Estimation of Power Spectra via Robust Filtering

B. Spangl and R. Dutter

BOKU - University of Natural Resources and Applied Life Sciences

and

Vienna University of Technology

(joint work with J. Pumprla and K. Howorka,
General Hospital of Vienna)

Österreichische Statistiktage 2003, Oct. 2003

Contents

Motivation

Analysis of Heart Rate Variability (HRV) as a non-invasive method for diabetic patients

Second-order Stationarity

Consider a time series $\{x_t : t \in \mathbb{Z}\}$ and assume that it satisfies the hypothesis of second-order stationarity:

(i) $E(x_t^2) < \infty \quad \forall\ t \in \mathbb{Z}$

(ii) $E(x_t) = \mu = \text{constant} \quad \forall\ t \in \mathbb{Z}$

(iii) $\text{cov}(x_{t+h}, x_t) = \gamma(h) \quad \forall\ t, h \in \mathbb{Z}$

where $\gamma(h)$ is the autocovariance function of x_t at lag h.

ARMA(p, q) Process

A univariate time series x_t, $t = 0, \pm 1, \pm 2, \ldots$, is said to be an ARMA($p, q$) process if x_t is stationary and

$$x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q} \ , \qquad (1)$$

with $\phi_p \neq 0$ and $\theta_q \neq 0$. The innovation process w_t in (1) is a white noise process with $\sigma_w^2 > 0$.

Additive Outliers (AO)

The process $\{y_t : t \in \mathbb{Z}\}$ is said to have additive outliers (AO) if it is defined by

$$y_t = x_t + v_t \tag{2}$$

where x_t is an ARMA(p, q) process and the v_t are independent identically distributed (iid) with common distribution $F_v = (1 - \varepsilon)\delta_0 + \varepsilon H$ where δ_0 is the degenerate distribution having all its mass at the origin and H is a heavy-tailed symmetric distribution with mean 0 and variance σ^2. Hence, the ARMA(p, q) process x_t is observed with probability $1 - \varepsilon$ whereas the ARMA(p, q) process plus an error v_t is observed with probability ε. We shall also assume that x_t and v_t are independent.

Spectral Representation Theorem

Any second-order stationary process x_t has the spectral representation

$$x_t = \int_{-1/2}^{1/2} \exp(-2\pi i t f)\, dZ(f) \tag{3}$$

where $Z(f)$, $-1/2 \leq f \leq 1/2$, has stationary uncorrelated increments. The process $Z(f)$ defines a monotone nondecreasing function $F(f)$ through

$$F(f) = E|Z(f)|^2, \quad dF(f) = E|dZ(f)|^2, \quad F(-1/2) = 0, \quad F(1/2) = \sigma^2 = \gamma(0)\,.$$

The function $F(f)$ is called the spectral distribution function. If the autocovariance function is absolutely summable, i.e., $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, the spectral distribution function is absolutely continuous with $dF(f) = S(f)\, df$. $S(f)$ is called the *spectral density function* of the process x_t. Other commonly used terms for $S(f)$ are *spectral density*, *spectrum* or *power spectrum*.

Definition of the Power Spectrum

It follows that any process with autocovariance function $\gamma(h)$ satisfying

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty \tag{4}$$

has the representation

$$\gamma(h) = \int_{-1/2}^{1/2} \exp(2\pi i h f) S(f)\, df, \qquad h \in \mathbb{Z}\ . \tag{5}$$

