FACTS AND FICTION IN SPECTRAL ANALYSIS OF STATIONARY STOCHASTIC PROCESSES

P.M.T. Broersen
Department of Applied Physics Delft University of Technology
P.O.Box 5046 2600 GA Delft The Netherlands
Phone + 31 15 2786419, Fax + 15 31 2784263, E-mail: broersen@tn.tudelft.nl

ABSTRACT:

New developments in time series analysis can be used to determine a better spectral representation for unknown data. Any stationary process can be modeled accurately with one of the three model types: AR (autoregressive), MA (moving average) or the combined ARMA model. Generally, the best type is unknown. However, if the three models are estimated with suitable methods, a single time series model can be chosen automatically in practice. The accuracy of the spectrum, computed from this single AR-MA time series model, is compared with the accuracy of many tapered and windowed periodogram estimates. The time series model typically gives a spectrum that is better than the best of all periodogram estimates.

1. INTRODUCTION

Two general methods for spectral analysis are parametric (or time series models) and non-parametric (or tapered and windowed periodograms) [1]. The choices for the type and the length of spectral or lag windows in periodogram spectra have only been developed for known spectra [1]. No statistical rules or window selection criteria are available for an optimal choice of windows or tapers in unknown data.

For stationary stochastic processes, at least one of the three parametric time series model types gives a good spectral description of the data. Any stationary stochastic process with a continuous spectral density can be written as an unique AR(∞) or MA(∞) process [1]. In practice, finite order MA or AR models for those infinite order processes are often accurate, because the true parameters decrease rapidly for most processes. Order selection for AR models has been studied with asymptotic criteria [1,2] and with finite sample equivalents [3,4,5,6] that are better if models of high orders are considered. For MA and ARMA models, a new development in time series analysis was necessary to have reliable estimation algorithms that perform well for all sample sizes [7,8,9,10]. That is the discovery of the optimal length of the long autoregressive intermediate model for Durbin's methods [7,8]. That long AR model is used to determine the MA parameters. With a sliding window technique for the practical selection of the long AR order, Durbin's improved methods yield accurate spectra.

So far, no automatic method for the selection of the model type, AR MA or ARMA, gives good results on unknown statistical data. A new criterion with the prediction error is given that can be used for the selection of the model type. Afterwards, the data spectrum is computed from the parameters of that *model, with selected type and order*. In simulations, an objective measure is used to compare the spectrum of this single model to a variety of tapered and windowed periodograms of the data. It is a fact that a single good time series model can be selected for unknown statistical data if suitable estimation algorithms have been used for the AR, MA and ARMA models. It is also a fact that even the best windowed periodogram is less accurate than the spectrum of the automatically selected single time series model.

2. FACTS and FICTION

The Fourier transform of a stationary stochastic process doesn't exist [1, p. 15]. More specific, the Fourier transform is not approximated better by taking more observations. This has influence on the square of its absolute value: the periodogram. It is fiction that looking at periodograms obtained from the same data with different types or lengths of windows and tapers may yield a statistically *reliable* choice between the alternatives. In those circumstances, a limited amount of details in the spectra is preferred by most people; that choice is not influenced by the level of details in the true spectrum. Theoretically, the choice of window type and size can only be made on deterministic grounds and it requires the exact knowledge of the true spectrum [1]. A periodogram can also be written as a finite order invertible MA model, with the finite length of the windowed correlation function as the MA model order [11]. The unique MA parameters of this invertible representation can be computed with an iterative algorithm [12]. In this way, periodograms can be compared with time series models.

All types of models, AR, MA and ARMA, must be estimated for many orders before a choice can be made for unknown data. Models of the unknown best order must be computed, but also models of higher orders. Only then it is possible to conclude that lower orders are better and should be selected. For every ARMA(p,q) model that gives a good fit, there are infinitely many ARMA(p+1,q+1) models with the same fit: all models where an additional pole and zero cancel. Therefore, the variance of the parameters of those overcomplete models will generally be much greater than the variance of the parameters in the correct model. The parameter variance may become so great that roots of estimated AR and/or MA polynomials fall outside the stationary or invertible region, with as boundary the unit circle. Order selection becomes difficult and uncertain if one candidate model has all roots inside the unit circle and an other candidate model has some roots outside. Mirroring of roots with respect to the unit circle or constraining them is no solution that solves this selection problem.

