

The Burg Algorithm for Segments

Stijn de Waele and Piet M. T. Broersen

Abstract—In many applications, the duration of an uninterrupted measurement of a time series is limited. However, it is often possible to obtain several separate segments of data. The estimation of an autoregressive model from this type of data is discussed. A straightforward approach is to take the average of models estimated from each segment separately. In this way, the variance of the estimated parameters is reduced. However, averaging does not reduce the bias in the estimate. With the Burg algorithm for segments, both the variance and the bias in the estimated parameters are reduced by fitting a single model to all segments simultaneously. As a result, the model estimated with the Burg algorithm for segments is more accurate than models obtained with averaging. The new weighted Burg algorithm for segments allows combining segments of different amplitudes.

Index Terms—Segmented data, time-series analysis.

I. INTRODUCTION

Autoregressive (AR) modeling is used for analyzing stationary stochastic processes for many different applications, e.g., radar [1], geophysics [2], and economics [3]. A comparison of various estimators of AR-parameters showed that the Burg algorithm [4], [5] is the preferred estimator for AR-parameters [6]. The Yule–Walker algorithm can be severely biased. After a large true reflection coefficient, estimates for higher order reflection coefficients suffer from a bias of order 1 instead of the smaller bias of order $1/N$ that is present in other estimation methods. The least squares estimator and the forward-backward least squares estimator have a greater variance than Burg. In addition, they may yield unstable models. In many applications, the duration of one uninterrupted measurement is submitted to practical limitations. However, it is often possible to obtain several separate segments of data. The Burg algorithm has been reformulated so that it can be used to analyze this type of data [7].

An example of the application of this generalized Burg algorithm is the estimation of a sea clutter model from radar data. The advantage of estimating a clutter model from the data at hand, rather than using a predefined model, is that the model is automatically adapted to the current environmental conditions. The clutter model can be used for target detection: Statistically significant deviations from clutter indicate a possible target [8]. The generalized Burg algorithm can be used to obtain an accurate clutter model by combining the information in adjacent range cells into a single clutter model. A more accurate clutter model results in more effective target detection.

Manuscript received July 21, 1999; revised June 5, 2000. The associate editor coordinating the review of this paper and approving it for publication was Dr. Lal C. Godara.

The authors are with the Department of Applied Physics, Delft University of Technology, Delft, The Netherlands.

Publisher Item Identifier S 1053-587X(00)07643-1.

The usual way of combining the information of segments is to take the average of the models estimated from the individual segments [9]. This will reduce the variance present in the estimate. However, AR estimation methods also contain a bias. The bias in the AR parameters is proportional to $1/N$, where N is the number of observations [10]. Taking the average of the models estimated from the individual segments does not reduce this bias.

The influence of the bias can be neglected when the number of observations per segment tends to infinity, whereas the number of segments is fixed. In this asymptotic regime, all methods for obtaining a model from segments are equivalent. Therefore, the asymptotic regime cannot be used to discern between the various methods for combining the information in the segments. To do this, the situation with many segments, each with a finite number of observations, has to be considered. In this regime, the bias in the estimated parameters becomes important.

II. THE BURG ALGORITHM FOR SEGMENTS

An autoregressive process is a (possibly complex-valued) stationary stochastic signal $x(n)$ generated by the following difference equation [11]:

$$x(n) - \mu + \sum_{i=1}^p a_i(x(n-i) - \mu) = \varepsilon(n) \quad (1)$$

where $\varepsilon(n)$ is a white noise signal. The mean of the signal is denoted μ . In this process, S segments are given. In each segment, the process is considered to be stationary. The Burg algorithm for segments can be derived for segments of varying length. For the sake of simplicity of formulae, the segments will be presumed to be of equal length N in the derivation given here. Observation n of segment s will be denoted $x^s(n)$, $n = 1, \dots, N$. All properties of the Burg algorithm for segments that are discussed are also valid for segments of varying length.

The Burg algorithm estimates the AR parameters by determining reflection coefficients k that minimize the sum of forward and backward residuals. The extension of the algorithm to segments is that the reflection coefficients are estimated by minimizing the sum of forward and backward residuals of all segments taken together. This means a single model is fitted to all segments in one time. This concept is also used for prediction error methods in system identification, where the input to the system is known, like in ARX modeling [12].

