

# Facts and Fiction in Spectral Analysis

P. M. T. Broersen

**Abstract**—This analysis is limited to the spectral analysis of stationary stochastic processes with unknown spectral density. The main spectral estimation methods are: parametric with time series models, or nonparametric with a windowed periodogram. A single time series model will be chosen with a statistical criterion from three previously estimated and selected models: the best autoregressive (AR) model, the best moving average (MA) model, and the best combined ARMA model. The accuracy of the spectrum, computed from this single selected time series model, is compared with the accuracy of some windowed periodogram estimates. The time series model generally gives a spectrum that is better than the best possible windowed periodogram. It is a fact that a single good time series model can be selected automatically for statistical data with unknown spectral density. It is fiction that objective choices between windowed periodograms can be made.

**Index Terms**—ARMA models, identification, order selection, parametric spectrum, spectral accuracy, spectral estimation, time series.

## I. INTRODUCTION

THE main methods for spectral analysis are parametric with time series models and nonparametric with tapered and windowed periodograms [1]. Generally, time series models are to be preferred if the true model type and order of the process under investigation are known [2]. However, application of the wrong model type or the wrong order can produce completely erroneous spectral estimates [2]–[4]. The variance of  $AR(p)$  spectral estimates and of Fourier transforms of the first  $p$  values of the correlation function is comparable [5], [6]. However, an alternative interpretation shows that the equivalent of an infinitely long measured correlation is transformed with correctly chosen  $AR(p)$  models [7]. Order selection can be interpreted as looking for the lowest order where the extrapolated autoregressive (AR) correlation function is not significantly different from the measured correlation; this is verified up to the highest order that is considered for AR order selection. Order selection for AR models has been studied with asymptotic criteria [8]–[10] and with finite sample equivalents [11]–[14] that are necessary if higher model orders are considered. Also for moving average (MA) and ARMA models, reliable algorithms have been developed that perform well for all sample sizes [15]–[18]. This is a new development in time series analysis. In text books ([19], [20] p. 103) no preferences have been formulated for specific MA and ARMA algorithms. But after the discovery of the optimal length of the long autoregressive intermediate model [15], [16], preference can be given to Durbin's methods [17], [18].

This paper deals with stationary stochastic processes with unknown spectra, not with deterministic or periodical signals for

which Fourier transform will be the obvious method to compute the spectrum. Moreover, the results of spectral estimation will only be accurate if the length where the correlation has values that are significantly different from white noise is less than half the length of the observations. Also, the inverse length in reconstructing the innovations from the data should be less than half the observation length. Otherwise, the number of observations is not sufficient to obtain good information about the structure of the data with time series analysis. For all stationary stochastic processes, at least one of the three model types gives a good spectral description of the data. The problem is to know which type is appropriate. So far, no automatic method for the selection of the model type, AR, MA, or ARMA, yielded good results on statistical data with unknown spectra.

Requirements for AR, MA, and ARMA algorithms are described. They are necessary to allow the selection of model order and type. The algorithms have to produce useful models of any order, independent of the sample size and the true process order and type. Only in this way, the best  $AR(p)$  model, the best  $MA(q)$ , and the best  $ARMA(r, r-1)$  model can be selected for given data. Afterwards, an automatic choice among the three previously estimated and selected models can be made with a statistical concept. A new criterion with the squared error of prediction is given that can be used for the selection of the single final model. Afterwards, the data spectrum is computed from the model parameters. In simulations, an objective quality measure [21] is used to compare the spectrum of this single model to a variety of tapered and windowed periodograms. This measure represents the squared error of prediction in the time domain; at the same time it represents a relative accuracy measure integrated over the frequency domain [22].

## II. TIME SERIES MODELS

Time series have three different models types: 1) autoregressive or AR, 2) moving average or MA, and 3) combined ARMA. An  $ARMA(P, Q)$  process can be written as [1]

$$\begin{aligned} x_n + a_1x_{n-1} + \cdots + a_Px_{n-P} \\ = \varepsilon_n + b_1\varepsilon_{n-1} + \cdots + b_Q\varepsilon_{n-Q} \end{aligned} \quad (1)$$

where  $\varepsilon_n$  is a purely random process, so a sequence of independent identically distributed stochastic variables with zero mean and variance  $\sigma^2$ . This process is purely AR for  $Q = 0$  and MA for  $P = 0$ . Any stationary stochastic process with a continuous spectral density can be written as an unique  $AR(\infty)$  or  $MA(\infty)$  process [1]. This is independent of the origin of the process; it may be the sum of a true AR process and several colored noise signals or an  $ARMA(P, Q)$  process. Several parametrizations for AR processes are known [4]. Parameters, poles, reflection coefficients, and correlation coefficients are a few of the ways

Manuscript received May 26, 1998; revised March 10, 2000.

The author is with the Department of Applied Physics, Delft University of Technology, Delft, The Netherlands (e-mail: broersen@tn.tudelft.nl).

