On Robust Estimation of Power Spectra

Bernhard Spangl[1]and Rudolf Dutter[2]
[1]University of Natural Resources and Applied Life Sciences,
Vienna, Austria
[2]Vienna University of Technology, Austria

Abstract: Let us consider the problem of robust spectral density estimation. Conventional methods to obtain estimates of spectral density function are not robust in the presence of outlying observations. We present different methods to robustly estimate the spectral density function that are insensitive against outliers. The proposed methods are applied to simulated and real data and the results are compared. As a special practical application we focus on the frequency-domain analysis of short-term heart rate variability measurements of diabetes patients.

Keywords: Power Spectrum, Spectral Density Function, Robustness.

1   Introduction

The spectral density function is a commonly used tool when analyzing time series in the frequency domain. Areas of applications are signal processing (cf. Thomson, 1994), geophysics (cf. Chave et al., 1987; Jones and Hollinger, 1997) and medicine (cf. Hartikainen et al., 1998). In medicine the analysis of heart rate variability as non-invasive method is increasingly used (cf. Pumprla et al., 2002). In the present survey we have in mind the frequency-domain analysis of short-term heart rate variability measurements.

A commonly used model for outliers in time series is the additive outlier model (AO model) introduced by Fox (1972). The AO model consists of a stationary core process, x_t, to which occasional outliers have been added. The observed process $\{y_t : t \in \mathbb{Z}\}$ is said to have additive outliers if it is defined by

$$y_t = x_t + v_t \tag{1}$$

where the contaminations v_t are independent, identically distributed with $F_v = (1-\varepsilon)\delta_0 + \varepsilon H$ where δ_0 is the degenerated distribution having all its mass at the origin and H is a heavy-tailed symmetric distribution with mean zero and variance σ_H^2. Hence, the core process x_t is observed with probability $1 - \varepsilon$ whereas the core process plus a disturbance v_t is observed with probability ε. We shall also assume that x_t and v_t are independent.

When analyzing heart rate variability data the AO model seems to be an appropriate model. To access the variability of heart rate in the frequency domain the spectral density function of the tachogram is estimated. The tachogram is the series of time intervals between consecutive heart beats (e.g. Figure 1). These intervals are also called *R-R*-intervals, i.e. the periods between an R-peak and the next R-peak in an electrocardiogram.

In the tachogram, outlying observations can be caused by non-sinus ectopic beats and other artefacts. For example, if in the electrocardiogram an R-peak is missed this will result in very large value in the tachogram. On the other hand, if an ectopic beat occurs,
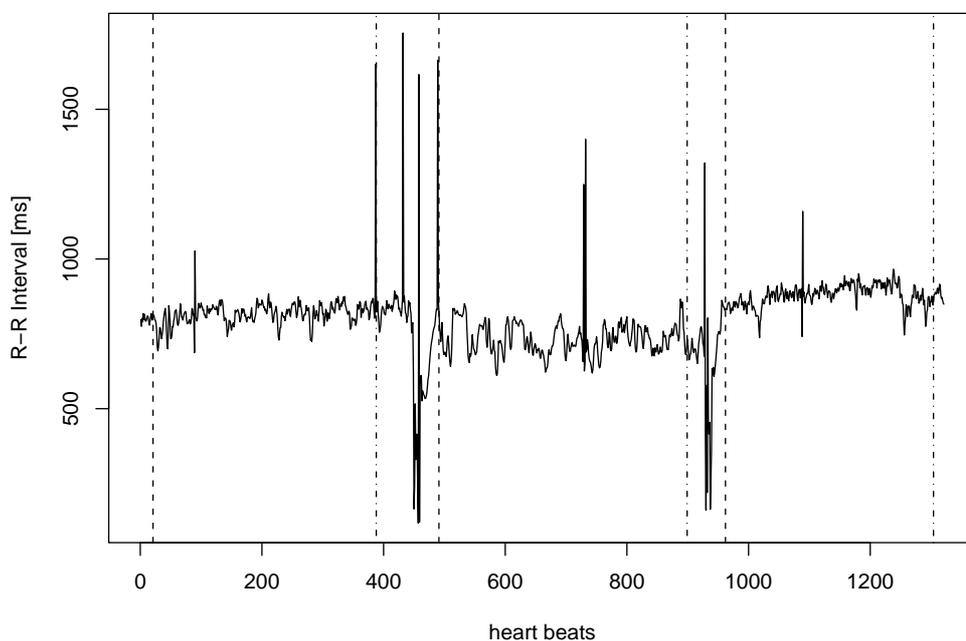
Figure 1: Tachogram of 1321 consecutive heart beats

i.e., if there is an extra heart beat between two regular beats, the heart beat following the ectopic beat will be very low and therefore usually missed. This results in a lower tachogram value followed by a higher one.