The advantage of this approach can be illustrated by looking at two ways of dealing with the signal mean μ . Before estimating the AR parameters of a signal with nonzero mean, the estimated mean must be subtracted. Due to the fact that an estimate of the

mean is subtracted instead of the true mean, this subtraction introduces bias in the estimated AR parameters [11]. Treatment of the mean for unsegmented data with maximum likelihood can be found in [13]. The mean can be estimated and subtracted for each segment individually. Alternatively, a single estimate of the mean can be obtained from all segments taken together. As this estimate of the mean is more accurate, the bias in the model parameters caused by subtraction of the estimated mean will be reduced. When the segments are treated separately, the bias will not decrease when the number of segments is increased. Obviously, the single estimate for the mean should only be used when all segments are obtained from signals with the same mean.

Some steps of the derivation of the Burg algorithm will be discussed to show the implications of fitting a model to all segments simultaneously. In the Burg algorithm, reflection coefficients k are estimated directly with a recursive algorithm [14]. In each recursion step, a single reflection coefficient is estimated. To estimate the p th reflection coefficient an AR(p) model is fitted to the data with the first $p - 1$ reflection coefficients k_1, \dots, k_{p-1} fixed to the value found in the previous steps. The parameters $\hat{a}_i^{[p]}$ of step p are related to the parameters of the previous step by the Levinson–Durbin algorithm [14]. The new reflection coefficient k_p is determined by minimizing the sum of the forward and backward residuals. The forward residuals $f^s[p]$ and the backward residuals $b^s[p]$ for a p th-order model and segment s are defined as

$$\begin{aligned} f^s[p](n) &= x^s(n) + \sum_{i=1}^p \hat{a}_i^{[p]} x^s(n-i) \\ n &= p+1, \dots, N \\ b^s[p](n) &= x^s(n) + \sum_{i=1}^p \hat{a}_i^{[p]*} x^s(n+i) \\ n &= 1, \dots, N-p. \end{aligned} \quad (2)$$

An essential characteristic of the Burg algorithm is that the number of residuals decreases with each recursion step. The criterion to be minimized with respect to k_p is the sum of absolute squares P of the forward residuals and the backward residuals

$$P(k_p) = \frac{1}{2} \sum_{s=1}^S \left(\sum_{n=p+1}^N |f^s[p](n)|^2 + \sum_{n=1}^{N-p} |b^s[p](n)|^2 \right). \quad (3)$$

The forward and backward residuals can be expressed in terms of the forward and backward residuals of the previous step:

$$\begin{aligned} f^s[p](n) &= f^s[p-1](n) + k_p b^s[p-1](n-p) \\ b^s[p](n) &= b^s[p-1](n) + k_p^* f^s[p-1](n+p). \end{aligned} \quad (4)$$

For the remainder of the derivation of the Burg estimate, it is convenient to express these equations in vector notation. The following vectors will be used:

$$\begin{aligned} \bar{f}^s[p] &= (f^s[p](p+1) \dots f^s[p](N)) \\ \bar{b}^s[p] &= (b^s[p](1) \dots b^s[p](N-p)). \end{aligned} \quad (5)$$

In addition, the residuals of the previous iteration step are used in (4). As stated before, the number of residuals decreases

with each iteration step. Therefore, the vectors of residuals from the previous step $\bar{f}^s[p-1]$ and $\bar{b}^s[p-1]$ do not have the same length as $\bar{f}^s[p]$ and $\bar{b}^s[p]$. The vectors of forward and backward residuals of the previous step with the appropriate length are denoted $\bar{v}^s[p]$ and $\bar{w}^s[p]$:

$$\begin{aligned} \bar{v}^s[p] &= (f^s[p-1](p+1) \dots f^s[p-1](N)) \\ \bar{w}^s[p] &= (b^s[p-1](1) \dots b^s[p-1](N-p)). \end{aligned} \quad (6)$$

With these vectors, (4) can be rewritten as

$$\begin{aligned} \bar{f}^s[p] &= \bar{v}^s[p] + k_p \bar{w}^s[p] \\ \bar{b}^s[p] &= \bar{w}^s[p] + k_p^* \bar{v}^s[p] \end{aligned} \quad (7)$$

Using the standard inner product $\langle \bar{x}, \bar{y} \rangle = (\sum x_i y_i)/N$ and norm $\|\cdot\|$ the criterion P can be written as

$$P(k_p) = \sum_s \frac{1}{2} (\|\bar{f}^s[p]\|^2 + \|\bar{b}^s[p]\|^2). \quad (8)$$

Substitution of (7) and some calculation yields that $P(k_p)$ is minimal for

$$k_p = \hat{k}_p = - \frac{\sum_s \langle \bar{v}^s[p], \bar{w}^s[p] \rangle}{\sum_s \frac{1}{2} (\|\bar{v}^s[p]\|^2 + \|\bar{w}^s[p]\|^2)}. \quad (9)$$

As is the case for the standard Burg algorithm for a single segment, the absolute value of \hat{k}_p is always smaller than unity. Therefore, the stability of the estimated AR model is guaranteed. As the proof of the stability property is analogous to that given in [14], it will not be given here.