Publisher Item Identifier S 0018-9456(00)04808-7.

to characterize an AR process completely with a finite number of parameters. The roots of

$$A(z) = 1 + a_1 z^{-1} + \dots + a_P z^{-P} \quad (2)$$

are denoted the poles of the ARMA( $P, Q$ ) process; the roots of

$$B(z) = 1 + b_1 z^{-1} + \dots + b_Q z^{-Q}$$

are the zeros. Processes are called stationary if all poles are strictly within the unit circle and invertible if all zeros are within the unit circle.

In estimation,  $N$  observations of a process with unknown characteristics are given. If the mean of the observed signal is not zero, the average of the given data is subtracted before parameters are estimated. An estimated ARMA( $p, q$ ) model is given by

$$\begin{aligned} x_n + \hat{a}_1 x_{n-1} + \dots + \hat{a}_p x_{n-p} \\ = \hat{\varepsilon}_n + \hat{b}_1 \hat{\varepsilon}_{n-1} + \dots + \hat{b}_q \hat{\varepsilon}_{n-q} \end{aligned} \quad (3)$$

where  $p$  and  $q$  may be arbitrary and  $\hat{\varepsilon}_n$  is some colored noise signal that is not purely random in general. Models are AR if  $q = 0$  and purely MA for  $p = 0$ . Parameters for different model orders can be estimated in principle by minimization of the residual variance. That is the variance of  $\hat{\varepsilon}_n$  if the parameters are estimated from the data  $x_n$ . The theory shows that the minimum of the residual variance gives at the same time the maximal flatness for the spectral density of  $\hat{\varepsilon}_n$ . Afterwards, orders  $p$  and/or  $q$  are selected. A selection criterion adds a penalty for each estimated parameter to the logarithm of the residual variance of each estimated model. It selects the model order with the minimum of the criterion. Finite order models are often quite well in describing true processes with infinite orders, because the true parameters are decreasing rapidly for many processes.

A recursive algorithm gives a relation between AR parameters  $\hat{a}_i$  and reflection coefficients  $k_m$  [4]. The parameters  $\hat{a}_{i,m}$  of the order  $m$  are related to the parameters  $\hat{a}_{i,m-1}$  of order  $m-1$  by the Levinson-Durbin recursion

$$\begin{aligned} \hat{a}_{i,m} &= \hat{a}_{i,m-1} + k_m \hat{a}_{m-i,m-1}, \quad i = 1, \dots, m-1 \\ \hat{a}_{m,m} &= k_m. \end{aligned} \quad (4)$$

By definition, the last parameter of a stationary AR model of order  $m$  is called a reflection coefficient. Reflection coefficients of stationary processes are always between  $-1$  and  $1$ . Vice versa, AR models are stationary if all reflection coefficients are less than  $1$  in magnitude.

The residual variance  $\text{RES}(m)$  is the variance of  $\hat{\varepsilon}_n$  in (3) for AR( $m$ ) models. It is given by

$$\text{RES}(m) = \text{RES}(0) \prod_{i=1}^m (1 - k_i^2) \quad (5)$$

where  $\text{RES}(0)$  is the variance of the observations  $x_n$ .

The covariance function and the power spectral density of the data can be computed from the parameters of the estimated model. The estimated power spectrum  $\hat{h}(\omega)$  of the ARMA( $p, q$ ) model is given by [1]

$$\hat{h}(\omega) = \frac{\sigma_{\varepsilon}^2}{2\pi} \left| 1 + \sum_{i=1}^q \hat{b}_i \exp(-j\omega i) \right|^2 / \left| 1 + \sum_{i=1}^p \hat{a}_i \exp(-j\omega i) \right|^2. \quad (6)$$

The covariance can be found as the inverse Fourier transform of (6). Also direct methods to derive the covariance from the parameters are given in the time series literature [1], [4].

### III. FICTION

It is fiction that looking at figures of periodograms obtained with different types or lengths of windows and tapers can yield a statistically reliable choice between the alternatives. Many examples show the contrary. It seems that the human eye or the human mind prefers to select a limited amount of details in the spectra but not too much; the choice is hardly influenced by the amount of details in the true spectrum. The theory about the choice of window type and size can only be developed on deterministic grounds and it requires the exact knowledge of the true spectrum [1].