Unfortunately, conventional spectral density estimators are not robust in the presence of additive outliers. For details see Kleiner et al. (1979) or Martin and Thomson (1982).

Hence, in the next section we present different methods to robustly estimate the spectral density function that are insensitive against outliers. In Section 3, we present results of the previously described methods applied to simulated and real data. Conclusions are given in Section 4.

2   Robust Spectral Density Estimation

Before analyzing heart rate variability data using conventional frequency domain methods, extensive editing and review of the electrocardiogram by an experienced operator or physician is usually required to remove outlying observations, such as ectopic beats and other artefacts (cf. Pumprla et al., 2002). In the following we present robust methods to estimate the spectral density function that are insensitive against outliers.

2.1   A Robust Filter-cleaner Algorithm

The procedure to obtain a robust spectral density estimate proposed by Martin and Thomson (1982) incorporates an important data-cleaning operation wherein the robustness is introduced.

Let $\{y_t, t = 1, \ldots, N\}$ denote the observed values of a second-order stationary process with mean zero. The cleaning operator C maps the original data y_t into the cleaned data Cy_t. In the context of the AO model (1), we want the Cy_t to reconstruct the core process x_t, i.e., $Cy_t = \widehat{x}_t$, where $\widehat{x}_t$ denotes an estimate of x_t at time t. For AO models with a fraction of contamination ε not too large, it turns out that the data cleaner has the property that $Cy_t = y_t$ most of the time, that is about $(1 - \varepsilon) \times 100$ percent of the time. The kind of data cleaner described here is a robust filter-cleaner which uses the past and present observations $y_1, \ldots, y_t$ to produce the cleaned data Cy_t, $t = 1, \ldots, N$.

The filter-cleaner procedure involves a robust estimation of an autoregressive approximation to the core process x_t of order p, with estimated coefficients $\widehat{\phi}_{1,p}, \ldots, \widehat{\phi}_{p,p}$. Now, the residual process

$$r_t = Cy_t - \sum_{i=1}^{p} \widehat{\phi}_{i,p} Cy_{t-i}\,, \quad t = p+1, \ldots, N\,, \tag{2}$$

can easily be formed. Since cleaned data are used to obtain these residuals, and the $\widehat{\phi}_{i,p}$ are robust estimates, the transformation (2) is called a robust prewhitening operation. The benefit in the use of prewhitening in the context of spectral density estimation is to reduce leakage (cf. Blackman and Tukey, 1958).

The robust spectral density estimate is based on the above robust prewhitening as follows. Let

$$\widehat{H}_p(f) = 1 - \sum_{j=1}^{p} \widehat{\phi}_{j,p} \exp(i2\pi j f) \tag{3}$$

be the transfer function of the prewhitening operator (2), and let $\widehat{S}_r^{(lw)}(f)$ denote a lag window spectral estimate based on the residual process r_t. Then the spectral density estimate is

$$\widehat{S}(f) = \frac{\widehat{S}_r^{(lw)}(f)}{|\widehat{H}_p(f)|^2}\,, \tag{4}$$

where $\widehat{S}(f)$ is evaluated at the Fourier frequencies $f_k = k/N$, $k = 0, \ldots, [N/2]$.

The filter-cleaner algorithm as presented in the paper of Martin and Thomson (1982) relies on the AR(p) approximation of the underlying process x_t, represented in the following state-space form, with $t = p+1, \ldots, N$:

$$\boldsymbol{X}_t = \boldsymbol{\Phi} \boldsymbol{X}_{t-1} + \boldsymbol{U}_t\,, \tag{5}$$

where

$$\boldsymbol{X}_t = (x_t, x_{t-1}, \ldots, x_{t-p+1})^\top\,, \tag{6}$$
$$\boldsymbol{U}_t = (\varepsilon_t, 0, \ldots, 0)^\top\,, \tag{7}$$

$$\text{with } \mathbf{\Phi} = \begin{pmatrix} \phi_{1,p} & \cdots & \phi_{p-1,p} & \phi_{p,p} \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}, \quad \text{cov}(\mathbf{U}_t) = \mathbf{Q} = \begin{pmatrix} \sigma_{\varepsilon,p}^2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{8}$$

and $y_t = x_t + v_t = (1, 0, \ldots, 0)\mathbf{X}_t + v_t$ with $\text{var}(v_t) = \sigma_0^2$.