The calculations required for the Burg algorithm for segments are very similar to the calculations for Burg applied to each segment separately. The difference is that the information contained in the segments is combined in a single estimate, as given by (9).

In the standard criterion (3), all segments are weighted equally. This can be modified by introducing weighting factors $(w^s)^2$ for each segment in (3). Like the unweighted algorithm, this weighted Burg algorithm for segments yields an estimated model that is guaranteed to be stable. This can immediately be seen from the fact that the weighted optimization problem can be translated into the original optimization problem with scaled observations $w^s x^s(n)$ instead of $x^s(n)$. Since stability of the original solution (9) is guaranteed for all sets of observations, this proves that the estimated model found with the modified criterion will also be stable.

An example of the application of the weighted algorithm is the situation where the various segments are realizations of a stochastic process with the same autocorrelation function but with a different variance. This can be taken into account by using the inverse of the standard deviation σ^s of the segment as weight w^s in the weighted algorithm.

The estimate found from several short segments will be less accurate than the estimate from one uninterrupted measurement. This is caused by the fact that for the p first observations of a segment the residuals are not calculated since not all previous observations required for these residuals are known. Therefore,

only NS-pS residuals are available. If one uninterrupted measurement of length NS is available, only the first p residuals are not calculated, leaving NS-p residuals. Moreover, the maximum model order that can be considered a candidate for order selection is higher for the uninterrupted measurement.

The number of parameters p that is used to describe the data can be obtained by order selection using an order selection criterion such as Akaike's information criterion (AIC) [15]. When only one segment of length N is available, the AIC is given by

$$\text{AIC}(p) = \ln(\text{RES}(p)) + \frac{2p}{N} \quad (10)$$

where $\text{RES}(p)$ is equal to the residual variance of the model of order p . When a single model is estimated from S independent segments of length N , the variance of the estimated parameters asymptotically decreases by a factor of S . This means the AIC for segments (AICS) is given by

$$\text{AIC}_s(p) = \ln(\text{RES}(p)) + \frac{2p}{SN}. \quad (11)$$

When the model order p , which is considered is greater than $0.1N$, finite sample effects must be taken into account [6]. If the segments are dependent, SN in (11) should be replaced by the effective number of degrees of freedom.

III. AVERAGING METHODS

With averaging methods, the information of the segments is combined by averaging the results of the individual segments. In this way, the variance of the estimate is reduced. However, averaging does not reduce the bias present in the estimates. The exact result of this approach depends on the model representation that is used. For a general representation, which is denoted α , the averaging method can be written as

$$\hat{\alpha} = \frac{1}{S} \sum_s \hat{\alpha}^s. \quad (12)$$

Some representations used for averaging are the following:

- p first autocorrelations ρ_1, \dots, ρ_p ;
- p first autocovariances R_1, \dots, R_p ;
- reflection coefficients k_1, \dots, k_p ;
- parameters in the difference equation a_1, \dots, a_p .

Another representation that can be used for averaging is the power spectrum $h(\omega)$ or, equivalently, the entire covariance function R . The average of the entire covariance function of S AR(p) processes cannot be written as the covariance function of an AR(p) process because the average of the entire covariance function of S AR(p) processes is the covariance function of an ARMA($Sp, (S-1)p$) process. However, an AR(p) process can be fitted to this average covariance function by determining the AR(p) model with the smallest error with respect to the average covariance function. This model is completely determined by the first p autocovariances. As a result, this procedure is equivalent to averaging the first p autocovariances.

Three distinct estimators for the autocovariance R could be used for averaging: the unbiased sample covariances, the biased sample covariances, and covariances calculated from the Burg reflection coefficients. The unbiased sample covariances may

yield an estimate that is not a valid covariance function because it is not positive definite [11]. The biased sample autocovariances contain the same large bias as the Yule-Walker estimator. The Burg covariance function is guaranteed to be positive definite and contains only a small bias. Therefore, the Burg covariance function will be used. Similarly, the correlation function ρ will be calculated from the Burg reflection coefficients.

