

FREQUENCY-DOMAIN IDENTIFICATION OF CONTINUOUS-TIME ARMA MODELS FROM SAMPLED DATA

Jonas Gillberg* Lennart Ljung*

* *Department of Electrical Engineering, Linköping
University
SE-581 83 Linköping, Sweden
Email: gillberg@isy.liu.se, ljung@isy.liu.se*

Abstract: This paper treats direct identification of continuous-time autoregressive moving average (CARMA) noise models. The approach has its point of origin in the frequency domain Whittle likelihood estimator. The discrete- or continuous-time spectral densities are estimated from equidistant samples of the output. For low sampling rates the discrete-time spectral density is modeled directly by its continuous-time spectral density using the Poisson summation formula. In the case of rapid sampling the continuous-time spectral density is estimated directly by modifying its discrete-time counterpart. *Copyright ©2005 IFAC*

Keywords: Continuous-time systems; Parameter estimation; Continuous-time ARMA; noise model; Whittle likelihood estimator

1. INTRODUCTION

Approaching parameter estimation from the discrete time domain has been the dominating paradigm in system identification. Identification of continuous-time models on the other hand is motivated by the fact that modeling of physical systems often take place in the continuous-time domain. In many practical applications there is also a genuine interest in the physical parameters. In the black-box discrete-time modeling framework however, the identified parameters often lack a physical interpretation.

Parameter identification of continuous-time systems is an important subject of its own. See for example the survey article by Unbehauen and Rao, (Unbehauen and Rao, 1990), the monograph by Sinha and Rao (Sinha and Rao, 1991) or the PhD thesis by Mensler (Mensler, 1999).

The research in identification of continuous-time models has primarily been concentrated to the

time-domain, with approaches such as: Poisson moment functionals, integrated sampling, orthogonal functions etc. A few authors on the other hand, have tackled the problem in the frequency domain. An early reference is by Shinbrot (Shinbrot, 1957) followed by the Fourier modulating function approach introduced by Pearson et al. (Pearson and Shen, 1993). Frequency-domain analysis for periodic and arbitrary signals has also been the starting point for the work by Pintelon et al. (Pintelon and Schoukens, 1997).

Recently there has been a renewed interest in continuous-time system identification in general and continuous-time noise models in particular, (Rao and Garnier, 2002), (Ljung, 2003a), (Larsson, 2003). See for instance the articles by Larsson and Söderström on continuous-time AR (Larsson and Söderström, 2002) and ARMA (Larsson and Mossberg, 2003) parameter estimation. The work on hybrid Box-Jenkins and ARMAX modeling by Pintelon et al (Pintelon and Schoukens, 2000) and

Johansson (Johansson, 1994) also concerns this problem. These approaches have in common that they use approximations of the noise or the noise model and consequently suffer from a bias in the model parameters.

The goal in this paper is to derive a method for the identification of continuous-time autoregressive moving average (CARMA) noise models. As a tool for this a frequency-domain Whittle likelihood approach together with some rudimentary theory for continuous- and discrete-time stochastic processes is used.

2. THE MODEL AND THE DATA

We shall consider continuous-time ARMA models represented as

$$y_t = G_c(p)e_t \quad (1)$$

where e_t is continuous time white noise such that

$$\begin{aligned} E[e_t] &= 0 \\ E[e_t e_s] &= \sigma^2 \delta(t-s) \end{aligned}$$

The operator p is here the differentiation operator. We assume that $G(p)$ is strictly proper, so y_t itself does not have a white-noise component, but is a well defined second order, stationary process. Its spectrum (spectral density) can be written as

$$\Phi_c(\omega) = \sigma^2 |G_c(i\omega)|^2 \quad (2)$$

We shall consider a general model parameterization

$$G_c(p, \theta) \quad (3)$$

where the model parameter vector θ includes the noise variance λ (whose true value is σ^2). The transfer function G can be parameterized by θ in an arbitrary way, for example by the conventional numerator and denominator parameters:

$$\begin{aligned} G(p, \theta) &= \frac{B(p)}{A(p)} \\ A(p) &= p^n + a_1 p^{n-1} + a_2 p^{n-2} + \dots + a_n \quad (4) \\ B(p) &= p^m + b_1 p^{m-1} + \dots + b_m \\ \theta &= [a_1 \ a_2 \ \dots \ a_n \ b_1 \ b_2 \ \dots \ b_m \ \lambda]^T \end{aligned}$$