The algorithm computes robust estimates $\widehat{\mathbf{X}}_t$ of the unobservable $\mathbf{X}_t$ according to the following recursion:

$$\widehat{\mathbf{X}}_t = \mathbf{\Phi}\widehat{\mathbf{X}}_{t-1} + \frac{\mathbf{m}_t}{s_t^2} \, s_t \, \psi\left(\frac{y_t - \hat{y}_t^{t-1}}{s_t}\right) \tag{9}$$

with $\mathbf{m}_t$ being the first column of $\mathbf{M}_t$ which is computed recursively as

$$\mathbf{M}_{t+1} = \mathbf{\Phi}\mathbf{P}_t\mathbf{\Phi}^\top + \mathbf{Q} \tag{10}$$

$$\mathbf{P}_t = \mathbf{M}_t - w\left(\frac{y_t - \hat{y}_t^{t-1}}{s_t}\right) \frac{\mathbf{m}_t\mathbf{m}_t^\top}{s_t^2}. \tag{11}$$

The weight function $w(r) = \psi(r)/r$ where ψ stands for some robustifying function. The scale s_t is defined by

$$s_t^2 = m_{11,t} \tag{12}$$

and $\hat{y}_t^{t-1}$ denotes a robust one-step prediction of y_t based on $\mathbf{Y}^{t-1} = (y_1, \ldots, y_{t-1})^\top$, and is given by

$$\hat{y}_t^{t-1} = (\mathbf{\Phi}\widehat{\mathbf{X}}_{t-1})_1. \tag{13}$$

Finally, the cleaned process at time t results in

$$\hat{x}_t = (\widehat{\mathbf{X}}_t)_1. \tag{14}$$

It should be noted that if ψ is the identity function and $w \equiv 1$, and (12) is replaced by $s_t^2 = m_{11,t} + \sigma_0^2$ with $\sigma_0^2 = \text{var}(v_t)$ in the AO model, the above recursions are those of the Kalman filter. Correspondingly, M_t and P_t are the prediction and filtering error-covariance matrices (cf., for example, Jazwinski, 1970).

Important: To use the filter-cleaner algorithm we need robust estimates $\widehat{\boldsymbol{\phi}}_p$ and $\hat{\sigma}_{\varepsilon,p}^2 = s_{\varepsilon,p}^2$ of $\boldsymbol{\phi}_p = (\phi_{1,p}, \ldots, \phi_{p,p})^\top$ and $\sigma_{\varepsilon,p}^2$, respectively. Martin and Thomson (1982) got initial estimates using bounded-influence autoregression (BIAR) via iteratively reweighted least squares (IWLS). Details about BIAR may be found in Martin (1980). Initial estimates may also be obtained by robust autoregression using LTS- and LMS-regression. Details about LTS- and LMS-regression are in Rousseeuw and Leroy (1987). An alternative way is to use a highly robust autocovariance function estimator (cf. Ma and Genton, 2000) and calculate the initial estimates via the Yule-Walker equations. All these different approaches lead to similar results.

2.2 The Biweight Filter-cleaner Algorithm

Tatum and Hurvich (1993a) propose a frequency domain approach to the problem of cleaning outliers in time series. Their high breakdown method assumes only that the core process is Gaussian and has a continuous spectrum. It is nonparametric in the sense that it does not assume a finite parameter model, e.g. an ARMA model, for the core process. The approach uses robust trigonometric regression to fit a sine and cosine coefficient at each Fourier frequency and, hence, to obtain a robustified discrete Fourier transform. These coefficients are then inverse Fourier transformed to get a filtered version of the data. This procedure is termed biweight filter (BF). On this basis, a cleaned version of the data is constructed in which observations that appear to be outliers are replaced and most of the original series remain unchanged. The replacement values are found by a linear interpolation of all "non-outlying" data points. The interpolation is based on an estimate of the autocovariance function of the filtered series. This additional step is called biweight filter-cleaner (BFC).