Theoretically, AR or MA roots and their image, mirrored with respect to the unit circle, produce the same covariance structure and the same spectral densities ([1] p. 146). In practice, model type and order are not known and selection is necessary. This is based on the decrease of the residual variance caused by introducing extra parameters in the model. Mirroring has influence on the residual variance. Hence, to favor order selection there is a preference for algorithms that give exclusively invertible and stationary roots of estimated models. If mirrored images of the roots are used, whenever noninvertible or nonstationary roots are estimated, order selection becomes dubious. A comparison between models of increasing orders with all roots inside the unit circle and a model with one or more roots outside the unit circle involves not only the number of parameters. Therefore, a statistical selection of the model order becomes impossible. Moreover, mirrored images of estimated roots that fall outside the invertible or stationary region will not concentrate around the location of the true roots. Many algorithms cannot guarantee that all estimated poles and zeros are certainly invertible or stationary. For AR, MA, and ARMA algorithms, the most important members of this unfavorable class are the least squares algorithms. For those algorithms, estimated roots can be outside the stationary or invertible region. This is sometimes prevented by using constrained least squares optimization, but in those algorithms the choice of the constraint determines the solution and not the data. Only if an algorithm is guaranteed to estimate stationary or invertible roots, the roots remain completely determined by the data. If algorithms produce undesirable roots, they can simply be outside the unit circle due to the estimation variance; their mirrored image is no useful estimate.

Algorithms that may produce wrong roots in small samples improve their behavior for an increasing number of observations

from the same process. But it is always possible to create examples where estimated roots can be wrong due to estimation variance. Examples can be given where the asymptotical approximation of the variance of reflection coefficients can become greater than 1. In those examples, the estimation variance is the reason that an unconstrained least squares algorithm can estimate roots on both sides of the unit circle, or even exactly on it. Take an AR(2) or MA(2) process with an arbitrary true reflection coefficient  $k_1$  and  $k_2 = -1 + 1/N$ . The variance of the first reflection coefficient  $k_1$  is  $2(1 - k_1^2)$  [23], independent of  $N$ . This means that estimated 'reflection coefficients' can become greater than 1 due to the variance and that is equivalent with non-stationary or noninvertible roots.

Algorithms that can produce poles outside the unit circle will sometimes do so. However, all types of models, AR, MA, and ARMA, must be estimated for different model orders, before a choice can be made for data with unknown spectral density. Not only models of the, afterwards selected, best order must be computed, but *also models of higher orders*. Only then it is possible to conclude that lower orders are the best 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: those are all models where an additional pole and zero cancel. Therefore, the variance of the least squares parameters of those overcomplete models is  $\infty$  in the asymptotic theory. Hence, estimated poles may be nonstationary and estimated zeros noninvertible and the residual variance cannot be used as a basis for order selection.

Especially for the overcomplete models, that are necessary for a good selection of the model order, convergence problems and roots outside the unit circle can be expected for least squares ARMA algorithms. It is fiction that least squares ARMA algorithms are capable of routinely estimating higher order ARMA( $p + 1, q + 1$ ) models that can be compared in a statistical reliable way with ARMA( $p, q$ ) models with order selection criteria. At least order selection is seriously hampered by mirroring roots or by constraining them; so many existing ARMA algorithms are not suitable for automatic use. That might be a reason why AR, MA, or ARMA has not been used before as an alternative to windowed periodograms for spectral analysis.

#### IV. FACTS

The Fourier transform of a stationary stochastic process does not exist [1, p.15] as an ordinary function. More specifically, the Fourier transform is not approximated any better by taking more observations. The quotient of two Fourier transforms is generally considered to be an estimate of the transfer function. However, for stationary stochastic processes the estimate based on the Fourier transform of the output signal divided by that of the input has an infinite variance [24]. This paper describes the consequences of this problem in estimating the periodogram, the square of the absolute value of a Fourier transform.

Any stationary stochastic process with a continuous spectral density can be written as an unique AR( $\infty$ ) or MA( $\infty$ ) process [1]. This does not depend on the origin of the process, e.g., it may be the sum of two or more ARMA processes or an AR

process with additive noise. Finite order MA or AR models give often an accurate description for those infinite order processes, because the true parameters decrease rapidly for most processes.

It is a fact that *combinations* of time series algorithms and order selection criteria are available that compute under all circumstances *useful* models for statistical data with unknown spectral density. Examples of combinations for AR, MA, and for ARMA modeling have been presented recently [11], [15], [16]. Algorithms for time series are useful for data with unknown spectral density if they *cannot* produce zeros that are not invertible and poles that are nonstationary. 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 objectively to MA and to ARMA models. Those will fit better in that case.

Four different AR model orders can be considered as theoretically optimal for the representation of a known MA( $q$ ) process, depending on the intended use or interpretation. The orders are

- $\infty$  for the *exact* theoretical equivalence between MA( $q$ ) and AR( $\infty$ );
- 0 (or  $\infty$ ) for  $q + 1$  or more steps ahead prediction if the exact MA( $q$ ) process is known;
- $K$  defined in [15] for one step ahead prediction with an *estimated* AR( $K$ ) model or for computing the spectral density with an AR approximation;
- the higher order  $M$  as intermediate order of an *estimated* AR( $M$ ) model for calculation of the MA parameters with Durbin's method [15].