In practice we have only sampled data $y(t_k)$ at our disposal to estimate θ . Although the case with irregularly sampled data is a prime target for our studies, we shall in this paper assume equidistantly sampled data:

$$y(t_k), t_k = kT_s, k = 1, \dots, N \quad (5)$$

3. THE ESTIMATION PROBLEM

The common way of modeling a general time series of the unstructured form (4) is to estimate a discrete-time model in the time domain and then

transform it to continuous time. If the parameterization (3) is tailor-made it would have to be transformed to discrete time

$$G_d(q, \theta) \quad (6)$$

by the well known sampling formulas, (Söderström, 1991), retaining the original parameters. Then the discrete time grey box model $y_t = G_d(q, \theta)e_t^d$ can be estimated by a straightforward prediction error method. See, e.g (Ljung, 1999), and implementations of this approach in (Ljung, 2003b). These methods work well and for Gaussian distributed signals they constitute the Maximum likelihood approach. Thus, in theory and asymptotically as the number of data tends to infinity, no method can perform better. However, for fast sampled data, they could be subject to numerically ill-conditioned calculations. Also, for irregularly sampled data, the computational burden could be substantial.

In this paper we shall consider frequency domain approaches as an alternative.

The spectrum of the sampled output according to (6) will be

$$\Phi_d(e^{i\omega T_s}, \theta) = \lambda_d |G_d(e^{i\omega T_s}, \theta)|^2 \quad (7)$$

where λ_d is the model's variance of the sampled driving noise e_t^d .

There is another way of writing the spectrum of the sampled output, which is obtained from Poisson's summation formula, see, e.g. (Papoulis, 1977)

$$\Phi_d(e^{i\omega T_s}, \theta) = \sum_{k=-\infty}^{\infty} \Phi_c(i\omega + i\frac{2\pi}{T_s}k, \theta) \quad (8)$$

4. THE WHITTLE LIKELIHOOD ESTIMATOR

4.1 Sampled data

From the sampled output (5) the periodogram of the signal can be computed:

$$\hat{\Phi}_d(e^{i\omega T_s}) = \left| \hat{Y}_d(e^{i\omega T_s}) \right|^2 \quad (9)$$

where

$$\hat{Y}_d(e^{i\omega_k T_s}) = \sqrt{\frac{T_s}{N}} \sum_{k=1}^N y(kT_s) e^{-i\omega_k T_s} \quad (10)$$

and

$$\omega_k = \frac{2\pi}{T_s} k. \quad (11)$$

If y has a Gaussian distribution, so will \hat{Y}_d have. It has zero mean and variance

$$E|\hat{Y}_d(e^{i\omega_k T_s})|^2 = \Phi_d(e^{i\omega_k T_s}, \theta)$$

under the model assumption (7). The pdf of $\hat{Y}_d(e^{i\omega_k T_s})$ is thus

$$\frac{1}{\sqrt{2\pi\Phi_d(e^{i\omega_k T_s}, \theta)}} e^{-\frac{|\hat{Y}_d(e^{i\omega_k T_s})|^2}{2\Phi_d(e^{i\omega_k T_s}, \theta)}}$$

At the DFT grid (11) the variables are also asymptotically independent. This means that a frequency domain procedure for estimating the parameter θ is

$$\hat{\theta}_N \triangleq \arg \min_{\theta} V_N^d(\theta) \quad (12a)$$

where

$$V_N^D(\theta) \triangleq \sum_{k=1}^{N_\omega} \frac{\hat{\Phi}_d(e^{i\omega_k T_s})}{\Phi_d(e^{i\omega_k T_s}, \theta)} + \log \Phi_d(e^{i\omega_k T_s}, \theta) \quad (12b)$$

This method is known as the *Whittle likelihood es-*

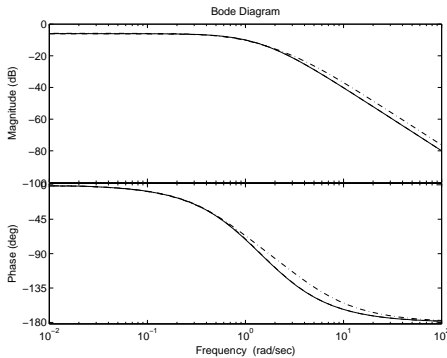


Fig. 1. Bode diagram comparing the Whittle likelihood estimator without (dashdot) and with (dashed) folding of the continuous-time spectrum to the true system (solid). The dashed line is almost identical to the solid.

timator (Whittle, 1961) and is an approximation of the corresponding time-domain ML-method. It has long been used to estimate parameters of discrete-time ARMA models.