2.2.1 The Biweight Filter

Any time series $\{y_t\}_{t=0}^{N-1}$ (note that we have changed the indexing in this section for ease of notation), can be represented as the sum of N cosines and sines at the Fourier frequencies $\omega_k = 2\pi k/N$, $k = 0, 1, \ldots, [N/2]$. The representation is

$$y_t = A_0 + \sum_{0 < k < N/2} \{A_k cos(\omega_k t) + B_k sin(\omega_k t)\} + (-1)^t A_{N/2} \, , \tag{15}$$

where the last term is only included if N is even.

The Fourier coefficients, $\{A_k\}_{k=0}^{[N/2]}$ and $\{B_k\}_{k=1}^{[N/2]-1}$, are identical to those that would be found by least squares regression of $\{y_t\}$ on the Fourier sinusoids, $\{cos(\omega_k t)\}$ and $\{sin(\omega_k t)\}$. If a core process, $\{x_t\}$, is subjected to contamination, then the Fourier transform of the resulting series $\{y_t\}$ will also reflect the contamination. Hence, inversion of the Fourier transform will simply return the contaminated data.

Tatum and Hurvich (1993a) propose a robust Fourier transform whose aim is to obtain estimates of the sine and cosine coefficients of the core process $\{x_t\}$ that are insensitive to the contaminating series $\{v_t\}$.

The robust regression is based on reducing the influence of large residuals by using Tukey's biweight ψ-function,

$$\psi(u) = \begin{cases} u(1 - u^2)^2 & |u| \leq 1 \\ 0 & |u| > 1 \, . \end{cases} \tag{16}$$

The cosine and sine coefficients at frequency ω_k, A_k' and B_k', are minimizing the function

$$\sum_{t=0}^{N-1} \rho(u_{k,t}) \tag{17}$$

where $\rho'(r) = \psi(r)$ and $u_{k,t} = \{y_t - A_k' cos(\omega_k t) - B_k' sin(\omega_k t)\}/(cS_k)$. S_k is a scale parameter of $\{u_{k,t}\}$, e.g. the median absolute deviation, and c is a tuning constant.

The regression can resist a high proportion of outliers as long as it is supplied with good starting values. Tatum and Hurvich (1993a) used Siegel's repeated median (cf. Siegel, 1982) to supply high breakdown initial values.

When the series is of prime length, the repeated median has an asymptotic breakdown point of 50%, as shown in Tatum and Hurvich (1993b). We will only consider this case in the following. In the case N is not prime, details can be found in Tatum and Hurvich (1993a).

The series is centered by removing a robust location estimate $\tilde{y}$. At each Fourier frequency, a sine and cosine coefficient is separately estimated by robust regression. The sum of the squared coefficients at each frequency gives a robust periodogram. Since the core process is assumed to have a continuous spectrum, the robust periodogram is smoothed using an appropriate lag window to obtain a lag window spectral estimate. Tatum and Hurvich (1993a) suggest to determine the amount of smoothing by the frequency domain version of the corrected AIC proposed by Hurvich and Beltrão (1990).

In the biweight filter the order in which the frequencies will be used is determined by this smoothed periodogram, largest periodogram values first. Using that order, a sinusoid is fitted at each Fourier frequency by robust regression and then swept out of the residuals from the previous step.

The cosine and sine coefficients found by the biweight filter algorithm are then inverse Fourier transformed to give a filtered series, y_t^F, $t = 0, \ldots, N-1$, where

$$y_t^F = \tilde{y} + \sum_{0 < k < N/2} \left\{ A_k' cos(\omega_k t) + B_k' sin(\omega_k t) \right\} . \tag{18}$$

2.2.2   The Biweight Filter-cleaner

On the basis of the biweight filter, Tatum and Hurvich (1993a) develop a biweight filter-cleaner which has output exactly equal to the input for most values and interpolates the remaining values.

The biweight filter-cleaner initially compares the filtered series to the original and flags discrepant points using the residuals relative to a robust scale estimate of the residuals. The flagged values are replaced by linear interpolation. Finally, the estimated interpolation variance is used to compare the distance between the interpolated points and their original values. When this distance is "relatively small", the original observation is reintroduced. "Relatively small" here means in respect to the estimation (or interpolation) error of the *filter-cleaned* value. This last step helps to counterbalance the tendency of the first step to flag too many uncontaminated observations.

