Pertanika J. Sci. & Technol. 10(1): 25-33 (2002)

ISSN: 0128-7680
© Universiti Putra Malaysia Press

The Performance of AICC as an Order Selection Criterion in ARMA Time Series Models

Liew Khim Sen & Mahendran Shitan[a]

Department of Mathematics
Universiti Putra Malaysia
43400 UPM, Serdang
Selangor, Malaysia

Received: 4 December 2000

ABSTRAK

Kajian ini bertujuan untuk menilai prestasi kriteria maklumat Akaike diperbaiki atau AICC (Akaike's Information Corrected Criterion) sebagai kriteria penentuan peringkat dalam pembentukan model Autoregresif Purata Bergerak (Autoregressive Moving-Average) atau ARMA $(p.q)$. Suatu penyelidikan simulasi dijalankan untuk menentukan kebarangkalian kriteria AICC minimum telah memilih model *sebenar* dengan tepat. Keputusan yang diperolehi menunjukkan bahawa prestasi AICC adalah sekadar sederhana. Masalah lebihan pembolehubah (over pemeterization) berada pada tahap yang minimum. Oleh itu, bagi sebarang dua model yang setanding, adalah lebih wajar untuk memilih model dengan peringkat p dan q yang lebih rendah.

ABSTRACT

This study is undertaken with the objective of investigating the performance of Akaike's Information Corrected Criterion (AICC) as an order determination criterion for the selection of Autoregressive Moving-Average or ARMA (p,q) time series model. A simulation investigation was carried to determine the probability of the AICC statistics picking up the *true* model. Result obtained showed that the probability of the AICC criterion picking up the correct model was moderately good. The problem of over parameterization existed but under parameterization was found to be minimal. Hence, for any two comparable models, it is always safe to choose the one with lower order of p and q.

Keywords: AICC, ARMA, under/over parameterization

INTRODUCTION

In the process of time series autoregressive moving-average or ARMA (p, q) modelling, we do not know the *true* order of the model generating the data. In fact it will usually be the case that there is no true ARMA (p, q) model, in which case our goal is simply to find one that represents the data optimally in some sense (Brockwell and Davis 1996). However, the challenge is to decide the optimal orders of p and q (Beveridge and Oickle 1994). In a given application,

Note: We are indebted to two anonymous referees for their helpful comments and suggestions on a previous draft of this paper. All errors remain our responsibility

the Box-Jenkins model selection procedure may suggest several specifications, each of which satisfies the diagnostic checks. Some kind of a measure of goodness of fit is therefore needed to distinguish between different models in these circumstances (Harvey 1993). Many criteria have been suggested for this reason by the past researchers. The Akaike's information corrected criterion (Hurvish and Tsai 1989) or AICC, among others, is a commonly used criterion. However, its performance must be evaluated. Therefore, the objective of the study is to evaluate the performance of AICC statistics in selecting the true ARMA time series model based on a simulation study.

The rest of this paper is organised as follows. The next section discussed the order determination criterion. This is followed by a description of simulation study and a report of simulation result. Finally, the conclusions of the study are presented.

ORDER DETERMINATION CRITERIA

Many criteria has been proposed for the purpose order determination by past researcher. These include the final prediction error (EPE) criterion, Schwarz-Rissanen criterion (SIC), Bayesian estimation criterion (BEC), Hannan- Quin criterion, Akaike's information criterion (AIC) and so on. The latest model selection criterion is the Akaike's information corrected criterion AICC, developed by Hurvish and Tsai in 1989.

There has been considerable literature published on order determination criteria. A brief discussion of these criteria is available in Beveridge and Oikle (1994); de Gooijer *et al.* (1985) and Stoica *et al.* (1986). Brockwell and Davis (1996) present greater theoretical and practical detail and additional references for many of these criteria.