The Yule-Walker method of AR estimation can have a severe bias, that is not present in the Burg method [25]. The bias is caused by the use of the biased estimate for the covariance of the data in the Yule-Walker method for AR estimation. Estimated reflection coefficients for orders higher than  $p$  are a factor 2 too small for models of processes where a true reflection coefficient of order  $p$  has the absolute value  $|1 - p/N|$  [26]. Precisely the same covariance bias is present in periodogram estimates for the spectrum; the influence of this bias on periodograms will be shown in Section VI.

Asymptotical AR order selection criteria can give wrong orders if the candidate orders are higher than  $0.1 N$  [13]. Using higher penalties in selection criteria or consistent criteria cannot cure this problem [12]. Taking the actual *expectation* of the logarithm of the residual variance into account helps according to the finite sample theory. A good criterion also depends on the *sampling variance* of the logarithm of the residual variance, as a function of the model order [11].

Any windowed periodogram can also be written as the spectrum of a finite order invertible MA model, with the finite length of the windowed correlation function as the MA model order [4]. The unique MA parameters of this invertible representation can be computed with an iterative algorithm [27]. Therefore, a periodogram can be represented as a MA model. This gives the opportunity to perform objective comparisons between periodograms and time series spectra with an accuracy measure that has been developed for time series [21].

The squared error of prediction PE is defined as the model fit to new and independent data or also as the square of the one step ahead error of prediction. It is also defined as the variance

of  $\hat{\varepsilon}_n$  in (3), under the condition that the parameters in (3) are not estimated from the data  $x_n$  in (3) but from independent data. In other words, the variance of  $\hat{\varepsilon}_n$  in (3) represents the residual variance if the parameters are estimated from the same data and the squared error of prediction if parameters and data are independent.

If the number of estimated parameters  $p$  and  $q$  are equal to or greater than the true orders of the generating process, the asymptotic expectation of the residual variance equals  $\sigma^2\{1 - (p + q)/N\}$ , where  $\sigma^2$  is the innovation variance of  $\varepsilon_n$  in (1). The asymptotic expectation of PE equals  $\sigma^2\{1 + (p + q)/N\}$ . PE has also an interpretation in the frequency domain as a measure for the spectral flatness: the model with minimum PE maximizes the spectral flatness of autoregressive models. A third interpretation is in terms of differences between the true and the estimated probability density function of the data, the Kullback-Leibler discrepancy. Some order selection criteria have been derived as approximations for this difference [14]. The best model is the model with the smallest estimate of the discrepancy.

An expression for PE in the frequency domain is the integral of  $S/\hat{S}$  [28]; an approximation is given by the integral of  $(S - \hat{S})^2/2S^2$  [9];  $S$  is the true spectrum and  $\hat{S}$  the estimate with an ARMA model. The relative spectral measures have an equivalent that is often used in speech coding: the spectral distortion  $SD^2$ . That is defined as the integral of  $[\ln S - \ln \hat{S}]^2$  and it equals  $2p/N$  for small variations in AR processes [5]. The model error ME is a scaled version of the squared error of prediction [11], [21]. ME can easily be computed in the time domain and its asymptotical expectation for complete unbiased models is  $p + q$ , independent of the sample size.

## V. SELECTION OF MODEL ORDER AND TYPE

The choice between the three previously selected best AR, MA, and ARMA models can only be made with an objective statistical criterion if the best model of each individual type is computed and selected with an algorithm that depends exclusively on the data, not on subjective choices of the experimenter. Order selection is for each model type only possible if the differences in the residual variance between successive model orders are purely determined by statistics. This means that all models must be obtained without constraints on the roots, without mirroring of roots, without elimination of undesirable model choices and without influence of the maximum candidate model order, provided that this order is taken high enough. Examples of algorithms and order selection criteria that fulfill the requirements are given here. They can be replaced by other methods, as long as only statistics is involved in parameter estimation and order selection.

AR models can be computed with Burg's method [25], with a finite sample order selection criterion  $CIC(p)$ , defined as [11]

$$CIC(p) = \ln\{\text{RES}(p)\} + \max \left[ \left( \prod_{i=1}^p \frac{1 + 1/(N + 1 - i)}{1 - 1/(N + 1 - i)} - 1 \right) \left( 3 \sum_{i=1}^p \frac{1}{N + 1 - i} \right) \right]. \quad (7)$$

$\text{RES}(p)$  is the variance of the residuals of the AR model of order  $p$ ;  $1/(N + 1 - i)$  is an empirical finite sample formula to approximate the variance of the reflection coefficient of order  $i$  estimated from  $N$  observations with the Burg algorithm [13].  $CIC(p)$  is a compromise between the optimal asymptotical penalty factor 3 [12] and the finite sample estimator for the Kullback-Leibler information [14] that gives a correction for the increasing variance of  $\ln\{\text{RES}(p)\}$  as a function of the model order  $p$ . The order  $K^*$  is selected if  $CIC(p)$  has its minimum at  $p = K^*$ .