3   Results

We concentrate on the robust filter-cleaner algorithm by Martin and Thomson (1982) and the biweight filter-cleaner transform by Tatum and Hurvich (1993a), and present results of these methods applied to simulated and real data. Both methods yield robust estimates of the spectral density function by cleaning the time series in a robust way first and calculating the spectral density function afterwards.
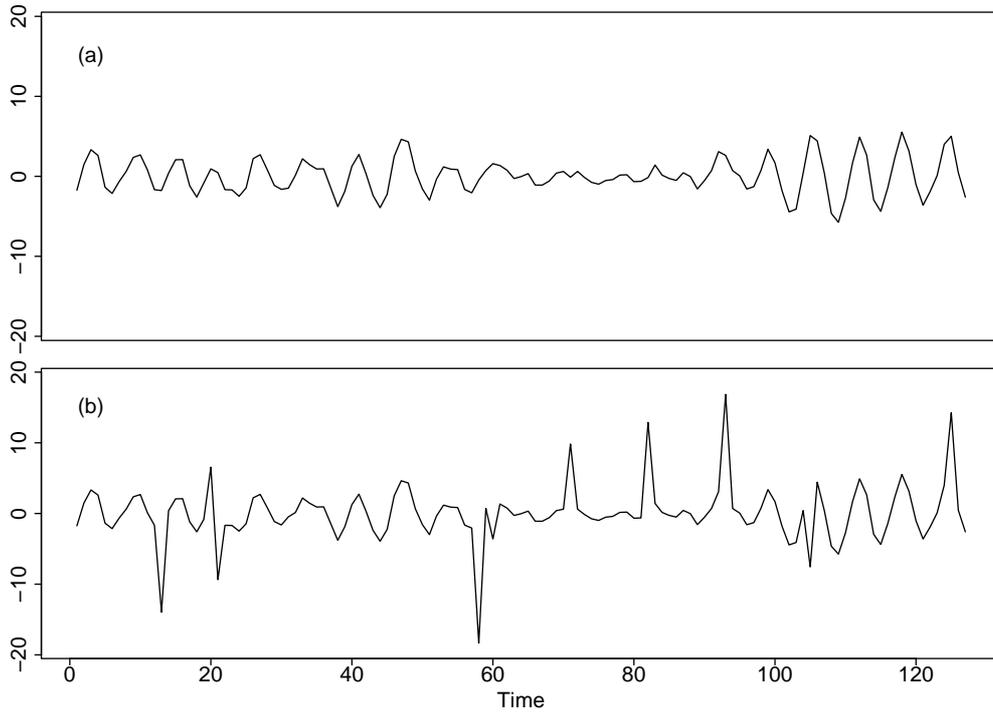
Figure 2: Plot of simulated AR(2) process: core process (a), with AO's (b)

For the robust filter-cleaner algorithm the order p of the autoregressive approximation is determined by AIC. As robustifying function Hampel's ψ-function is used.

3.1 Simulated Data

We consider a simulated AR(2) process of length $N = 127$ given by

$$x_t = x_{t-1} - 0.9x_{t-2} + \varepsilon_t \, , \tag{19}$$

with $\varepsilon_t \sim N(0, 1)$. To obtain the process y_t contaminated by additive outliers, noise from $0.9\delta_0 + 0.1N(0, 100)$ is added to the core process x_t (cf. Figure 2).

As shown in Figure 3, the resulting series obtained by both filter-cleaner algorithms approximate the underlying core process x_t quite well. All outliers are revealed, even relatively small ones, compared to the variability of the neighboring data points, e.g., at time $t = 60$ and $t = 105$. However, the biweight filter-cleaner algorithm slightly underestimates the core process, especially at time $t = 82$, $t = 93$, and $t = 125$.