Thus the autocovariance at lag h, $\gamma(h)$, $h \in \mathbb{Z}$, are the Fourier coefficients of $S(f)$ and so

$$S(f) = \sum_{h=-\infty}^{\infty} \gamma(h) \exp(-2\pi i h f)\ . \tag{6}$$

Non-parametric Estimation

The non-parametric method of estimating the spectrum is based on smoothing the periodogram. Let x_t, $t = 1, \ldots, n$, be the observed time series. Then the discrete Fourier transform of x_t

$$X(f_k) = n^{-1/2} \sum_{t=1}^{n} x_t \exp(-\frac{2\pi i t k}{n}) \ , \quad f_k = \frac{k}{n} \ , \quad k = -\left[\frac{n-1}{2}\right], \ldots, \left[\frac{n}{2}\right] \tag{7}$$

is computed using the fast Fourier transformation and the periodogram is formed by

$$I(f_k) = |X(f_k)|^2 \ . \tag{8}$$

The periodogram is smoothed to get an estimate of the power spectrum

$$\widehat{S}(f_k) = \sum_{j=-L}^{L} w_j I(f_{k+j}) \ , \quad w_j = w_{-j} \ , \quad \sum_{j=-L}^{L} w_j = 1 \ . \tag{9}$$

Parametric Estimation

We now consider basing a spectral estimator on the parameters of a p-th order autoregressive model, i.e., an AR(p) model

$$x_t - \sum_{k=1}^{p} \phi_k x_{t-k} = z_t \ , \tag{10}$$

where z_t is a white noise process with mean zero and variance σ_z^2.

Substituting the maximum likelihood or least squares estimators of the model parameters, denoted by $\widehat{\phi}_1, \ldots, \widehat{\phi}_p$ and $\widehat{\sigma}_z^2$, we obtain a parametric estimate of the power spectrum

$$\widehat{S}(f) = \frac{\widehat{\sigma}_z^2}{\left|1 - \sum_{k=1}^{p} \widehat{\phi}_k \exp(-2\pi i k f)\right|^2} \ . \tag{11}$$

Estimates Based on Robust Prewhitening and Filtering

Let $\{y_k, k = 1, \ldots, n\}$ denote the observed values of a second-order stationary process with mean zero. Kleiner et al. (1979) prefer the following autoregression prewhitened spectral density estimate which was originally suggested by Blackman & Tukey (1958):

$$\widehat{S}(f) = \frac{\widehat{S}_r(f)}{|\widehat{H}_p(f)|^2} \,, \tag{12}$$

where $\widehat{S}_r(f)$ is a smoothed periodogram estimate of the prediction residuals $r_k = y_k - \sum_{j=1}^{p} \widehat{\phi}_j y_{k-j}$, $k = p + 1, \ldots, n$ and

$$\widehat{H}_p(f) = 1 - \sum_{j=1}^{p} \widehat{\phi}_j \exp(i2\pi jf) \,. \tag{13}$$

The basic proposal is to robustify both the numerator $\widehat{S}_r(f)$ and the denominator $\widehat{H}_p(f)$ in (12).

Robust Filter-Cleaner (Part I)

The filter-cleaner algorithm rely on the AR(p) approximation of the process x_t, represented in the following state-space form with $t = p+1, \ldots, n$:

$$\boldsymbol{X}_t = \boldsymbol{\Phi} \boldsymbol{X}_{t-1} + \boldsymbol{U}_t \ , \tag{14}$$

where

$$\boldsymbol{X}_t = (x_t, x_{t-1}, \ldots, x_{t-p+1})^\top \ , \tag{15}$$

$$\boldsymbol{U}_t = (\varepsilon_t, 0, \ldots, 0)^\top \ , \tag{16}$$

$$\text{with } \boldsymbol{\Phi} = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \ , \quad \boldsymbol{Q} = \begin{pmatrix} \sigma_\varepsilon^2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \text{ and} \tag{17}$$

$$y_t = x_t + v_t = (1, 0, \ldots, 0) \boldsymbol{X}_t + v_t \quad \text{with} \quad \boldsymbol{R} = \mathrm{var}(v_t) = \sigma_0^2 \ . \tag{18}$$

Robust Filter-Cleaner (Part II)

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

$$\widehat{\boldsymbol{X}}_t = \boldsymbol{\Phi}\widehat{\boldsymbol{X}}_{t-1} + \frac{\boldsymbol{m}_t}{s_t^2}\ s_t\ \psi\left(\frac{y_t - \widehat{y}_t^{t-1}}{s_t}\right) \tag{19}$$