MA( $q$ ) models can be computed with Durbin's method [17] that uses a long AR model as intermediate. The order of that AR model, which should be estimated with Burg's method, is essential for a good result [15]. It is chosen with a sliding window algorithm [15] as twice the selected AR order plus the number of MA parameters that is to be estimated, so  $2K^* + q$ . The MA order  $q$  is selected as the minimum of the asymptotical selection criterion  $GIC(q, 3)$  defined as

$$GIC(q, 3) = \ln\{\text{RES}(q)\} + 3q/N. \quad (8)$$

ARMA( $r, r - 1$ ) models can be computed with Durbin's second method [18]. This method has a poor reputation in the literature, but its performance is much improved by taking a proper order for the intermediate AR model. A sliding window choice  $3K^* + 2r - 1$  has been described for the AR order to estimate ARMA( $r, r - 1$ ) models [16]. 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  $GIC(2r - 1, 3)$ . The restriction of orders to  $r, r - 1$  has the computational advantage that not all combinations of  $p$  and  $q$  have to be evaluated.

The residual variance is minimized during AR estimation, but it is computed afterwards in MA and ARMA estimation. Therefore, different order selection criteria,  $CIC(p)$  and  $GIC(q, 3)$ , have been recommended.  $GIC(q, 3)$  can also be used to choose between the selected MA( $q$ ) and ARMA( $r, r - 1$ ) models, but a new principle is necessary to include the AR( $p$ ) model. The finite sample behavior of AR residuals [13] prevents the selection of the model type with a single selection criterion. The principle is found by realizing that in practice a model is desired with a small prediction error. For a measured and given value of the residual variance, the conditional expectation of the squared error of prediction for the selected AR( $p$ ) model is found with the finite sample theory as [13]

$$PE(p) = \{\text{RES}(p)\} \prod_{i=1}^p \frac{1 + 1/(N + 1 - i)}{1 - 1/(N + 1 - i)}. \quad (9)$$

The conditional expectations of the prediction error for MA( $q$ ) and ARMA( $r, r - 1$ ) models are based on the same asymptotical theory as the Final Prediction Error. This is Akaike's first order selection criterion [29]

$$PE(m) = \{\text{RES}(m)\} \frac{1 + m/N}{1 - m/N} \quad (10)$$

where  $m$  denotes the number of estimated parameters. From the three selected models, the type with the smallest estimate of the prediction error PE with (9) or (10) is chosen. This gives a *single*

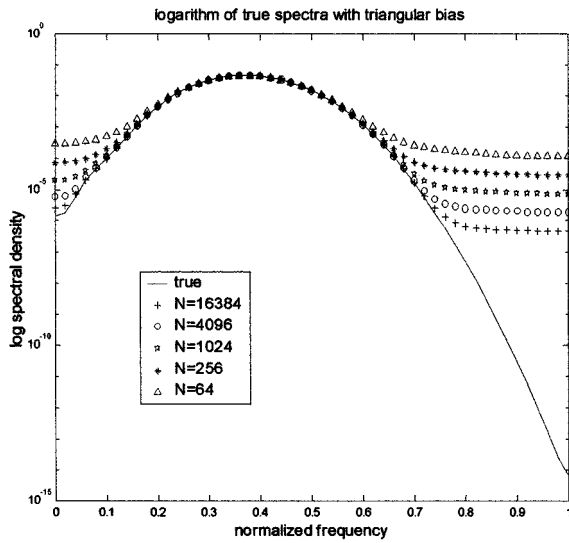


Fig. 1. Spectra computed with the true covariance function multiplied with a triangular bias of length  $N$ .

*time series model*, called ARMAse1, that is selected on purely statistical arguments.

## VI. SIMULATIONS

Periodograms can be seen as Fourier transforms of  $N$  points of the biased estimated autocovariance function. The bias is given by a multiplication with the triangular window  $1 - \tau/N$  [1]. Removing this bias is undesirable, because the positive semi-definite property would be lost. This property is guaranteeing positive estimates for the spectral density [1]. The influence of this bias is investigated by applying the bias to the true covariance function  $R(\tau)$ , without any estimation errors. The same bias is present in the Yule-Walker algorithm for AR estimation.

An example is presented here, which has been used before [6] to study the accuracy of time series models. Its autocovariance function is a Gaussian bell shape as envelope of a cosine function

$$R(\tau) = e^{-0.5(0.25\tau)^2} \cos(0.38\pi\tau) \quad (11)$$

for  $\tau = 0, 1, \dots, 20$ . For ease of computation and for a convenient generation of simulation data, the correlation is made equal to zero for values of the shift  $\tau$  greater than 20, neglecting only covariance smaller than  $10^{-6}$ . In this way, the generating process for the simulations is MA(20).