Figure 4 shows spectral density estimates of the simulated process. The dot-dashed line represents the spectral density estimate of the filter-cleaned process obtained by the robust filter-cleaner algorithm, whereas the thin solid line is the one of the filter-cleaned data set computed by the biweight filter-cleaner procedure. Since both filter-cleaner yield similar results that approximate the underlying core process well (cf. Figure 3), both spectral density estimates are similar in shape and power and approximate the true spectral density function of the core process (thick solid line) quite well, too.
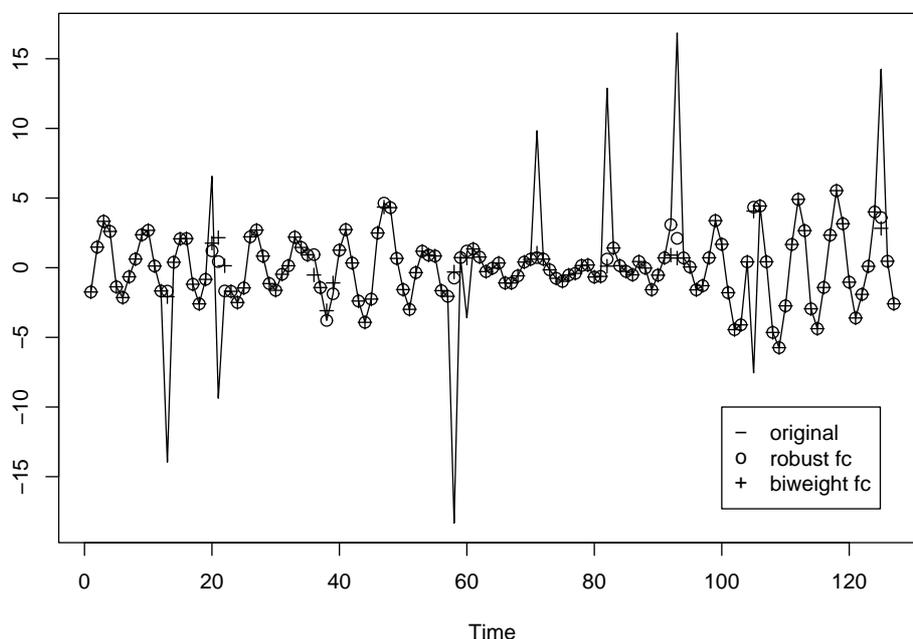
Figure 3: Plot of the filter-cleaned process

3.2 Heart Rate Variability Data

The heart rate variability data consist of *R-R*-intervals of 1321 consecutive heart beats and were provided by J. Pumprla and K. Howorka, Department of Biomedical Engineering and Physics, General Hospital of Vienna. A tachogram of the data is plotted in Figure 1.

For further analysis only the first 256 seconds of the heart rate variability data are considered. In Figure 5 this part of the original tachogram along with the results of both filter-cleaner algorithms is shown. The only outlying observations in this tachogram sequence that were identified by an expert are those around 55 seconds. This contamination is also revealed by both filter-cleaner algorithms. However, the biweight filter-cleaner procedure identifies additional outliers around 10, 20, and 220 seconds. Again, it tends to underestimate the underlying core process, whereas the robust filter-cleaner algorithm approximates the core process in a better way.

Figure 6 shows spectral density estimates of the tachogram sequence. The spectral density estimate of the filter-cleaned process obtained by the robust filter-cleaner algorithm (dot-dashed line) is very similar in shape and power to the spectral density estimate of the tachogram that was cleaned by the physician (thick solid line). This corresponds to the extremely good approximation of the core process by the robust filter-cleaned tachogram sequence (cf. Figure 5).
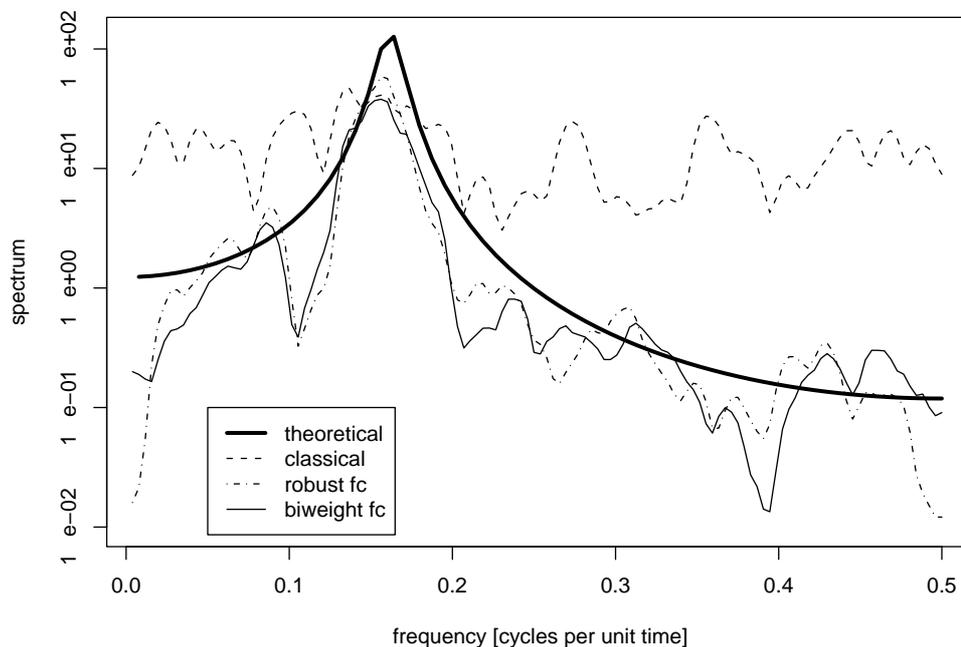
Figure 4: Plot of the spectral density estimates of the simulated process

4   Conclusions

In order to get a robust estimate of the spectral density function, that is insensitive against outlying observations, it turned out that cleaning the time series in a robust way first and calculating the spectral density function afterwards leads to encouraging results.

