12	-0.23
	1.25 (.21)	4.39	0.84	0.40	-0.21	-0.18
	1 (.32)	3.29	0.62	0.19	-0.21	-0.11
$\varphi = .3$ $\theta = -.3$	1.96 (.05)	11.98	1.66	1.34	0.70	-0.23
	1.75 (.08)	4.59	0.24	0.54	0.39	-0.05
	1.5 (.13)	1.82	0.61	0.84	0.24	0.13
	1.25 (.21)	1.26	1.14	0.47	0.17	0.23
	1 (.32)	0.80	0.46	0.22	0.24	0.02
$\varphi = .5$ $\theta = -.3$	1.96 (.05)	2.93	-0.15	1.33	0.73	0.63
	1.75 (.08)	1.61	-0.43	0.36	0.68	1.01
	1.5 (.13)	0.43	-0.72	0.05	0.26	0.38
	1.25 (.21)	-0.04	-0.46	-0.04	0.14	0.29
	1 (.32)	0.21	-0.06	-0.15	-0.11	-0.05

Il semble donc que si les valeurs des paramètres sont relativement grandes, le seuil du test n'a pas d'influence sur la qualité de l'ajustement ni sur la qualité des prévisions. Par contre, dans la situation opposée, un test avec r plus près de 1 que de 2 peut nous garantir contre un moins bon ajustement et des prévisions à court terme moins précises. Les mêmes constatations furent faites à partir de modèles de type AR(2),

$$Z_t - \varphi_1 Z_{t-1} - \varphi_2 Z_{t-2} = a_t \tag{2}$$

où $\varphi_1 \neq 0$ et $\varphi_2 \neq 0$.

Dans tous les exemples considérés jusqu'ici l'effet possible des tests, effectués par souci de parcimonie, est de conduire possiblement à un modèle contenant moins de paramètres que le modèle exact (sous-paramétrisation). Afin d'analyser la situation inverse considérons un modèle AR(2) de la forme

$$Z_t - \varphi_2 Z_{t-2} = a_t,$$

c'est-à-dire que $\varphi_1 = 0$ en (2), et procédons à l'estimation des paramètres comme si les données générées provenaient d'un modèle AR(2) sans contrainte a priori sur les paramètres. Les tests t_r peuvent alors avoir pour effet de conserver à tort le paramètre φ_1, ce qui conduit à une surparamétrisation si φ_2 est aussi conservé. Les résultats obtenus pour deux valeurs de φ_2 sont présentés aux Tableaux III et IV. On constate que pour une petite valeur de $|\varphi_2|$, un test avec r élevé a trop tendance à nettoyer complètement le modèle pour en faire un bruit blanc, entrainant un ajustement global inférieur à celui correspondant à la surparamétrisation dûe à l'absence de test.

TABLEAU III
Résultats relatifs à l'ajustement des modèles AR (2) avec $\varphi_1 = 0$.

	TEST r (α(r))	D ou D (r)	% de AR (2) avec $\varphi_1 = 0$	% de AR (2)	% de AR (1)	% de AR (0)
$\varphi_1 = 0$ $\varphi_2 = .2$	1.96 (.05)	0.50	42.4	1.6	3.2	52.8
	1.75 (.08)	0.52	50.4	2.0	4.0	43.6
	1.5 (.13)	0.53	56.0	5.2	6.0	32.8
	1.25 (.21)	0.55	56.8	16.0	7.2	20.0
	1 (.32)	0.57	54.8	24.4	8.0	12.8
	aucun	0.55	—	—	—	—
$\varphi_1 = 0$ $\varphi_2 = -.5$	1.96 (.05)	0.53	96.4	3.6	0	0
	1.75 (.08)	0.53	91.6	8.4	0	0
	1.5 (.13)	0.54	88.4	11.6	0	0
	1.25 (.21)	0.54	79.6	20.4	0	0
	1 (.32)	0.55	67.2	32.8	0	0
	aucun	0.53	—	—	—	—

Lorsque $|\varphi_2|$ augmente nous constatons que si r diminue le test conduit moins souvent au bon modèle alors que le degré global d'ajustement mesuré par D (r) demeure constant ou augmente légérement, cela peut sembler paradoxal. En examinant la situation de plus près nous avons constaté que ce phénomène s'explique par le fait qu'il y a des cas où un modèle AR(2) avec $\varphi_1 = 0$, détecté par $t_{1.96}$, apporte un ajustement inférieur à un modèle AR(2) complet, et ces cas sont progressivement abandonnés (en faveur d'un AR(2) complet) lorsque r augmente. Au niveau des prévisions, la valeur de r ne semble pas très importante ici.