The spectra in Fig. 1 are computed for different values of  $N$  with the true covariance  $R(\tau)$  multiplied by the triangular window  $1 - \tau/(N+1)$ ,  $\tau = 0, 1, \dots, 20$ . Those spectra are the expectation of the periodogram and also of the AR spectrum that would be obtained with the Yule-Walker algorithm. This figure shows that the influence of the inevitable bias on the spectral estimate can be serious, even for more than 10 000 observations. The accuracy of this spectra of  $N$  points can be expressed in the model error ME. This is a scaled version of the squared error of prediction [21]

$$ME = N \left( \frac{PE}{\sigma_\varepsilon^2} - 1 \right). \quad (12)$$

TABLE I  
MODEL ACCURACY AS A FUNCTION OF  $N$  FOR A TRIANGULAR BIAS IN THE COVARIANCE, FOR A PROCESS WITH A COSINE WITH A GAUSSIAN BELL ENVELOPE AS CORRELATION

N	64	256	1024	4096	16384
ME	2071	4819	11934	31158	84905
PE/ $\sigma^2$	33.35	19.82	12.65	8.60	6.18

PE was estimated as a transformation of the residual variance in (9) and (10). In simulations, the PE can be computed using estimated model parameters and true process parameters. In the application here, the true MA(20) process defined for (11) is compared with the parameters of the MA(20) model that generates  $R(\tau)\{1 - \tau/(N+1)\}$ . The usual application of ME is to models estimated from  $N$  observations in simulations. For models of the true process type and with orders equal to the true order or higher, the asymptotical expectation of ME equals the number of estimated parameters, independent of the number of observations. ME is particularly suited to describe the model quality for different sample sizes.

The expectation of ME for estimating the 20 parameters of a MA(20) model equals 20. Table I gives much higher values for ME, without estimation but only with applying the inevitable triangular bias that appears in windowed periodograms. In practice still a smoothing window would be required to improve the variance of the estimated periodogram. Table I demonstrates that the influence of the triangular bias in the theoretical covariance for the periodogram is much greater than the total expected inaccuracy due to estimating the time series model. Also the values for PE are given. The value of the variance of the process in (11) is  $443\sigma^2$ , where  $\sigma^2$  is the variance of the innovations of the MA(20) process. In other words, the PE is 0.08 times the variance of the process for  $N = 64$  in Table I and 0.014 times that variance for  $N = 16384$ .

Simulation experiments have been conducted with a double intention: first of all to investigate the quality of the single time series model, chosen with the minimum of (9) and (10) and secondly to compare this quality with different windowed periodograms. Those are expressed as invertible MA models [4], with parameters computed with an iterative algorithm [27]. In this way, the assessment of the quality is made with the same measure as used for time series. Many simulations with numerous examples have been carried out. Low order AR, MA, and ARMA processes give the result that the true model structure and model order give the best spectrum and that windowed periodogram transforms perform less. The true spectrum and some estimates of the process defined in (11) are presented in Fig. 2, for one simulation run with  $N = 128$ . The ME value of the selected AR(36) time series model is 150 in this spectrum. To compute periodogram estimates, the data have been tapered with a cosine bell over the first and the last 10% of the sample interval, leaving the middle 80% undisturbed [1]. Tapering was used throughout for periodogram estimation because it improves the quality. A Parzen window over  $N/4$  points of the estimated correlation function had ME = 1260. Both num-

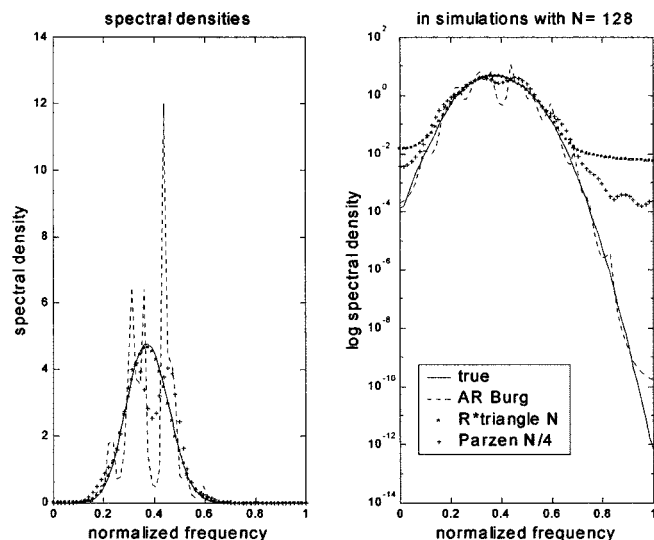


Fig. 2. Spectral density and its logarithm for the true spectrum, the selected AR model, the spectrum with triangular bias and a periodogram with Parzen window, for  $N = 128$ .

bers of ME are smaller than in Table I. This means that estimated Burg models are better than AR models obtained with the Yule–Walker bias. Moreover, using tapers in periodogram estimation largely reduces the influence of the triangular bias, at the cost of a slightly increased variance. The estimated logarithm of the tapered and windowed periodogram has the same appearance as in Fig. 1. Fig. 2 clearly demonstrates the problem of making a choice visually, by looking at measured spectra: it is very likely that people prefer the periodogram estimate in the left-hand representation and the AR estimate in the right-hand picture, although both describe the same data. Therefore, an objective criterion is necessary to make a proper choice for data with unknown spectral density. Section IV showed the relationship between PE and the logarithmic spectral distortion  $SD^2$ . This explains why an accurate spectral estimate with a small PE is the same as an estimate with a small error in the logarithmic spectrum.