In the asymptotic regime, all representations (ρ, R, k and a) used for averaging are equivalent. However, for finite samples, the various approaches have some distinguishing properties. The average of parameters a is not guaranteed to yield a stationary model as the average model. The remaining representations mentioned do result in a stationary model under all conditions. A comparison of reflection coefficients k and parameters a as a representation for averaging showed that averaging reflection coefficients is more accurate [9]. For speech processing, a comparison of a wider range of averaging methods showed that averaging autocovariances R provides the most accurate model [16].

A common disadvantage of averaging methods is that the bias in the estimated parameters is not reduced as a result of averaging. The bias in the parameters is dependent on the number of observations. Exact results for the bias are difficult to obtain. Approximate results using Taylor expansion for the least-squares AR estimator show that the bias b for the reflection coefficient of an AR(1) process is inversely proportional to the number of observations per segment N [10]

$$b = E(\hat{k}) - k \sim \frac{1}{N}. \quad (13)$$

This bias is called Taylor bias. The bias will increase for higher model orders and for a greater dynamic range in the spectrum. This is proved in [10] for processes of order p equal to 2. For the Burg estimator, no theoretical results have been obtained. However, simulations showed that the bias in the Burg estimator is similar to the bias in the least-squares estimate [10].

The variance of the estimated parameter for an AR(1) process is proportional to $1/N$. This is reduced by a factor of S as a result of averaging S independent contributions

$$\text{var}(\hat{a}) \sim \frac{1}{SN}. \quad (14)$$

Since the contribution of the bias to an error measure such as the prediction error PE is given by the square of the bias

$$b^2 \sim \frac{1}{N^2} \quad (15)$$

the bias will be negligible with respect to the variance if the number of segment S is small compared with the number of observations N . However, as the number of segments increases, the variance is reduced, whereas the contribution of the bias will remain the same. Then, the influence of the bias will become important. The lack of an exact theoretical description is not a great problem for the current investigation. It is sufficient to acknowledge the existence of the Taylor bias and the fact that averaging does not diminish this bias.

Since the Burg algorithm for segments fits a model to all segments simultaneously, it is expected that the Taylor bias b_S in

this estimate will be small. An expression for the bias in \hat{k}_1 can be obtained by using a second-order Taylor approximation of the estimate (9) for k [17]. For simplicity of formulae, the analysis is again restricted to segments of equal length. Each segment's contribution to the expectation of the denominator is assumed to be equal to the signal variance σ^2 . This is a sufficiently accurate approximation for the current calculation. The resulting expression for the bias b_s is

$$b_s \approx \frac{1}{\sigma^4 S^2} \text{cov} \left(\sum_s \langle \bar{v}^{s[1]}, \bar{w}^{s[1]} \rangle, \sum_s \frac{1}{2} (\|\bar{v}^{s[1]}\|^2 + \|\bar{w}^{s[1]}\|^2) \right) + k_1 \frac{1}{\sigma^4 S^2} \text{var} \left(\sum_s \frac{1}{2} (\|\bar{v}^{s[1]}\|^2 + \|\bar{w}^{s[1]}\|^2) \right). \quad (16)$$

For independent segments, the (co)variances in this expression reduce to S times the (co)variance in a single segment. As a result, the bias b_s can be expressed in terms of the bias in a single segment

$$b_s \approx \frac{1}{S} \left(\frac{1}{\sigma^4} \text{cov} \left(\langle \bar{v}^{[1]}, \bar{w}^{[1]} \rangle, \frac{1}{2} (\|\bar{v}^{[1]}\|^2 + \|\bar{w}^{[1]}\|^2) \right) + k_1 \frac{1}{\sigma^4} \text{cov} \left(\frac{1}{2} (\|\bar{v}^{[1]}\|^2 + \|\bar{w}^{[1]}\|^2) \right) \right) = \frac{1}{S} b. \quad (17)$$

Substitution of (13) for b yields that the bias is inversely proportional to the total number of observations NS

$$b_s \sim \frac{1}{NS}. \quad (18)$$

IV. COMPARISON OF AVERAGING METHODS AND BURG

Burg for segments is compared with the averaging methods in two simulation experiments. In the first experiment, the two approaches are compared in an example with a moderate number of segments. In the second experiment, a very large number of segments is analyzed to study the bias present in both methods. Finally, some comments are made regarding the number of computations required for the various methods.