TABLEAU IV
Résultats relatifs à l'ajustement des modèles AR (2) avec $\varphi_1 = 0$.
(—— indique une valeur négligeable)

	r (α(r))	E (r, ℓ) ; ℓ = 1, ..., 5: 1	2	3	4	5
$\varphi_1 = 0$, $\varphi_2 = .2$	1.96 (.05)	-0.89	0.52	0.37	0.40	0.08
	1.75 (.08)	-0.79	0.33	0.60	0.25	0.08
	1.5 (.13)	0.48	-1.14	0.74	0.19	0.05
	1.25 (.21)	-0.63	-0.42	0.25	0.09	——
	1 (.32)	-1.30	-0.07	-0.10	0.03	0.04
$\varphi_1 = 0$, $\varphi_2 = -.5$	1.96 (.05)	-0.82	-0.47	-0.83	0.58	0.15
	1.75 (.08)	-0.90	0.05	-0.78	0.69	0.05
	1.5 (.13)	-1.52	-0.07	-0.24	0.42	-0.05
	1.25 (.21)	-1.04	-0.62	-0.24	——	-0.11
	1 (.32)	-0.11	-0.42	-0.24	-0.17	-0.09

Des modèles de type MA (2) et ARMA (2, 2) furent également analysés et on observa les mêmes phénomènes que ceux illustrés plus haut.

2 - MODELES ARMA MULTIDIMENSIONNELS

Tiao et Box [11] indiquent comment des modèles ARMA multidimensionnels peuvent être ajustés à des données. Dans le cas d'un vecteur aléatoire de dimension k, U_t, un tel modèle s'écrit

$$U_t - \varphi_1 U_{t-1} - \dots - \varphi_p U_{t-p} = \mu + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \tag{3}$$

où : $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$ sont des matrices $k \times k$, μ est un vecteur composé de k constantes, les vecteurs aléatoires a sont indépendants et suivent une loi multinormale avec vecteur moyen nul et matrice de covariance Σ. Comme pour les modèles ARMA unidimensionnels les modèles (3) trouvent beaucoup d'applications. Par exemple, Moore [7] indique que de tels modèles bidimensionnels $(k = 2)$ peuvent parfois être utilisés pour représenter les trajectoires suivies par des icebergs. Dénotons par (X_t, Y_t) la position d'un iceberg au temps t, cela par rapport à un point fixe (par exemple une plateforme de forage) et un système d'axes orthogonaux donné. Une fois les composantes X_t et Y_t correctement différenciées (pour obtenir la stationnarité) un modèle ARMA peut être ajusté et celui-ci utilisé pour prédire le déplacement du iceberg. Dans le cas bidimensionnel chaque paramètre matriciel introduit apporte quatre paramètres, la parcimonie est donc de rigueur.

Lors de notre étude sur le déplacement des icebergs nous disposions de données relatives à plusieurs icebergs (Whitmore et Gentleman [12]). Nous voulons illustrer ici, dans le cas de trois icebergs, l'influence d'un test t_r, utilisé tel que décrit à la section 1 et cela pour chacune des composantes des paramètres matriciels. Nous considérons l'influence sur la qualité de l'ajustement et sur la précision des prévisions.