4.2 Continuous time signals

Let us for a moment assume that we have available the whole, continuous time signal y over the time interval $[0 T]$ ($T = NT_s$). Then the periodogram estimate of the continuous time spectrum could be computed as

$$\hat{\Phi}_c^T(i\omega) = |Y_c^T(i\omega)|^2.$$

where the truncated continuous-time Fourier transform is

$$Y_c^T(i\omega) = \frac{1}{\sqrt{T}} \int_0^T y(t) e^{-i\omega t} dt.$$

Just as in the sampled case Y_c^T will have a Gaussian distribution with zero mean and variance equal to the continuous time spectrum (once T is

large enough so that transients and non-periodic effects can be neglected). Moreover the Fourier transforms will be asymptotically independent for frequencies that are further apart than the frequency resolution $2\pi/T$. See e.g. (Brillinger, 1981) and (Gillberg, 2004), Section 3. Just as for (12) we thus have the continuous-time Whittle-type estimator

$$\hat{\theta} \triangleq \arg \min_{\theta} V_N^T(\theta, \hat{\Phi}_c^T) \quad (13a)$$

where

$$V_N^T(\theta, \hat{\Phi}_c^T) \triangleq \sum_{k=1}^{N_\omega} \frac{\hat{\Phi}_c^T(i\omega_k)}{\Phi_c(i\omega_k, \theta)} + \log \Phi_c(i\omega_k, \theta). \quad (13b)$$

See also (Gillberg, 2004), Chapter 3 for a more detailed description.

In practice it gives better robustness to truncate the summation over frequencies in (12b) and (13b) at a lower frequency \bar{N}_ω , since the Fourier transforms of the signals typically are less reliable at higher frequencies.

5. INDIRECT FREQUENCY-DOMAIN ESTIMATION

By an indirect estimation method we mean using the discrete time spectrum to estimate the continuous time parameters. In the frequency domain it means that the criterion (12) is used. The question is how to calculate the discrete time spectrum Φ_d . A possibility that apparently has not been so much discussed is to use the Poisson summation type formula (8). A practical benefit is that only a limited number of terms in this sum may be needed to produce a good approximation of its discrete-time counterpart in the frequency range of interest.

5.1 Numerical Illustration

In Figure 1 we have estimated the second-order continuous-time AR-model

$$y_t = \frac{\sigma}{p^2 + a_1 p + a_2} e_t \quad (14)$$

with $\sigma = 1$, $a_1 = 3$ and $a_2 = 2$. The duration of the data set was $T = 1000s$ with the sampling time $T_s = 1s$. The figure illustrates the frequency-domain bias which could occur if only the term $k = 0$ in (8) is used in the expression for Φ_d , that is if the folding is not taken into account. In Table 1 the mean parameter values for $N_{MC} = 250$ Monte Carlo simulations are illustrated. Here $N_f = 0$ and $N_f = 5$ terms around $k = 0$ have been included in the sum (8). From the figure and the table we see that ignoring the effects of folding can produce very biased estimates.

Table 1. Mean Values of Parameters Estimates

System/Method	a_1	a_2	σ
True System	3	2	1
Folded ($N_f = 5$)	3.090567	2.037784	1.023398
Unfolded ($N_f = 0$)	4.978773	3.141028	1.639744

6. DIRECT CONTINUOUS-TIME ESTIMATION

By a direct method we mean one that only works directly with continuous time signals and spectra. In this case it will be the method (13).