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

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}\boldsymbol{P}_t\boldsymbol{\Phi}^\top + \boldsymbol{Q} \tag{20}$$

$$\boldsymbol{P}_t = \boldsymbol{M}_t - w\left(\frac{y_t - \widehat{y}_t^{t-1}}{s_t}\right)\frac{\boldsymbol{m}_t\boldsymbol{m}_t^\top}{s_t^2}\ . \tag{21}$$

The scale s_t is defined by $s_t^2 = m_{11,t}$ and $\widehat{y}_t^{t-1}$ denotes a robust one-step prediction of y_t based on $\boldsymbol{Y}^{t-1} = (y_1, \ldots, y_{t-1})^\top$, and is given by

$$\widehat{y}_t^{t-1} = (\boldsymbol{\Phi}\widehat{\boldsymbol{X}}_{t-1})_1\ . \tag{22}$$

Finally, the cleaned process at time t is

$$\widehat{x}_t = (\widehat{\boldsymbol{X}}_t)_1\ . \tag{23}$$

Robust Filter-Cleaner (Part III)

To use the filter-cleaner algorithm we need robust estimates $\widehat{\boldsymbol{\phi}}$ and $\hat{\sigma}_\varepsilon^2 = \hat{s}_\varepsilon^2$ of $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)^\top$ and σ_ε^2.

Until now, we have tried two different approaches to obtain initial estimates

- using bounded-influence autoregession (BIAR) via iterated re-weighted least squares (IWLS) (cf. Martin & Thomson, 1982) or

- a highly robust autocovariance function estimator (cf. Ma & Genton, 2000) and the Yule-Walker equations

$\Rightarrow$ similar results

Highly Robust Autocovariance Function

Ma & Genton (2000) suggest a highly robust autocovariance function estimator $\widehat{\gamma}_Q(h, \boldsymbol{x})$ which is defined as follows. Extract the first $n - h$ observations of $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ to produce a vector $\boldsymbol{u}$ of length $n - h$ and the last $n - h$ observations of $\boldsymbol{x}$ to produce a vector $\boldsymbol{v}$ of length $n - h$. Then:

$$\widehat{\gamma}_Q(h, \boldsymbol{x}) = \frac{1}{4}[Q^2_{n-h}(\boldsymbol{u} + \boldsymbol{v}) - Q^2_{n-h}(\boldsymbol{u} - \boldsymbol{v})] \ . \tag{24}$$

The function $Q_n(\boldsymbol{z})$ (cf. Croux & Rousseeuw, 1992) is a highly robust estimator of scale and is defined by

$$Q_n(\boldsymbol{z}) = c\{|z_i - z_j| : i < j\}_{(k)} \ , \tag{25}$$

where $\boldsymbol{z} = (z_1, \ldots, z_n)^\top$ is the sample and

$$k = \binom{h}{2} \ , \ \text{with } h = \left[\frac{n}{2}\right] + 1 \ ,$$

where [.] denotes the integer part. The factor $c = 2.2191$ is for consistency.