The final prediction error, RPE criterion was original proposed by Akaike (1969, 1970) for AR (p) order determination and was extended to ARMA (p,q) models by Söderströ, in 1977 (Beveridge and Oickle 1994). This criterion was established on the basis of minimizing the one-step-ahead mean square forecast error after incorporating the inflating effects of estimated coefficients. The criterion to be minimized is

$$\text{FPE} = \hat{\sigma}^2 \frac{n+p+q}{n-p-q} \tag{1}$$

Where $\hat{\sigma}^2$ is estimated variance of white noise,
n is number of observation,
p is order of the autoregressive component,
and q is order of the moving average component.

In 1970, Akaike found that FPE is asymptotically inconsistent and in 1973 he employed information-theoretic considerations to develop the Akaike's information criterion, AIC. This was designed to be an asymptotically unbiased estimate of the Kullback-Leibler index of the fitted model relative to the true model (Akaike 1973). The AIC statistics is defined as

$$\text{AIC} = -2 \text{ In Likelihood } (\hat{\phi}, \hat{\theta}, \hat{\sigma}^2) + 2(p + q + 1) \tag{2}$$

where $\hat{\phi}$ are estimated autoregressive parameters,
$\hat{\theta}$ are estimated moving average parameters,
and $\hat{\sigma}^2$, n, p and q are as defined in equation (1).

A criterion like AIC that penalizes the likelihood for the number of parameters in the model attempts to choose the most parsimonous model. However, AIC is only asymptotically unbiased and Jones (1975) and Shibata (1976) showed empirical evidence that AIC has the tendency to pick models which are over-parameterized. In view of this, Akaike applied a Bayesian modification to AIC and finally in 1978, he came up with a consistent order selection criterion, known as Bayesian information criterion or BIC (see Akaike 1979). If the data $\{X_1, ..., X_n\}$ are in fact observations of an ARMA (p, q) process, then a Bayesian information criterion is defined to be

$$\text{BIC} = (n - p - q)\ln \frac{n\hat{\sigma}^2}{n-p-q} + n(1 + \ln 2n) + (p+q)\ln \left[\frac{\sum_{t=1}^{n} X_t^2 - n\hat{\sigma}^2}{p+q} \right] \tag{3}$$

There is evidence to suggest that the BIC is more satisfactory than the AIC as an ARMA model selection criterion since the AIC has a tendency to the pick models, which are over-parameterized (Hannan 1980).

Schwarz (1978) used a Bayesian analysis and Rissanen (1978) applied an optimal data-recording scheme to independently arrive at the same criterion, later known as Schwarz-Rissanen criterion, SIC. The criterion to be minimized is given by

$$\text{SIC} = \ln \hat{\sigma}^2 + (\frac{p+q}{n}) \ln n \tag{4}$$

Geweke and Mease (1981) suggested approximating SIC by Bayesian estimation criterion, BEC.

$$\text{BEC} = \hat{\sigma}^2 + (p_x + q_x)\hat{\sigma}^2{}_x \ln \frac{n}{n - p_x - q_x} \tag{5}$$

where x denotes a quantity from pre-assigned high order ARMA model that includes all potential models.

Hannan and Quinn (1979) and Hannan (1980) constructed Hannan-Quinn criterion from the law of the iterated logarithm. It provides a penalty function, which decreases as fast as possible for a strongly consistent estimator, as sample size increases. Hannan-Quin criterion is given by

$$HQ = \ln \hat{\sigma}^2 + 2(p+q)\frac{\ln(\ln n)}{n} \qquad (6)$$

Hannan and Rissanen (1982) replace the term Ln (lnn) by lnn to speed up the covergence of HQ. This revised version of HQ, however, was found to overestimate the model orders (Kavaleris 1991).