It is a fact that combinations of time series programs and order selection criteria are now available that compute *useful* models for unknown statistical data under all circumstances. Some combinations for AR, for MA and for ARMA modeling have recently been presented [4,9,10] with arguments why those combinations are useful. Algorithms for time series are useful for unknown data if they *cannot* produce zeros that are not invertible and poles that are not stationary. The best AR model of a true MA or ARMA process may still be a model of a poor quality; this AR model is nevertheless *useful* if it can be compared to MA and to ARMA models, which will be better then.

The Yule-Walker method of AR estimation can have a bias of order 1 for AR(p) models of processes where a true reflection coefficient has the absolute value |1-p/N| [13]. To reduce this possible bias contribution, it is better to use Burg's AR estimation method [14] for unknown data.

Durbin's method for MA estimation [7] and his second method for ARMA [8] use the parameters of a long intermediate autoregressive model to compute MA parameters. This avoids the non-linear optimization with problems of convergence and of non-invertible zeros. Durbin's algorithms fulfill all requirements to be useful under all circumstances: they have no problems with convergence and always produce an invertible solution. A recent improvement in Durbin's methods is that the theoretically optimal order has been defined for the intermediate AR model, so the performance of Durbin's algorithms in practice has been improved [9,10]. In many examples, the Cramér-Rao lower bound for the parameter accuracy is approximated in estimated MA and ARMA models for reasonable sample sizes.

Asymptotical AR order selection criteria can give wrong orders if the candidate orders become higher than 0.1N [6]. Using higher penalties or consistent criteria cannot cure this problem [5]. Taking the actual expectation of the logarithm of the residual variance into account helps according to the finite sample theory. The Combined Information Criterion CIC is based on the expectation and on the variance of the logarithm of the residual variance, as a function of the model order [4].

The prediction error PE is defined as the square of the one step ahead error of prediction or as the model fit to new and independent data. It is an obvious measure for the accuracy of time series models. Its asymptotical expectation equals $\sigma^2(1+p/N)$, where p is the number of parameters and σ^2 is the innovation variance. It also has an interpretation in the time domain as a measure for the spectral flatness. The model error ME is a scaled version of the excess prediction error due to the combination of model selection and parameter estimation:

$$\mathbf{ME} = \mathbf{N}\left(\frac{\mathbf{PE}}{\sigma_\varepsilon^2} - \mathbf{1}\right). \tag{1}$$

ME can easily be computed in the time domain [15] and its asymptotical expectation for unbiased models is the number of estimated parameters p, independent of the sample size. The practical significance of this measure ME can be seen in speech coding. An equivalent measure is often used there: the spectral distortion. SD^2 is defined as the average integral of $[\ln S - \ln \hat{S}]^2$; its asymptotical expectation equals 2p/N for small variations in AR processes [16], the same value as ME apart from a constant.

3. SELECTION OF MODEL TYPE

If the best model of each individual type is computed and selected with an algorithm that depends only on the data , the choice between the best AR, MA and ARMA models can be made with an objective statistical criterion. Examples of algorithms and order selection criteria that fulfill the requirements are given here.

AR models can preferably be computed with Burg's method [14], with a finite sample order selection criterion CIC, defined as [4]

$$\mathbf{CIC} = \ln\{\mathbf{RES}(\mathbf{p})\} + \max\left[\prod_{i=1}^{p} \frac{1+1/(N+1-i)}{1-1/(N+1-i)} - 1\ ,\ 3\sum_{i=1}^{p} \frac{1}{N+1-i}\right].$$

where RES(p) denotes the residual variance for order p. CIC is a compromise between the optimal asymptotical penalty factor 3 [5] and the finite sample estimator for the Kullback-Leibler information that gives a correction for the increasing variance of ln{RES(p)} as a function of the model order p.

Good and useful MA models can be computed with Durbin's method [7] that uses a long AR model as intermediate. The order of that AR model is chosen with a sliding window algorithm [9] as twice the AR order as selected with CIC plus the number of MA parameters that is to be estimated. The MA order q can be selected with the asymptotical selection criterion GIC(q,3) defined as:

$$\mathbf{GIC(q,3)} = \ln\{\mathbf{RES}(\mathbf{q})\} + 3\mathbf{q}/\mathbf{N}.$$

The penalty 3 is a compromise between the famous factor 2 in Akaike's criterion [2] with too much risk of overfit and consistent criteria where the selection of underfitting models with too low order is the main problem [5].