Artificial Data

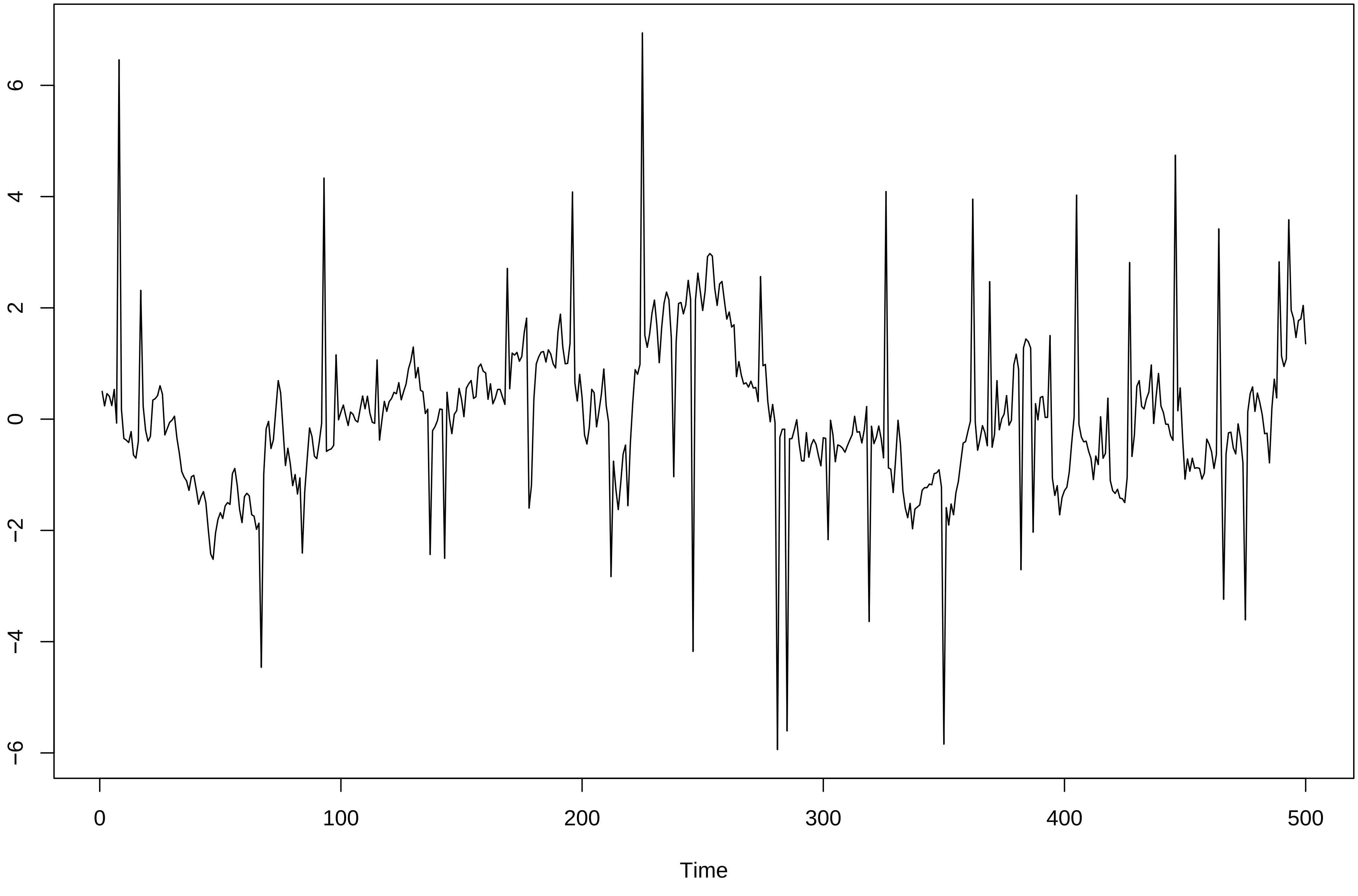
This example is a artificial process with AO composed of the following three autoregressive processes:

$$
\begin{aligned}
u_k &= 0.975 u_{k-1} + \varepsilon_k \ , \\
w_k &= 0.95 w_{k-1} - 0.9 w_{k-2} + \eta_k \ , \\
z_k &= 0.33 z_{k-1} - 0.9 z_{k-2} + \zeta_k \ ,
\end{aligned}
$$