The problem is that the sampled data periodogram $\hat{\Phi}_d(e^{i\omega T_s})$ can be readily found from the sampled data in (9) and (10), while we need the continuous-time periodogram $\hat{\Phi}_c(i\omega)$ in the criterion (13b). How can the latter be estimated or computed from $\hat{\Phi}_d(e^{i\omega T_s})$? We will be looking for relationships like

$$\hat{\Phi}_c(i\omega) = H(i\omega)\hat{\Phi}_d(e^{i\omega T_s}) \quad (15)$$

for a suitable function H . We could for instance assume that the signal is ZOH or FOH between the samples, but to do better than that, we argue as follows: If the true signal parameters θ_0 were known, we would have

$$\hat{\Phi}_d(e^{i\omega T_s}) \approx \Phi_d(e^{i\omega T_s}, \theta_0)$$

and

$$\hat{\Phi}_c(i\omega) \approx \Phi_c(e^{i\omega T_s}, \theta_0).$$

This means that the ideal transformation filter H in (15) would be

$$H(e^{i\omega T_s}) = \frac{\Phi_c(i\omega, \theta_0)}{\Phi_d(e^{i\omega T_s}, \theta_0)}. \quad (16)$$

Since θ_0 is unknown, we cannot construct H in this way, but the point is that when $T_s \rightarrow 0$, $H(e^{i\omega T_s})$ in (16) will approach something that does not depend on the signal parameters θ_0 , but only on the relative degree (pole excess) $l = n - m$ of the signal model in (1). This is related to the theory of sampling zeros, see e.g. (Åström and Sternby, 1984), (Wahlberg, 1988), (Weller *et al.*, 2001). This is what we will show now.

6.1 Estimation Method

Let the system in (1) be strictly proper, stable and $l = n - m$ be its *relative degree* (or *pole excess*), i.e. the difference between the number of poles and zeros of the system. Further assume that ω is below the Nyquist frequency. Define

$$\Phi_f^{(l)}(e^{i\omega T_s}) \triangleq \frac{\left| \frac{e^{i\omega T_s} - 1}{i\omega T_s} \right|^{2\ell}}{\left| \frac{B_{2\ell-1}(e^{i\omega T_s})}{(2\ell-1)!} \right|} \quad (17)$$

where $B_{2\ell-1}(z)$ are the so called Euler-Frobenius polynomials (Weller *et al.*, 2001). The polynomials are

$$B_\ell(z) = b_1^\ell z^{\ell-1} + b_2^\ell z^{\ell-2} + \dots + b_\ell^\ell \quad (18)$$

where

$$b_k^\ell = \sum_{m=1}^k (-1)^{k-m} m^\ell \binom{n+1}{k-m}, \quad k = 1, \dots, \ell.$$

Below is a list of polynomials for different values of ℓ

$$B_1(z) = 1$$

$$B_2(z) = z + 1$$

$$B_3(z) = z^2 + 4z + 1$$

$$B_4(z) = z^3 + 11z^2 + 11z + 1.$$

In Figure 2 we see that there is a very good correspondence between $\Phi_c(i\omega, \theta_0)/\Phi_d(i\omega, \theta_0)$ and $\Phi_f^{(l)}(i\omega)$ for the system in (14). This observation is verified by the following theoretical result.

Theorem 1. Consider the continuous time model (1) and assume that its pole excess is ℓ . Its spectrum is $\Phi_c(i\omega)$ given in (2). Let the spectrum of the sampled process (with sampling period T_s) be $\Phi_d(e^{i\omega T_s})$. Assume $\Phi_f^{(l)}$ is defined as in (17), that ω is less than the Nyquist frequency and that $\ell \geq 1$. Then

$$\left| \frac{\Phi_c(i\omega)}{\Phi_d(e^{i\omega T_s})} - \Phi_f^{(l)}(e^{i\omega T_s}) \right| < CT_s^{2\ell}$$

PROOF. Since ω is less than the Nyquist frequency

$$\frac{\Phi_c(i\omega + i\frac{2\pi}{T_s}k)}{\frac{\sigma^2}{|\omega + i\frac{2\pi}{T_s}k|^{2\ell}}} \rightarrow 1$$

as $T_s \rightarrow 0$ if $k \neq 0$. This has the consequence that

$$\frac{\Phi_c(i\omega)}{\Phi_d(e^{i\omega T_s})} \rightarrow \frac{\Phi_c(i\omega)}{\Phi_c(i\omega) + \sum_{k \neq 0} \frac{\sigma^2}{|\omega + i\frac{2\pi}{T_s}k|^{2\ell}}}$$

as $T_s \rightarrow 0$. From Lemma 3.2 in (Wahlberg, 1988)

$$\Phi_f^{(l)}(e^{i\omega T_s}) = \frac{1}{|\omega|^{2\ell} + \sum_{k \neq 0} \frac{1}{|\omega + i\frac{2\pi}{T_s}k|^{2\ell}}}.$$