ARMA models can be computed with Durbin's second method [8]. A sliding window of length "twice the AR order + 2r-1" for the intermediate AR model has been described for ARMA(r,r-1) models [10] and the same penalty as for MA can be used for selection of the 2r-1 parameters of the ARMA(r,r-1) model order, so GIC(2r-1,3).

In Durbin's MA and ARMA estimation algorithms, the residual variance is computed after the estimation by substituting the parameters. In contrast, AR parameters are estimated by minimization of the residual variance. This different role of the residual variance is the reason to recommend different order selection criteria, CIC for AR and GIC for MA and ARMA selection. The typical finite sample behavior of AR residuals [6] prevents the selection of the model type with a single selection criterion in all circumstances. It would be possible to use the criterion GIC to choose between the selected MA(q) and ARMA(r,r-1) models. However, a new principle is necessary to include the AR(p) model. This is found by looking at the prediction error of the three different models. For a measured and given value of the residual variance, the conditional expectation of the prediction error for the selected AR(p) model is found with the finite sample theory as [6]:

$$\mathbf{PE}(\mathbf{p}) = \{\mathbf{RES}(\mathbf{p})\} \prod_{\mathbf{i}=1}^{\mathbf{p}} \frac{\mathbf{1}+\mathbf{1}/(\mathbf{N}+\mathbf{1}-\mathbf{i})}{\mathbf{1}-\mathbf{1}/(\mathbf{N}+\mathbf{1}-\mathbf{i})} \ . \qquad (2)$$

The constants 1/(N+1-i) in numerator and denominator represent the finite sample variance coefficients for Burg's estimation method. The conditional expectations of the prediction error for MA(q) and ARMA(r,r-1) models are based on the asymptotical theory and they are given by :

$$\mathbf{PE}(\mathbf{m}) = \{\mathbf{RES}(\mathbf{m})\} \frac{\mathbf{1}+\mathbf{m}/\mathbf{N}}{\mathbf{1}-\mathbf{m}/\mathbf{N}} \ , \qquad (3)$$

where m denotes the number of estimated parameters. From the three previously selected models, AR(p), MA(q) and ARMA(r,r-1), the type with the smallest estimate of the prediction error PE with (2) or (3) is chosen as the model type. This gives a single time series model where the model type and the model order are selected on purely statistical arguments. This model is denoted AR-MA.

4. TRIANGULAR BIAS IN PERIODOGRAMS

Periodograms can be derived as the square of the FFT of the data, but they can also be derived as Fourier transforms of N points of the **biased** covariance function [1]. This covariance bias is caused by an inevitable window on the estimated covariance in order to obtain a positive definite estimate in practice. The influence of the same covariance bias on Yule-Walker estimates of AR parameters has been evaluated before [13], showing a parameter bias of order of magnitude 1. The influence on periodograms is investigated here by applying this bias to the true covariance function for a MA(4) process with parameters [1 -1.68 1. 88 -1.39 .6], with agrees with reflection coefficients [-.6 .6 -.6 .6]. For a given innovation variance of 1, the covariance R(0)-R(4) in this example is [9.66 -8.29 5.35 -2.4 .6], followed by zeros for i>4. The spectra that are shown in Fig.1 have been computed with the true covariance R(i), multiplied by the triangle { 1-i/(N+1) } for increasing N. For N > 4000, the difference between the

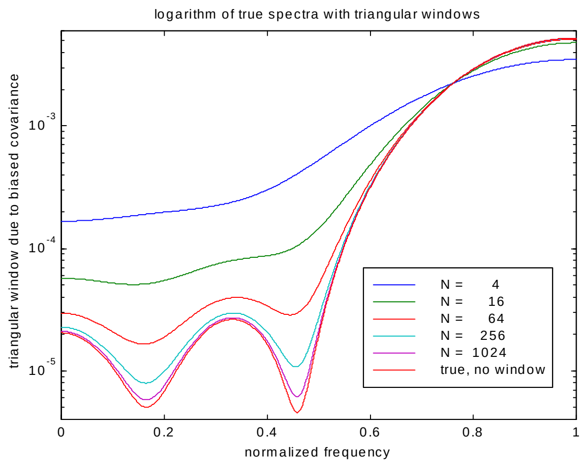

Fig.1 Spectrum of true covariance function, multiplied with the triangular window of length N due to the biased covariance used in periodograms.