In 1989, Hurvish and Tsai found that BIC, which was modified from AIC, is not asymptotically efficient. Hence, they suggested a biased corrected version of AIC, known as Akaike's information corrected criterion or AICC. AICC statistic is given by

$$AICC = -2 \ln \mathit{Likelihood}\,(\hat{\phi},\hat{\theta},\hat{\sigma}^2) + [2n(p+q+l)]/[n-(p+q)-2] \qquad (7)$$

where $\hat{\phi}$ are estimated autoregressive parameters,
$\hat{\theta}$ are estimated moving average parameters,
$\hat{\sigma}^2$ is estimated variance of white noise,
n is number of observations,
p is order of the autoregressive component,
q is order of the moving average component,

and *Likelihood* $(\hat{\phi},\hat{\theta},\hat{\sigma}^2)$ is the likelihood of the data under the Gaussion ARMA model with parameters $(\phi,\theta,\hat{\sigma}^2)$.

The penalty factors $2n(p+q+1)/[n-(p+q)-2]$ and $2(p+q+1)$, for AICC statistics and AIC statistics respectively, are asymptotically equivalent as $n \rightarrow \infty$. Moreover, AICC, as AIC or PE, is asymptotically efficient for autoregressive process. The AICC statistics however, has a more extreme penalty for large order models, which counteract the over fitting nature of the AIC (Brockwell and Davis 1996). Today, the AICC statistics, as its earlier version(AIC), has been widely used as one of the order selection criteria in ARMA time series as well as the lag-length selection criteria in econometric modelling processes. Due to its popularity, Brockwell and Davis (1994) for instance, have included the AICC statistic in their computer software package known as "*Iterative Time Series Modelling (ITSM)*". As the AICC statistics is an important criterion for the selection of order in time series models, its performance must be evaluated. The study hence takes the initiative to explore the probability of minimum AICC criterion in picking up the *true* model based on a simulation study.

SIMULATION STUDY

In this study, a total of 10,000 simulated data series from 10 autoregressive moving average processes were investigated. These processes were AR(1), AR(2), AR(3), AR(4), MA(1), MA(2), ARMA(1,1), ARMA(1,2), ARMA(2,1)and ARMA(2,2). From there, 100 models were formulated in such a way that each process was assigned a number of 10 models. These models are summarized in the Appendix. For illustration, the 10 models for AR(1) process were those with a parameter ϕ value of 0.10. 0.30. 0.50, 0.70, 0.90, -0.30, -0.50, -0.60, -0.80 and -0.95 respectively. Each of these 10 models is in turn replicated into 100 random data series using a different random seed number (less than 10 digits) for each replication. To be consistent in comparison, every random series has 555 observations with a mean value of 111 and unit variance. No element of seasonality or trend is involved in this simulated data. The data series are randomly generated using the "Generation of the Simulated Data" option of the ITSM software.

The process of time series model fitting in this study involves identification of appropriate models, estimation of parameters and validation of the model. In the process of model fitting, ITSM automatically selected a minimum AICC model for each of the data series generated from the AR(1), AR(2), AR(3) and AR(4) processes. As for each of the remaining series, 4 to 9 appropriate models were fitted for model selection purpose. The estimated models are appropriate in the sense that, besides they are stationary and invertible, they are also required to pass the following formal diagnostic tests of randomness.

1. Ljung-Box portmanteau test, which uses the autocorrelations of the residuals to test for the null hyphotesis that the residuals are independently and identically distributed (iid);
2. Mcleod-Li portmanteau test, which tests whether the residuals are from an iid sequence of normally distributed random variables, by using the autocorrelations of the squarred-residuals;
3. Turning point test, which is normality test based on the number of turning points;
4. Different sign test, which is used to detect whether a linear trend (implies non-stationary) is present in the residuals;
5. Rank test, which is also a stationary rest for the residuals.

These test are easily checked by "Tests of Randomness of the Residuals" option in the software mentioned earlier. The order of the Yule-Walker model for the residuals is also estimated by this option, to asses whether the residuals of the each estimated model are compatible with the white plotting the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) are performed by the "Model ACF/PACF" option of ITSM software. The details on these diagnostic tests are available in Brockwell and Davis (1996). Out of a class of appropriate models, the order p and q of the minimum AICC model were recorded for each series.