The stationary stochastic process used in these simulations is generated by an AR(10) process with ten equal reflection coefficients h . The value of h is varied to generate a range of signals with different properties. From this signal, an AR(10) model is estimated with Burg for segments and the averaging methods. The mean of the signal is considered known. The difference between the true signal properties and the estimated signal properties is expressed using the model error ME [18]. The model error is a normalized version of the one-step-ahead prediction error PE

$$\text{ME}(y, x) = NS \frac{\text{PE} - \sigma_\varepsilon^2}{\sigma_\varepsilon^2} \quad (19)$$

where PE is the expectation of the squared difference between $x(n)$ and the one-step-ahead prediction $\hat{x}(n)$ based on the previous observations and the estimated model

$$\begin{aligned} \text{PE} &= E\{(x(n) - \hat{x}(n))^2\} \\ &= E\{(x(n) + \hat{a}_1 x(n-1) + \dots + \hat{a}_p x(n-p))^2\}. \end{aligned} \quad (20)$$

TABLE I
COMPARISON OF THE BURG ALGORITHM FOR SEGMENTS (BURG-S) AND THE AVERAGING METHODS FOR $S = 10$ SEGMENTS OF $N = 20$ OBSERVATIONS EACH. THE SIGNAL IS GENERATED BY AN AR(10)-PROCESS CHARACTERIZED BY THE PARAMETER h . THE MODEL ERROR ME WAS AVERAGED OVER 50 000 SIMULATION RUNS FOR EACH VALUE OF h

h	Model Error ME				
	Burg-S	Averaging Methods			
		ρ	R	k	a
-0.5	16	47	41	138	370
-0.4	15	26	23	79	149
-0.3	14	18	16	46	149
-0.2	14	15	14	30	52
-0.1	15	13	14	23	34
0.0	15	13	14	19	24
0.1	15	14	14	18	22
0.2	15	17	16	20	25
0.3	16	24	22	25	34
0.4	17	45	38	37	50
0.5	18	102	84	64	84

The expectation $E\{\cdot\}$ is the average over all possible realizations of the process $x(n)$ and not the fit to the realization used for parameter estimation.

The model error contains contributions of both the variance and the squared bias in a way that is significant for the use of the models. The behavior of the Burg algorithm can be described using the finite sample formulae for the prediction error PE [6]. If one uninterrupted measurement is available, the model error would be approximately equal to the number of parameters 10.

In the first simulation experiment, the various methods have been compared for $S = 10$ segments of $N = 20$ observations. This example is representative for radar processing practice. The results are given in Table I. The value for the model error calculated using the finite sample formulae is 14.7. The results for the new Burg algorithm remain more or less the same for all values of h , whereas the averaging methods deteriorate for large values of h .

As stated before, there are no theoretical results for the bias in the Burg estimator. However, the bias can be examined in simulations by taking a very large number of segments. Then, the variance of the estimates is greatly reduced by averaging, whereas the bias remains present. In this way, a clear picture of the bias in the averaging methods is obtained. This has been done in the second simulation experiment, where the various methods have been compared for $S = 10\,000$ segments of $N = 20$ observations. The results are given in Table II. The value for the model error that is calculated using the finite sample formulae is 14.4.

The Burg algorithm for segments effectively combines the information in the segments, thus reducing the Taylor bias. The model error is only a result of the variance of the estimated parameters as predicted by the finite sample formulae. The averaging methods are biased. For all of the averaging methods, the bias increases for higher values of h . This agrees with the behavior of the bias in simple model structures, as described in [10]. The exact amount of bias depends on the model representation used for averaging. A comparison of the results of the two simulation experiments clearly shows that the deterioration of the averaging methods in the first experiment can be ascribed to the bias.

TABLE II

COMPARISON OF THE BURG ALGORITHM FOR SEGMENTS (BURG-S) AND THE AVERAGING METHODS FOR $S = 10\,000$ SEGMENTS OF $N = 20$ OBSERVATIONS EACH. THE SIGNAL IS GENERATED BY AN AR(10)-PROCESS CHARACTERIZED BY THE PARAMETER h . THE MODEL ERROR ME WAS AVERAGED OVER 100 SIMULATION RUNS FOR EACH VALUES OF h

h	Burg-S	Model Error ME			
		p	R	k	a
-0.5	15	25044	21592	105860	69951
-0.4	13	8216	6179	57099	50490
-0.3	14	2705	1535	27523	35843
-0.2	14	1076	356	13107	22869
-0.1	13	309	57	7115	11505
0.0	14	22	23	3961	4256
0.1	14	606	248	2696	2075
0.2	15	2775	1345	3615	3991
0.3	15	9139	5236	7251	9720
0.4	15	30324	19438	15497	21128
0.5	17	104306	71926	35287	44210