The robust filter-cleaner algorithm proposed by Martin and Thomson (1982) approximates the underlying core process very well, especially, if there are only few outliers present. The biweight filter-cleaner procedure proposed by Tatum and Hurvich (1993a) leads also to good approximations, but tends to underestimate the core process.

Varying the tuning constants and remaining parameters of the biweight filter-cleaner algorithm to solve the problem of underestimation seems worth trying. Moreover, further research and additional simulation studies have to be done.

References

R.B. Blackman and J.W. Tukey. *The Measurement of Power Spectra.* Dover, New York, 1958.

A.D. Chave, D.J. Thomson, and M.E. Ander. On the robust estimation of power spectra, coherences, and transfer functions. *J. Geophys. Res.*, 92(B1):633–648, 1987.

A.J. Fox. Outliers in time series. *J. Royal Statist. Soc.*, 34(3):350–363, 1972.
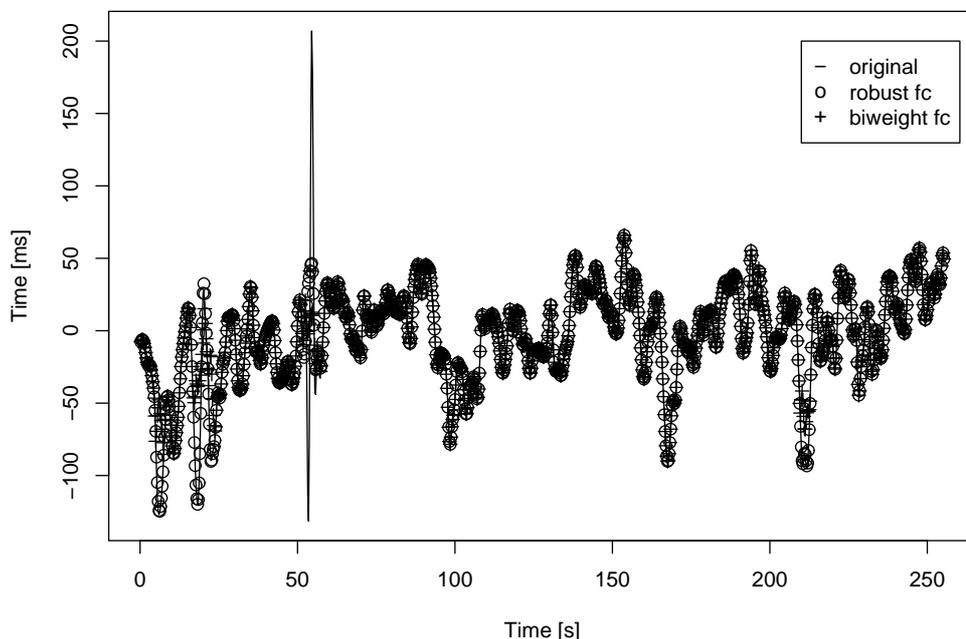
Figure 5: Plot of first 256 seconds of the filter-cleaned tachogram

J.E.K. Hartikainen, K.U.O. Tahvanainen, and T.A. Kuusela. Short-term measurement of heart rate variability. In Malik, editor, *Clinical Guide to Cardiac Autonomic Tests*, pages 149–176. Kluwer, Dordrecht, 1998.

C.M. Hurvich and Beltrão. Cross-validatory choice of a spectrum estimate and its connections with AIC. *J. Time Series Analysis*, 11(2):121–137, 1990.

A. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, New York, 1970.

A.G. Jones and K. Hollinger. Spectral analysis of the KTB sonic and density logs using robust nonparametric methods. *J. Geophys. Res.*, 102(B8):18391–18403, 1997.

R. Kleiner, R.D. Martin, and D.J. Thomson. Robust estimation of power spectra. *J. Royal Statist. Soc. B*, 41(3):313–351, 1979.