If the estimated p and q of the minimum AICC model matches the simulated model, we say that the AICC criterion has picked up the correct model. If it failed to pick up the correct model, further investigation was carried out to determine whether over parameterization or under parameterization has occurred. Due to the fact that in the computation of AICC statistics the sum of p and q exceeding sum of the true order p and q, whereas under parameterization happened when sum of the true order p and q exceeding sum of the estimated order p and q. With these definitions, a minimum AICC model might fail to pick up the correct model, due to neither over parameterization not under parameterization, however. For instance, ARMA(1,2), ARMA(3,0) and ARMa(0,3) models were clearly different from ARMA(1,2) model, but neither of them was considered over parameterization or under parameterization. This paradox stemmed from the deficiency in the computation of AICC statistics, which regarded $p + q$ as one term. In this study, these models are treated as misspecified models.

In this study, for every 100 series of the same model, the probability that the minimum AICC model picks up the correct model, denote by P_c was computed as

$$P_c = \frac{\text{number of time “pick up” occurred}}{100}. \tag{8}$$

The probability that the event "over parameterization" happened, *Po was* calculated as

$$P_o = \frac{\text{number of time “over parameterization” occurred}}{100}. \tag{10}$$

Finally, the probability that the event "mis-specification" occurred, P_m was determined by

$$P_m = \frac{\text{number of time “mis-specification” occurred}}{100}. \tag{11}$$

SIMULATION RESULT

Amongst the 10 models of AR(1) process P_r ranged from 0.63 to 0.81 with a mean value of 0.721; P_o ranged from 0.19 to 0.37 with a mean value of 0.268, while P_u ranged from 0 to 0.99 with a mean value of 0.011. This mean that out of all the 1000 series of AR(1) process, the minimum AICC model matches the correct model 721 of the time; over parameterization occurs 268 of the time and under parameterization happens only 11 of the time. The result for AR(1) process and other procesesses in this study was summarized in Table 1. From this criterion, with a probability of picking the true model ranging from 0.366 to 0.795 and a mean value of 0.613. However, changes of over parameterization still exist and in very 100 models, around 17 to 50 models will be over parameterized. As compared to Autoregressive of Moving-Average models, over parameterization was found relatively serious in mixed Autoregressive Moving-Average models, where the AICC statistics could pick up at most 60 percent of the correct models. The AICC statistics in picking up the "mis-specified" model was negligible in only 4 out of 100 models (not shown). This result suggests that

TABLE 1
Summary of simulation's results

No	Process	Correctly estimated			Over parameterization			Under parameterization		
		Low	High	Mean	Low	high	Mean	Low	High	Mean
1	AR(1)	.63	.81	.721	.19	.37	.268	.00	.09	.011
2	AR(2)	.52	.84	.751	.16	.25	.219	.00	.25	.030
3	AR(3)	.60	.79	.714	.19	.32	.255	.00	.16	.031
4	AR(4)	.25	.78	.631	.15	.33	.233	.00	.60	.097
5	MA(1)	.43	.79	.670	.19	.41	.256	.00	.04	.005
6	MA(2)	.56	.84	.733	.16	.44	.265	.00	.00	.000
7	ARMA(1,1)	.20	.87	.601	.11	.80	.358	.00	.13	.013
8	ARMA(1,2)	.45	.74	.594	.26	.55	.406	.00	.00	.000
9	ARMA(2,1)	.01	.84	.320	.11	.71	.302	.00	.84	.246
10	ARMA(2,2)	.01	.65	.393	.22	.82	.413	.00	.62	.116
	Overall	.366	.795	.613	.174	.500	.298	.000	.273	.055

whenever the minimum AICC criterion failed to pick up the *true* model correctly, it was due to over parameterization. This fact that AICC over parameterized could be perceived as supportive to the proponents of parsimonous model such as Box and Jenkins (1976). Hence for any two comparable models, it is always safe to choose the one with lower order p and q.