It can be concluded that the Burg algorithm for segments provides more accurate models than any of the averaging methods. How does it compare with the averaging methods in terms of numerical workload? As remarked in Section II, the calculations for Burg for segments are very similar to the calculations for Burg applied to each segment separately. This means that the computational effort required for averaging methods and Burg for segments are practically the same. However, averaging covariances, correlations, and parameters requires a transformation of the estimated reflection coefficients to the representation used for averaging. Therefore, the number of calculations required for these methods will be somewhat greater. A drawback of the Burg algorithm for segments is that exchanging a single segment for a new segment requires new processing of all segments. The estimate of the first reflection coefficients will be different as a result of the new segment. This reflection coefficient is used to find the residuals with (4). In this way, the new segment influences the contribution of all other segments for subsequent reflection coefficients.

V. CONCLUSIONS

The Burg algorithm for segments effectively combines the information contained in segments of data. The new weighted Burg algorithm allows combining segments of different amplitudes. The estimated model is stable under all conditions. The Burg algorithm for segments has been compared with averaging methods. The numerical workload is comparable with that of the averaging methods. With averaging methods, the variance in the estimated parameters is effectively reduced. However, the Taylor bias is not diminished. The influence of this bias becomes more prominent when the number of segments is large or when the process under consideration deviates considerably

from white noise. With the Burg algorithm for segments, both the variance and the bias in the estimated parameters are reduced. As a result, the Burg algorithm for segments provides more accurate models than the averaging methods.

REFERENCES

- [1] S. Haykin, *Nonlinear Methods of Spectral Analysis*, S. Haykin, Ed. Berlin, Germany: Springer-Verlag, 1979.
- [2] P. Yuou, E. Baert, and M. F. Loutre, "Spectral analysis of climate data," *Survey Geophys.*, vol. 17, pp. 619–663, 1996.
- [3] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. San Francisco, CA: Holden-Day, 1980.
- [4] J. P. Burg, "Maximum entropy spectral analysis," in *Proc. 37th Meet. Soc. Explorational Geophys.*, Oklahoma City, OK, 1967.
- [5] S. M. Kay and S. L. Marple, "Spectrum analysis—A modern perspective," *Proc. IEEE*, vol. 69, pp. 1380–1419, 1981.
- [6] P. M. T. Broersen, "The ABC of autoregressive order selection criteria," in *Proc. Sysid Conf.*, Kitakyushu, Japan, July 8–11, 1997, pp. 231–236.
- [7] S. Haykin, B. W. Currie, and S. B. Kesler, "Maximum-entropy spectral analysis of radar clutter," *Proc. IEEE*, vol. 70, 1982.
- [8] M. I. Skolnik, *Introduction to Radar Systems*. Auckland, New Zealand: McGraw-Hill, 1981.
- [9] A. A. Beex and M. D. A. Rahman, "On averaging Burg spectral estimators for segments," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1473–1484, Dec. 1986.
- [10] D. Tjøstheim and J. Paulsen, "Bias of some commonly-used time series estimates," *Biometrika*, vol. 70, pp. 389–399, 1983.
- [11] M. B. Priestley, *Spectral Analysis and Time Series*. London, U.K.: Academic, 1981.
- [12] L. Ljung, *System Identification for the User*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.
- [13] T. W. Anderson, *The Statistical Analysis of Time Series*. New York: Wiley, 1971.
- [14] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1997.
- [15] H. Akaike, "A new look at statistical model identification," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 716–723, 1974.
- [16] J. S. Erkelens and P. M. T. Broersen, "LPC interpolation by approximation of the sample autocorrelation function," *IEEE Trans. Speech Audio Processing*, vol. 6, pp. 569–573, Nov. 1998.
- [17] M. Kendall, A. Stuard, and J. K. Ord, *The Advanced Theory of Statistics, Vol. 3, Design and Analysis, Time Series*, 4th ed. London, U.K.: Charles Griffin, 1987.
- [18] P. M. T. Broersen, "The quality of models for ARMA processes," *IEEE Trans. Signal Processing*, vol. 46, pp. 1749–1752, 1998.

Stijn de Waele was born in Eindhoven, The Netherlands, in 1973. He received the M.Sc. degree in applied physics in 1998 from the Delft University of Technology, Delft, the Netherlands, where he is currently pursuing the Ph.D. degree.

His research is aimed at the development of new time series analysis algorithms and its application to radar processing.

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 with 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.