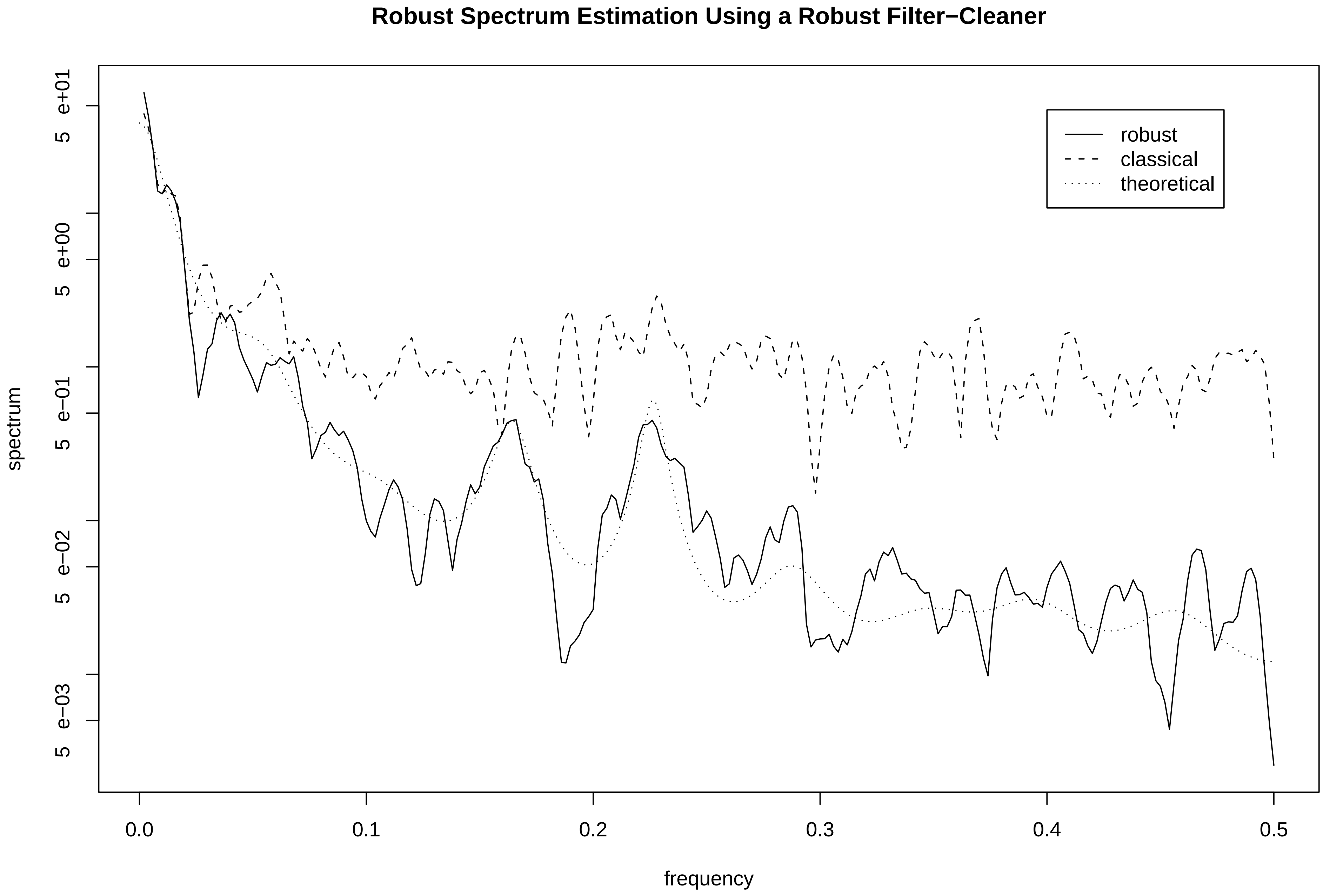
with $\varepsilon_k, \eta_k, \zeta_k \sim N(0,1)$. We standardize u_k, w_k and z_k, and compute

$$y_k = \sqrt{75} u_k + w_k + z_k \ , \quad k = 1, \ldots, n \ ,$$

and then standardize y_k again. Additionally we add some noise from $0.9\delta_0 + 0.1N(0,10)$.