Table 1: MODEL ERROR **ME** AS A FUNCTION OF THE SAMPLE SIZE FOR SPECTRA OF A MA(4) PROCESS OBTAINED WITH BIASED TRUE COVARIANCE

N	4	16	64	256	1024	4096
ME	6.1	6.6	6.4	4.3	1.8	.55

true unwindowed spectrum and the triangular window becomes small. Fig.1 shows that the influence of the bias on the spectral estimate can be serious, even without any estimation inaccuracies. The legend gives the lines in the correct sequence, so N=4 is the upper line and the true unbiased result is the bottom line. The accuracy of a spectrum estimated from N observations can be expressed in the model error ME, defined in (1) and particularly suited to describe the model quality for different sample sizes. The expectation of ME for estimating the 4 parameters of a MA(4) model equals 4. Table 1 demonstrates that the influence of the triangular bias in the theoretical covariance for the periodogram is, for N less than 256, greater than the total expected inaccuracy due to estimating the time series model. In other examples, with more difference between the parts with low and with high power in the spectrum, the influence of this bias could be much worse. As a consequence, this triangular bias prevents the accurate estimation of spectra with periodogram methods. The fact that still further windows are required in practice to reduce the statistical variance shows that inaccurate spectral estimates may be expected for windowed periodograms.

5. SIMULATIONS

Simulation experiments have been conducted with a double intention: first of all to investigate the quality of the single time series model, chosen with (2) and (3) and secondly to compare this quality with different windowed periodogram estimates. Those are expressed as invertible MA models [11], with parameters computed with an iterative algorithm [12]. In this way, the assessment of the quality is made with the same measure ME as used for time series.

Many simulations with numerous examples have been carried out. The average results of multiple simulation runs with the MA(4) example of Section 4 are presented in Table 2. This example is chosen to demonstrate the accuracy of periodograms on a process that can very well be represented by only a few values of the covariance function. The first four

Table 2: AVERAGE MODEL ERROR **ME** AS A FUNCTION OF THE SAMPLE SIZE FOR ESTIMATED SPECTRA OF A MA(4) PROCESS

N	32	64	128	256	512	1024
AR	12.1	17.5	22.5	23.6	27.9	32.7
MA	9.4	9.1	8.7	6.8	6.0	5.6
ARMA	11.5	16.1	19.9	17.3	13.9	12.5
true	8.0	8.3	8.0	5.7	4.9	4.6
AR-MA	10.8	12.1	10.9	8.1	6.6	5.9
P N/2	11.3	13.2	21.3	42.0	84.4	170.3
P N/4	15.7	13.8	13.8	21.0	40.5	82.3
P N/8	52.7	25.5	20.5	16.2	21.0	40.1
P N/16	186.6	95.4	47.8	34.5	21.7	22.3
P N/32	XX	369.5	187.9	92.3	63.1	33.6
P N/64	XX	XX	738.1	371.6	181.9	121.1
P N/128	XX	XX	XX	1474.6	740.2	361.7

rows give ME of the selected AR, MA and ARMA model and of the true MA(4) model. Row 5 gives the ME of the single selected time series model with smallest estimated PE with (2) and (3). The ME results of Parzen windows [1] with lengths between N/2 and N/128 follow. Of course, all windowed results use the biased estimate of the correlation function [1], with the triangular bias of section 4. Tapers on the data have been used throughout in the analysis of periodograms, because that was always an improvement, with lower values for the ME.

For N greater than 1024, the ME of the MA model in the row *true* approaches the theoretical minimum obtainable value 4: the true MA order. The average ME for the MA model with selected order in row 2 is only about 1 higher than the MA for the true order. Finally, the ME of the AR-MA model with *selected model type and model order* decreases with the sample size N. So the choice based on the PE of (2) and (3) introduces hardly an additional error in comparison with the situation in row 2 where the model type was considered to be known. In other words, it is not necessary to know the type of the time series in advance because the choice can be made with the method of this paper.