Y. Ma and M.G. Genton. Highly robust estimation of the autocovariance function. *J. Time Series Analysis*, 21(6):663–684, 2000.

R.D. Martin. Robust estimation of autoregressive models. In D.R. Brillinger and G.C. Tiao, editors, *Directions in Time Series*, pages 228–254. Inst. Math. Statist. Publications, Haywood, CA, 1980.

R.D. Martin and D.J. Thomson. Robust-resistant sprectrum estimation. In *Proceedings of the IEEE*, volume 70, pages 1097–1115. IEEE, 1982.
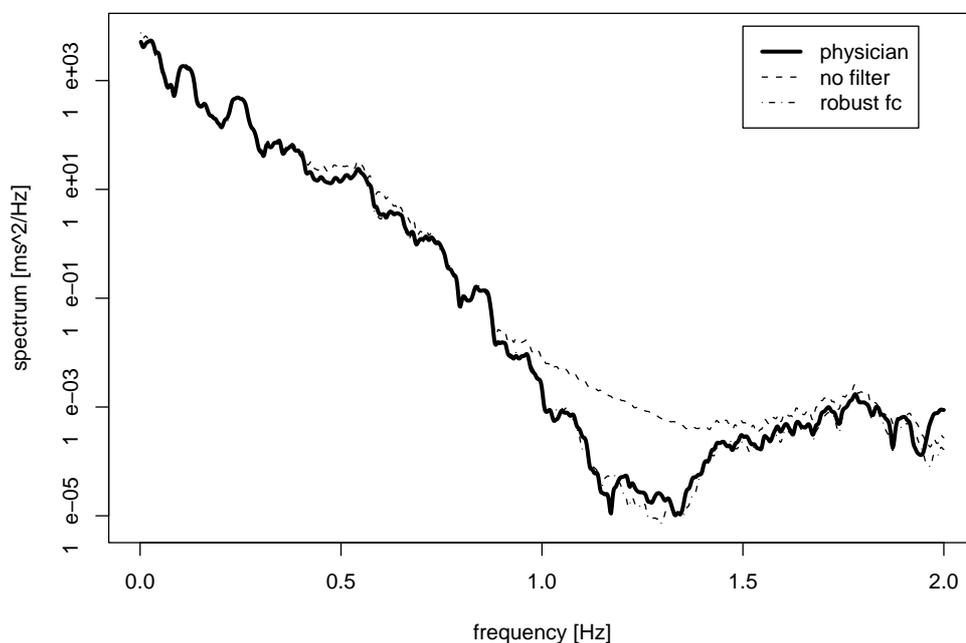
Figure 6: Plot of the spectral density estimates of the tachogram

J. Pumprla, K. Howorka, D. Groves, M. Chester, and J. Nolan. Functional assessment of heart rate variability: physiological basis and practical applications. *Int. J. Cardiology*, 84:1–14, 2002.

P.J. Rousseeuw and A.M. Leroy. *Robust Regression and Outlier Detection.* John Wiley & Sons, New York, 1987.

A.F. Siegel. Robust regression using repeated medians. *Biometrika*, 69:242–244, 1982.

L.G. Tatum and C.M. Hurvich. A frequency domain approach to robust time series analysis. In Morgenthaler, Ronchetti, and Stahel, editors, *New Directions in Statistical Data Analysis and Robustness*. Birkhäuser-Verlag, Basel, 1993a.

L.G. Tatum and C.M. Hurvich. High breakdown methods of time series analysis. *J. Royal Statist. Soc. B*, 55(4):881–896, 1993b.

D.J. Thomson. An overview of multiple-window and quadratic-inverse spectrum estimation methods. In *Proceedings of the IEEE ICASSP*, volume 6, pages 185–194. IEEE, 1994.

Authors' addresses:

Bernhard Spangl
Group Applied Statistics and Computing
Dept. of Spatial, Landscape, and Infrastructure Sciences
University of Natural Resources and Applied Life Sciences
Gregor-Mendel-Str. 33, A-1180 Vienna, Austria

Tel. +43 1 47654 / 5063
E-mail: Bernhard.Spangl@boku.ac.at

Rudolf Dutter
Department of Statistics and Probability Theory
Vienna University of Technology
Wiedner Hauptstr. 8-10
A-1040 Vienna
Austria

Tel. +43 1 58801 / 10730
Fax +43 1 58801 / 10799
E-mail: R.Dutter@tuwien.ac.at