CONCLUSION

The AICC statistics, as its earlier version (AIC) has been widely used as one of order selection criteria in ARMA time series as well as the lag-length selection criterion in econometric processes. As the AICC statistics is important in ARMA time series modelling and related fields, its performance must be evaluated. This paper evaluates the performance of AICC by determining the probability of the minimum AICC criterion in picking up the true model based on a simulation study. A total of 100 models from 10 ARMA processes were used in this study, with 100 replicants for each model giving to a total of 10,000 data series. The probability if interest was found to be only 0.613, even though we had use considerably large sample size. Hence, the performance of AICC in picking up the true models is expected to decline in the case of smaller sample size, which usually happens in empirical research. In addition, the minimum AICC criterion, which tries to overcome the over parameterization of the minimum AIC criterion, still has the tendency to overestimate the model orders. This implies that applying AICC criterion in either time series modelling or the selection of lag-length for any lag-length sensitive test such as unit root and cointegration test in the related fields would weaken the credibility of the ultimate result.

This study investigation only 10 of the commonly used ARMA (p,q) processes. It could be improved by including more variations of process, especially those with moderately high order, to produce a more influential result. The sample size could also be varied such that the actual performance of the minimum AICC criterion in conjunction with various sample size could be uncovered. A computer search algorithm could also be designed to determine a new empirically sound order selection criterion.

REFERENCES

AKAIKE, H. 1969. Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics* **21**: 243-247.

AKAIKE, H. 1970. Statistical predictor identification. *Annals of the Institute of Statistical Mathematics* **22**: 20-217.

AKAIKE, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In Peron, B.N. and Csaki, F. (eds), *2nd International Symposium in Information Theory*, p. 207-261. Budapest: Akademial Kiodd.

AKAIKE, H. 1979. A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrica* **66(2)**: 237-242.

BEVERIDGE, S. and C. OICKLE. 1994. A comparison Box-Jenkins and objective methods for determining the order of a non-seasonal ARMA model. *Journal of Forecasting* **13**: 419-434.

BOX, G. E. P. and G. M. JENKINS. 1976. *Time Series Analysis.* Revised edition. San Francisco: Holden-day.

BROCKWELL, P. J. and R. A. Bavis. 1994. *ITSM for Windows.* New York: Springer.

BROCKWELL, P. J. and R. A. Davis. 1996. *Introduction to Time Series and Forecasting.* New York: Springer.

DE GOOIJER, G., B. ABRAHAM, A. GOULD and L. ROBINSON. 1985. Methods for determining the order of an autoregressive moving average process: a survey. *International Statistical Review* **53**: 301-329.

GEWEKE, J. F. and R. A. MEASE. 1981. Estimating regression of finite but unknown order. *International Economic Review* **22**: 55-77.

HANNAN, E. J. 1980. The estimation of the order of an ARMA process. *Annals of Statistics* **8**: 1071-1081.

HANNAN, E. J. and B. G. QUINN. 1979. The determination of the order of an autoregression. *Journal of Royal Statistical Society* **41(2)**: 190-195.

HANNAN, E. J. and J. RISSANEN. 1982. Recursive estimation of mixed autoregressive moving average order. *Biometrica* **69(1)**: 81-94.

HARVEY, A. C. 1993. *Time Series Model.* 2nd ed. UK: Harvester Wheatsheaf.

HURVISH C. M and C. L. TSAI. 1989. Regression and time series model selection in small samples. *Biometrica* **76**: 297-307.

JONES, R. H. 1975. Fitting autoregressions. *Journal of American Statistics Association* **70**: 590-592.

KAVALIERIS, L. 1991. A note on estimating the dimension of a model. *The Annals of Statistics* **6(2)**: 461-464.

SHIBATA, R. 1976. Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrica* **63(1)**: 117-126.

STOICA, P., P. EYKHOFF, P. JANSEN and T. SÖDERSTRÖM. 1986. Model selection by cross validation. *International Journal of Control* **43**: 1841-1878.

RISSANEN, J. 1978. Modelling by shortest data description. *Automatica* **14**: 467-471.