L'analyse proposée est la suivante. Si pour un iceberg on dispose de n observations consécutives, (x_t, y_t), au sujet de sa trajectoire, on utilise les $(4/5)n$ premières observations pour ajuster un modèle ARMA et avec celui-ci on prédit les observations restantes. La qualité de l'ajustement est évaluée en examinant les matrices de corrélations croisées calculées à partir des résidus, et en calculant la statistique du portemanteau proposée par Li et McLeod [5]. Afin d'évaluer la capacité du modèle à prévoir la position du iceberg nous calculons

$$SDA_x = \sum_{\ell=1}^{n-t_0} \left| \hat{X}_{t_0}(\ell) - x_{t_0+\ell} \right|$$

$$SDA_y = \sum_{\ell=1}^{n-t_0} \left| \hat{Y}_{t_0}(\ell) - y_{t_0+\ell} \right|$$

$$SD = \sum_{\ell=1}^{n-t_0} \left[\left(\hat{X}_{t_0}(\ell) - x_{t_0+\ell} \right)^2 + \left(\hat{Y}_{t_0}(\ell) - y_{t_0+\ell} \right)^2 \right]^{1/2},$$

t_0 étant le temps correspondant à la dernière observation utilisée pour ajuster le modèle.

Les tableaux V, VI et VII présentent les résultats obtenus pour les trois icebergs considérés. Trois scénarios furent envisagés, ne faire aucun test et faire des tests avec $r = 1$ ou $r = 2$. Tous les calculs furent effectués à l'aide du progiciel SCA [10] qui est particulièrement bien adapté pour l'analyse des modèles ARMA multidimensionnels. La deuxième colonne indique le nombre de paramètres apparaissant dans le modèle, on constate que l'application d'un test réduit considérablement le nombre de paramètres. A la troisième colonne on mentionne les écarts auxquels on retrouve des corrélations significatives dans les résidus (autocorrélations ou corrélations croisées), et la probabilité de dépassement (probabilité de signification) associée à la statistique du portemanteau. Finalement, à la dernière colonne on trouve les valeurs de SDA_x, SDA_y et SD.

TABLEAU V
Analyse de l'iceberg A
n = 99, 79 observations utilisées pour l'ajustement, modèle ARMA (0,2)

TEST	NOMBRE DE PARAMETRES	QUALITE DE L'AJUSTEMENT	QUALITE DES PREVISIONS (Km)
t_1	4	7ième , 12ième 0.21	$SDA_x = 44.0$ $SDA_y = 39.8$ $SD = 60.8$
t_2	1	N.D	N.D
aucun	10	7ième 0.22	$SDA_x = 55.8$ $SDA_y = 32.7$ $SD = 65.6$

(N.D. : non disponible, voir le texte)

TABLEAU VI
Analyse de l'iceberg B
n = 81, 65 observations utilisées pour l'ajustement, modèle ARMA (5,0)

TEST	NOMBRE DE PARAMETRES	QUALITE DE L'AJUSTEMENT	QUALITE DES PREVISIONS (Km)
t_1	15	9ième , 20ième 0.15	$SDA_x = 18.5$ $SDA_y = 15.1$ $SD = 24.7$
t_2	7	9ième , 19iéme 0.11	$SDA_x = 18.9$ $SDA_y = 23.6$ $SD = 31.1$
aucun	22	9ième , 20ième 0.22	$SDA_x = 29.0$ $SDA_y = 17.4$ $SD = 34.9$

TABLEAU VII
Analyse de l'iceberg C
n = 78, 65 observations utilisées pour l'ajustement, modèle ARMA (2,0)

TEST	NOMBRE DE PARAMETRES	QUALITE DE L'AJUSTEMENT	QUALITE DES PREVISIONS (Km)
t_1	5	3iéme 0.63	SDA_x = 5.8 SDA_y = 11.3 SD = 12.8
t_2	4	3iéme 0.39	SDA_x = 10.1 SDA_y = 10.6 SD = 14.7
aucun	10	—— 0.94	SDA_x = 2.4 SDA_y = 5.7 SD = 6.8

De l'analyse de ces résultats, et de ceux obtenus pour plusieurs autres icebergs, il semble se dégager les points suivants :

i) Il peut s'avérer dangereux de trop réduire le nombre de paramètres en utilisant un test de faible seuil. Des cas furent rencontrés où une trop grande valeur de r entraînait la non convergence de l'algorithme d'estimation, ce fut le cas avec t_2 pour l'iceberg A.

ii) Plus souvent SD a une plus grande valeur avec t_2 qu'avec t_1.

iii) La qualité de l'ajustement est en général meilleure avec t_1 qu'avec t_2.

iv) Les écarts-type associés aux prévisions (non rapportés ici) sont en général inférieurs si t_1 est utilisé plutôt que t_2.