The average results of multiple simulation runs are presented in Table II. The first four rows give the selected AR, MA, and ARMA model and the true MA(20) model. Row 5 gives the ME of the single selected ARMAse1 time series model with smallest estimated PE with (2) and (3). The ME results of periodograms with Parzen windows [1] with lengths  $M$  between  $N/2$  and  $N/128$  follow; they are indicated as  $P : M$  in the Table, with  $M$  for the length of the Parzen window. Of course, all windowed results use the biased estimate of the correlation function [1], with expectation  $R(i)\{1 - i/N\}$ . This bias effect is greatly reduced by tapering. No selected periodogram is given, because no objective global selection criterion for windowed periodograms exists. Locally, a variable bandwidth selector principle has been used for smoothing the raw periodogram with polynomial models [30]. No specific model for stationary stochastic processes is used in this local smoothing. Therefore, it is not possible to evaluate local results in terms of statistical model accuracy.

TABLE II  
AVERAGE MODEL ERROR ME AS A FUNCTION OF THE SAMPLE SIZE FOR DIFFERENT ESTIMATED SPECTRA OF A PROCESS WITH A COSINE WITH A GAUSSIAN BELL ENVELOPE AS CORRELATION

	N	32	64	128	256	512	1024
AR	241	114	88	80	79	82	
MA	185	716	2732	215	24	21	
ARMA	192	375	272	85	28	26	
true	131	69	60	45	33	27	
ARMAse1	209	105	80	41	25	22	
P: N/2	1208	875	762	870	1070	XX	
P: N/4	2706	1685	1317	1350	1570	1880	
P: N/8	8689	4886	2962	2470	2570	2960	
P: N/16	13189	17405	9258	5720	4850	5020	
P: N/32	XX	26449	34568	18190	11340	9560	
P: N/64	XX	XX	52856	68890	36270	22540	
P: N/128	XX	XX	XX	105650	137890	72340	

For  $N$  is 1024, the ME of the MA model approaches the theoretical minimum obtainable value 20: the true MA order. It is remarkable that estimated covariances in the periodograms of Table II produce a lower ME than the biased true covariances in Table I; this is caused by the application of a taper.

The ME of the selected ARMAse1 model in Table II is often smaller than the average of the three separate types: AR, MA, or ARMA. Therefore, the choice of the type based on the PE of (9) and (10) hardly introduces an additional error. 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 ME in Table II for MA estimates for  $N = 64, 128$  and 256 has some high values because the computed residual variance was quite distorted in some simulation runs, which caused the selection of the wrong MA order. However, in those runs always the best AR or ARMA was selected as the preferred time series model and that turned out to be a good alternative. This shows that a useful method must at least give a good indication of the accuracy of each estimated model.

The length of the true correlation function is only 20. Nevertheless, the best window length with the smallest ME in Table II was always the greatest length, so 256 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 ME of all periodogram based spectra in this example is worse than the ME of the single selected time series model. This has always been found. Tapers have been used throughout in the periodogram analysis, because that was an improvement, with lower values for ME than without a taper. An application shows that the ARMAse1 model also gives the best estimate for the autocovariance function [31], especially for the summation of all covariances.

The best time series model type depends on the number of observations, as is shown in Table III. In this MA example, the best ARMAse1 choice was the AR type for small samples. Only for  $N$  greater than 250, a preference for MA models is found. This shows that it is good or even necessary to have the three

TABLE III  
THE PERCENTAGE OF SELECTION OF AR, MA, AND ARMA MODEL AS A  
FUNCTION OF  $N$

N	32	64	128	256	512	1024
AR	77	83	88	16	0	0
MA	13	15	8	57	83	89
ARMA	10	2	4	27	17	11

types of time series models available for data with unknown spectral density.

## VII. CONCLUDING REMARKS

A single ARMAse1 time series model can be selected from the three previously computed and selected  $AR(p)$ ,  $MA(q)$ , and  $ARMA(r, r-1)$  models. The quality of that ARMAse1 model is excellent if the models of the three different types have been estimated and selected with suitable algorithms. For statistical data with unknown spectral density, the quality of this single model cannot be reached by even the best windowed periodogram, in all examples. Moreover, the time series model is selected with an objective statistical criterion. Such criteria are not available for windowed periodograms. Even without estimation, the bias in the true covariance function of periodogram estimates can cause inadmissible distortion in the spectral density.