By putting the two previous expressions on a common denominator, we get the the following relation

$$\frac{\Phi_c(i\omega)}{\Phi_d(e^{i\omega T_s})} - \Phi_f^{(l)}(e^{i\omega T_s}) \rightarrow \Phi_f^{(l)}(e^{i\omega T_s}) \Phi_r(i\omega) \Phi_s(i\omega)$$

where

$$\Phi_r(i\omega) = \frac{1 - \Phi_c(i\omega) \frac{|\omega|^{2\ell}}{\sigma^2}}{\Phi_c(i\omega) + \sum_{k \neq 0} \frac{\sigma^2}{|\omega + i\frac{2\pi}{T_s}k|^{2\ell}}}$$

and

$$\Phi_s(i\omega) = \sum_{k \neq 0} \frac{\sigma^2}{|\omega + i\frac{2\pi}{T_s}k|^{2\ell}}.$$

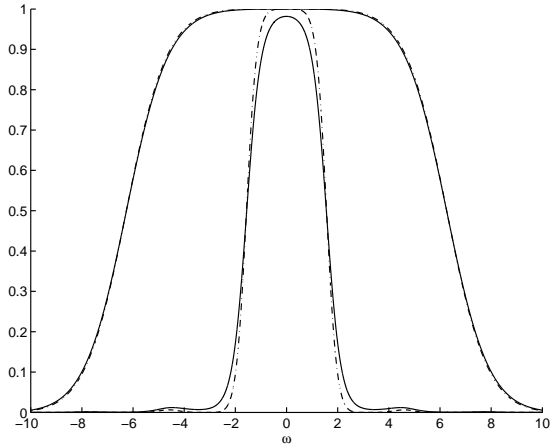


Fig. 2. Comparison of Φ_c/Φ_d (solid) to $\Phi_f^{(\ell)}$ (dash-dot) for the system in (14). In the inner plot $T_s = 2$ and in the outer plot $T_s = 0.5$.

Since $\Phi_f^{(\ell)}$ and Φ_r are bounded and

$$\frac{1}{|i\omega + i\frac{2\pi}{T_s}k|^{2\ell}} \leq C \left(\frac{T_s}{k}\right)^{2\ell}$$

if $k \neq 0$, the result

$$\left| \frac{\Phi_c(i\omega)}{\Phi_d(e^{i\omega T_s})} - \Phi_f^{(\ell)}(e^{i\omega T_s}) \right| \leq CT_s^{2\ell}$$

will follow.

Therefore a reasonable estimate of the continuous-time spectrum would be

$$\hat{\Phi}_c(i\omega) = \frac{\left| \frac{e^{i\omega T_s} - 1}{i\omega T_s} \right|^{2\ell}}{\left| \frac{B_{2\ell-1}(e^{i\omega T_s})}{(2\ell-1)!} \right|} \hat{\Phi}_d(i\omega).$$

6.2 Numerical Illustration

In Table 6.2 we have compared the performance of the estimator in the section above, Method 2, compared to the one described in Section 5.1, Method 1, for different sampling intervals. The means for each different sampling interval have been estimated by $N_{MC} = 250$ Monte Carlo simulations. The system is given by (14) and the correspondence between the mean parameter estimates in the table seems to be good.

7. CONCLUSIONS

Two parametric frequency-domain identification algorithms for continuous-time ARMA noise models have been presented. For low sampling rates, the Poisson summation formula is used to establish the exact distribution of the discrete-time Fourier transform of the output of the CARMA model. In the case of rapid sampling

T_s	Method	$a_1(=3)$	$a_2(=2)$	$\sigma(=1)$
0.6	1	3.0355	2.0189	1.0071
0.6	2	2.9935	2.0166	1.0007
0.5	1	3.0186	2.0206	1.0054
0.5	2	3.0476	2.0376	1.0150
0.4	1	3.0266	2.0150	1.0055
0.4	2	3.0449	2.0240	1.0110
0.3	1	3.0276	2.0143	1.0071
0.3	2	3.0327	2.0168	1.0086
0.2	1	3.0009	2.0118	1.0009
0.2	2	3.0017	2.0122	1.0011
0.1	1	3.0245	2.0210	1.0071
0.1	2	3.0246	2.0211	1.0072

Table 2. Comparison of mean values of parameter estimates from the indirect (Method 1) and direct (Method 2) estimators versus the sample time T_s . The system and circumstances are the same as in (14). The statistics for each sampling time has been estimated by $N_{MC} = 250$ Monte-Carlo Simulations.

the continuous-time spectrum is estimated from the discrete-time spectrum by means of a certain compensator. Further, numerical examples illustrate the efficiency of the different approaches. An interesting generalization study is the case of non-uniformly sampled data.