Il nous semble donc que dans ce cas, l'utilisation du test t_1 représente un compromis valable par rapport aux deux situations extrêmes consistant à ne pas faire de test d'une part et à appliquer le test t_2 d'autre part. Ce compromis semble avantageux relativement à la qualité des prévisions et acceptable au niveau de l'ajustement du modèle et de la stabilité des prévisions.

3 - CONCLUSION

Cette étude ne répond bien sûr pas à la question soulevée dans l'introduction relativement à la valeur de α représentant un bon compromis. Toutefois, des simulations conduites et des analyses de données réelles effectuées, il se dégage clairement que l'utilisation d'un test de seuil 5% peut parfois présenter des inconvénients au niveau de la qualité de l'ajustement et de la qualité des prévisions ; par contre l'utilisation d'un test de seuil de l'ordre de 30% n'a jamais présenté d'inconvénients importants. Bien sûr l'utilisation d'un seuil de 5% n'a jamais généré de résultats catastrophiques, mais nous croyons utile de savoir que lors d'un test au sujet d'un paramètre dont la valeur est relativement petite, il semble préférable d'utiliser un critère de la forme "plus ou moins un écart-type" plutôt que "plus ou moins deux écarts-type".

REMERCIEMENTS

Ce travail fut réalisé grâce aux subventions OGP0008211 du CRSNG et CRP-2093 du Fonds FCAR. Les auteurs sont reconnaissants envers les deux rapporteurs qui ont aidé, par leurs remarques constructives, à clarifier plusieurs points.

BIBLIOGRAPHIE

[1] **Box, G.E.P., Jenkins, G.M.**, Time series analysis : Forecasting and control.
Revised edition, Holden-Day, San Francisco, 1976.

[2] **DeGroot, M.H.**, A conversation with Erich L. Lehmann, Statistical Science, 1986, Vol. 1, n° 2, pp. 243-258.

[3] **Kumitomo, N., Yamamoto, T.**, Properties of predictors in misspecified autogressive time series models, Journal of the American Statistical Association, 1985, Vol. 80, n° 392, pp. 941-950

[4] **Ledolter, J., Abraham, B.**, Parsimony and its importance in time series forecasting, Technometrics, 1981, Vol. 23, n° 4, pp. 411-414.

[5] **Li, W.K., McLeod, A.I.**, Distribution of the residual autocorrelation in multivariate ARIMA time series models, J. Roy. Statist. Soc. Ser. B, 1981, Vol.43, n° 2, pp. 231-239.

[6] **Ljung, G.M., Box, G.E.P.**, On a measure of lack of fit in time series models,
Biometrika, 1978, Vol. 65, n° 2, pp. 297-303.

[7] **Moore, M.**, Modeling iceberg motion : a multiple-time series approach,
La Revue Canadienne de Statistique, 1985, Vol. 13, n° 2, pp. 88-94.

[8] **Pankratz, A.**, Forecasting with Univariate Box-Jenkins Models Concepts and Cases, John Wiley and Sons, New York, 1983.

[9] SAS/ETS User's Guide, SAS Institute Inc., Cary, North Carolina, 1982.

[10] SCA-Scientific computing Associates, De Kalb. Illinois, 1985.

[11] **Tiao, G.C., Box, G.E.P.**, Modeling multiple time series with applications, Journal of the American Statistical Association, Vol. 76, n° 376, pp. 802-816.

[12] **Whitmore, G.A., Gentleman, J.F.**, Iceberg paths and collision risks for fixed marine structures, La Revue Canadienne de Statistique, 1985, Vol. 13, n° 2, pp. 84-87.