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

## REFERENCES

- [1] M. B. Priestley, *Spectral Analysis and Time Series*. London, U.K.: Academic, 1981.
- [2] N. Beamish and M. B. Priestley, "A study of autoregressive and window spectral estimation," *Appl. Stat.*, vol. 30, pp. 41–58, 1981.
- [3] P. R. Gutowski, E. A. Robinson, and S. Treitel, "Spectral estimation: Fact or fiction," *IEEE Trans. Geosci. Electron.*, vol. GE-16, pp. 80–84, 1978.
- [4] S. M. Kay and S. L. Marple, "Spectrum analysis—A modern perspective," *Proc. IEEE*, vol. 69, pp. 1380–1419, 1981.
- [5] B. Friedlander and B. Porat, "A general lower bound for parametric spectrum estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 728–733, 1984.
- [6] M. Kaveh and G. R. Cooper, "Empirical investigation of the properties of the autoregressive spectral estimator," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 313–323, 1976.
- [7] A. van den Bos, "Alternative interpretation of maximum entropy spectral analysis," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 493–494, 1971.
- [8] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 716–723, 1974.
- [9] E. Parzen, "Some recent advantages in time series modeling," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 723–730, 1974.

- [10] R. Shibata, "Selection of the order of an autoregressive model by Akaike's information criterion," *Biometrika*, vol. 63, pp. 117–126, 1976.
- [11] P. M. T. Broersen, "The ABC of autoregressive order selection criteria," in *Preprints Sysid '97 Conf.*, Kitakyushu, Japan, July 8–11, 1997, pp. 231–236.
- [12] P. M. T. Broersen and H. E. Wensink, "On the penalty factor for autoregressive order selection in finite samples," *IEEE Trans. Signal Processing*, vol. 44, pp. 748–752, 1996.
- [13] —, "On finite sample theory for autoregressive model order selection," *IEEE Trans. Signal Processing*, vol. SP-41, pp. 194–204, 1993.
- [14] —, "Autoregressive model order selection by a finite sample estimator for the Kullback-Leibler discrepancy," *IEEE Trans. Signal Processing*, vol. SP-46, pp. 2058–2061, 1998.
- [15] P. M. T. Broersen, "The best order of long autoregressive models for moving average estimation," in *Signal Processing VIII, Proc. Eusipco Conf.*, Trieste, Italy, 1996, pp. 799–802.
- [16] —, "On orders of long AR models for ARMA estimation," in *Proc. ECSAP'97 Conf.*, Prague, Czech Republic, June 24–27, 1997, pp. 83–86.
- [17] J. Durbin, "Efficient estimation of parameters in moving average models," *Biometrika*, vol. 46, pp. 306–316, 1959.
- [18] —, "The fitting of time series models," *Revue Inst. Int. de Stat.*, vol. 28, pp. 233–243, 1960.
- [19] S. M. Kay, *Modern Spectral Estimation, Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
- [20] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice Hall, 1997.
- [21] P. M. T. Broersen, "The quality of models for ARMA processes," *IEEE Trans. Signal Processing*, vol. 46, pp. 1749–1751, 1998.
- [22] —, "The performance of spectral quality measures," in *Proc. IMTC Conf.*, Venice, Italy, May 24–26, 1999, pp. 751–756.
- [23] S. Kay and J. Makhoul, "On the statistics of the estimated reflection coefficients of an autoregressive process," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. SP-31, pp. 1447–1455, 1983.
- [24] P. M. T. Broersen, "A comparison of transfer function estimators," *IEEE Trans. Instrum. Meas.*, vol. 44, pp. 657–661, 1995.
- [25] J. P. Burg, "Maximum likelihood spectral analysis," in *Proc. 37th Meeting Soc. of Exploration Geophysicists*, vol. 6, Oklahoma City, OK, 1967, pp. 1–6.
- [26] J. S. Erkelens and P. M. T. Broersen, "Bias propagation in the autocorrelation method of linear prediction," *IEEE Trans. Speech Audio Processing*, vol. 5, pp. 116–119, 1997.
- [27] G. Wilson, "Factorization of the covariance generating function of a pure moving average process," *SIAM J. Numer. Anal.*, vol. 6, pp. 1–7, 1969.
- [28] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, 1976.
- [29] H. Akaike, "Statistical predictor identification," *Ann. Inst. Statist. Math.*, vol. 22, pp. 203–217, 1970.
- [30] J. Fan and I. Gijbels, *Local Polynomial Modeling and Its Applications*. London, U.K.: Chapman & Hall, 1996.
- [31] P. M. T. Broersen, "Estimation of the accuracy of mean and variance of correlated data," *IEEE Trans. Instrum. Meas.*, vol. 47, pp. 1085–1091, 1998.

**Piet M.T. Broersen** was born in Zijdewind The Netherlands, in 1944. He received the M.Sc. degree in applied physics in 1968 and the Ph.D. degree in 1976, both from the Delft University of Technology, Delft, The Netherlands.

He is currently in the Department of Applied Physics, Delft University. His research interests are the selection of order and type of time series models and the application to spectral analysis, model building, and feature extraction.