REFERENCES

- Brillinger, D.R. (1981). *Time Series*. Holden-Day. San Francisco, California.
- Gillberg, J. (2004). Methods for frequency domain estimation of continuous-time models. Lic. Dissertation Linköping Studies in Science and Technology Thesis No. 1133. Department of Electrical Engineering, Linköping University. SE-581 83 Linköping, Sweden.
- Johansson, R. (1994). Identification of continuous-time models. *IEEE Trans. on Signal Processing* **42**(4), 887–897.
- Larsson, E. K. and T. Söderström (2002). Continuous-time AR parameter estimation by using properties of sampled systems. In: *Proc. 15th IFAC World Congress, Barcelona, Spain*.
- Larsson, E.K. (2003). Identification of Stochastic Continuous-time Systems. PhD thesis. Department of Information Technology, Division of Systems and Control, Uppsala University, Uppsala, Sweden.
- Larsson, E.K. and M. Mossberg (2003). On possibilities for estimating continuous-time arma parameters. In: *Proc. IFAC Symposium on System Identification 2003, Rotterdam, The Netherlands*.
- Ljung, L. (1999). *Identification: Theory for the User*. Prentice-Hall. Upper Saddle River.
- Ljung, L. (2003a). Initialization aspects for subspace and output-error identification meth-

- ods. In: *Proc. 5th European Control Conference (ECC), Cambridge, UK.*
- Ljung, L. (2003b). *System Identification Toolbox for use with MATLAB. Version 6.* 6th ed.. The MathWorks, Inc. Natick, MA.
- Mensler, M. (1999). Analyse et etude comparative de methodes d'identification des systemes a representation continue. PhD thesis. Centre de Recherche en Automatique de Nancy, Universite Henri Poincare, Nancy.
- Papoulis, A. (1977). *Signal Analysis.* McGraw-Hill.
- Pearson, A. E. and Y. Shen (1993). Wiegthed least squares/MFT algorithms for linear differential system identification. In: *Proc. 32nd Conference on Decision and Control, San Antonio, Texas.*
- Pintelon, R. and J. Schoukens (1997). Identification of continuous-time systems using arbitrary signals. *Automatica* **33**(5), 991–994.
- Pintelon, R. and J. Schoukens (2000). Box-jenkins continuous-time modeling. *Automatica* **36**, 983–991.
- Rao, G. P. and H. Garnier (2002). Numerical illustrations of the relevance of direct continuous-time model identification. In: *Proc. 15th IFAC World Congress, Barcelona, Spain.*
- Söderström, T. (1991). Computing stochastic continuous-time models from ARMA models. *Int. J. Control* **53**(6), 1311–1326.
- Shinbrot, M. (1957). On the analysis of linear and nonlinear systems. *Trans. ASME* **79**, 547–552.
- Sihna, N. K. and G. P. Rao (1991). *Identification of continuous-time systems : methodology and computer implementation.* Kluwer. Dordrecht.
- Åström, K.J., Hagander P. and J. Sternby (1984). Zeros of sampled systems. *Automatica* **20**(1), 31–38.
- Unbehauen, M. and G.P. Rao (1990). Continuous-time approaches to system identification - a survey. *Automatica* **26**(1), 23–35.
- Wahlberg, B. (1988). Limit results for sampled systems. *International Journal of Control* **48**(3), 1267–1283.
- Weller, S. R., W. Moran, B. Ninnes and A. D. Pollinton (2001). Sampling zeros and the Euler-Frobenius polynomials. *IEEE Transactions on Automatic Control* **46**(2), 340–343.
- Whittle, P. (1961). Gaussian estimation in stationary time series. *Bulletin of the International Statistical Institute* **39**, 105–130.