The length of the true correlation function is only 4 in the simulation example. Nevertheless, the best length of the Parzen windows in the periodogram estimates of the spectral density, so the window with the smallest ME in Table 2, was always much greater than 4, e.g. 64 for N=512 and 1024. This indicates that the deformation of the correlation function by the window shape has a strong influence. ME was still much greater if a Bartlett or triangular window was applied instead of the Parzen window. The average ME of all FFT based spectra in this example is worse than the ME of the single selected time series model. This result has always been found, in all examples. Typically, the quality of the best of all periodogram estimates is not as good as the quality of the spectrum of the single AR-MA model found with the methods of section 3. The difference in quality between windowed periodogram and AR-MA depends on the true process: The greater the difference between maxima and minima in the true spectrum, the greater the difference in ME of the periodogram and the time series model. Two applications of the AR-MA model have been reported. The variance of the mean depends on the sum of all covariances and is found with the AR-MA model [17]. Also irregularly sampled data can also be treated [18].

6. CONCLUDING REMARKS

A single time series model can be selected with objective statistical criteria from the three previously computed and selected AR(p), MA(q) and ARMA(r,r-1) models. The quality of that AR-MA model is excellent if the models of the three different types, (AR, MA and ARMA) have been estimated and selected with suitable algorithms and selection criteria. For unknown statistical data, the quality of the power spectral density of this automatically selected single model is better than even the best windowed periodogram spectral estimate.

FACT: use time series analysis for the spectral analysis of unknown statistical data to obtain the best accuracy.

REFERENCES

[1] M.B. Priestley, *'Spectral Analysis and Time Series'*, London, Ac. Press, 1981.

[2] H. Akaike, 'A New Look at the Statistical Model Identification', *IEEE Trans. Autom. Control*, vol AC-19, p 716-723, 1974.

[3] P.M.T. Broersen and H.E. Wensink, 'Autoregressive model order selection by a finite sample estimator for the Kullback-Leibler discrepancy', *IEEE Trans. on Signal Process.*, vol 46, in press, 1998.

[4] P.M.T. Broersen, 'The ABC of autoregressive order selection criteria', *Preprints Sysid '97 Conf., Kitakyushu, Japan*, 8-11 July 1997, p 231-236.

[5] P.M.T. Broersen and H.E. Wensink, 'On the penalty factor for autoregressive order selection in finite samples', *IEEE Trans. on Signal Process.*, vol 44, p 748-752, 1996.

[6] P.M.T. Broersen and H.E. Wensink, 'On Finite Sample Theory for Autoregressive Model Order Selection', *IEEE Trans. on Signal Processing*, vol SP-41, p 194-204, 1993.

[7] J. Durbin, 'Efficient Estimation of Parameters in Moving Average Models', *Biometrika*, vol 46, p 306-316, 1959.

[8] J. Durbin, 'The Fitting of Time Series Models', *Revue Inst. Int. de Stat.*, vol 28, p 233-243, 1960.

[9] P.M.T. Broersen, 'The best order of long autoregressive models for moving average estimation', *Signal Processing VIII, Proc. Eusipco Conf., Trieste, Italy*, p 799-802, 1996.

[10] P.M.T. Broersen, 'On orders of long AR models for ARMA estimation', *Proc. ECSAP '97 Conf., Prague, Czech Republic*, p 83-86, 24-27 June 1997.

[11] S.M. Kay and S.L. Marple, 'Spectrum Analysis-A Modern Perspective', *Proc. IEEE*, vol 69, p 1380-1419, 1981.

[12] G. Wilson, 'Factorization of the Covariance Generating Function of a Pure Moving Average Process', *SIAM Journal Numer. Anal.*, vol 6, p 1-7, 1969.

[13] J.S. Erkelens and P.M.T. Broersen, 'Bias Propagation in the Autocorrelation Method of Linear Prediction', *IEEE Trans. on Speech and Audio*, vol 5, p 116-119, 1997.

[14] J.P. Burg, 'Maximum Likelihood Spectral Analysis', *Proc. 37th Meeting Society of Exploration Geophysicists,* Oklahoma City, 6 p, 1967.

[15] P.M.T. Broersen, 'The quality of models for ARMA processes'. *IEEE Trans. on Signal Process.*, vol 46, in press, 1998.

[16] B. Friedlander and B. Porat, 'A General Lower Bound for Parametric Spectrum Estimation', *IEEE Trans. Acoust., Speech, Signal* Processing, vol ASSP-32, p 728-733, 1984.

[17] P.M.T. Broersen, 'Estimation of the Accuracy of Mean and Variance of Correlated Data'. *Proc. IEEE IMTC Conference*, St Paul, May 18-21,1998.

[18] S. de Waele and P.M.T. Broersen, 'Accurate LDA-spectra by Resampling and ARMA-modeling', *Proc. EUSIPCO-98 Conference,* 1998.