6
4
2
0
−2
−4
−6
0
100
200
300
400
500
Time

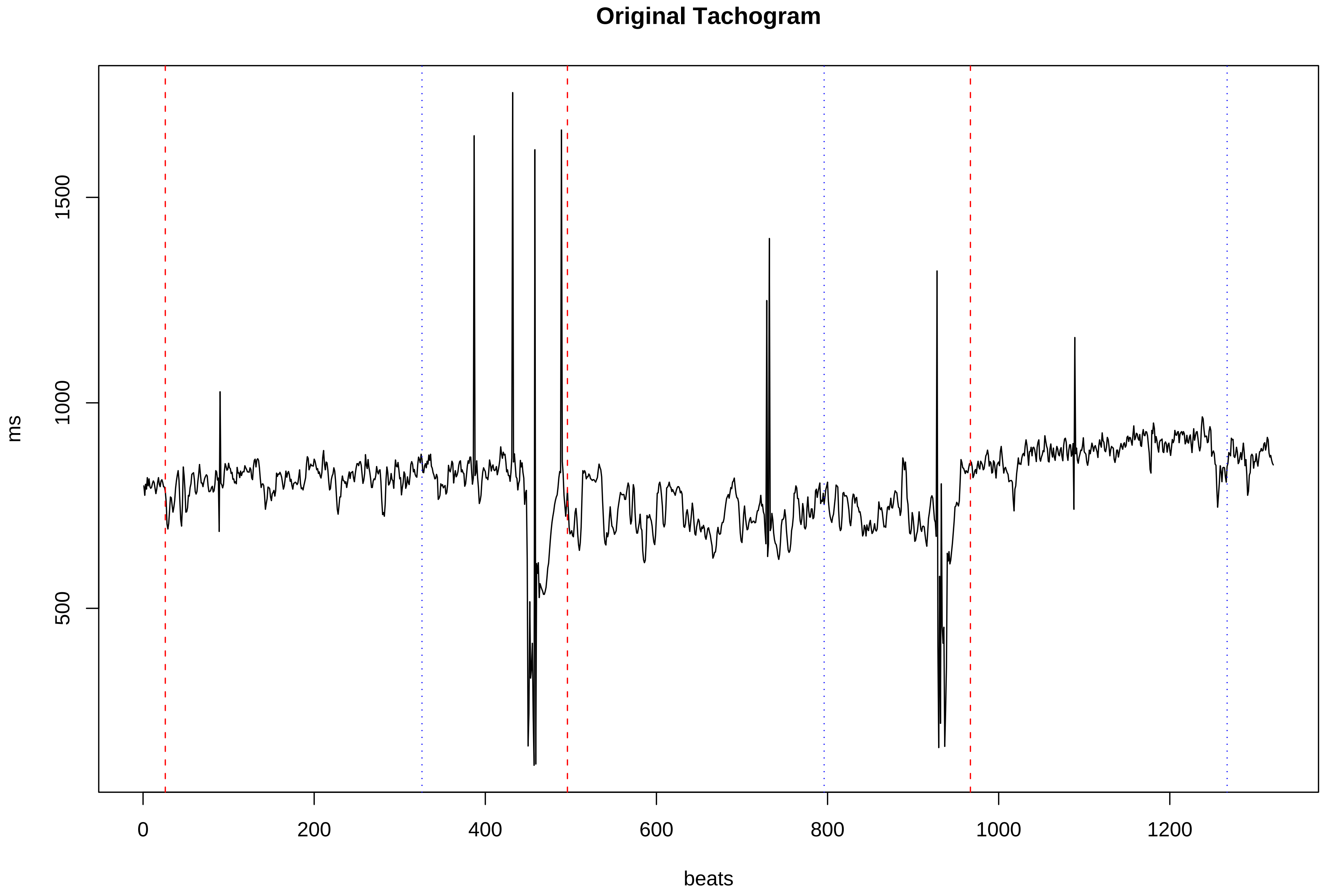


Robust Spectrum Estimation Using a Robust Filter−Cleaner
robust
classical
theoretical
spectrum
frequency
0.0
0.1
0.2
0.3
0.4
0.5
5 e−03
e−03
5 e−02
e−02
5 e−01
e−01
5 e+00
e+00
5 e+01
e+01

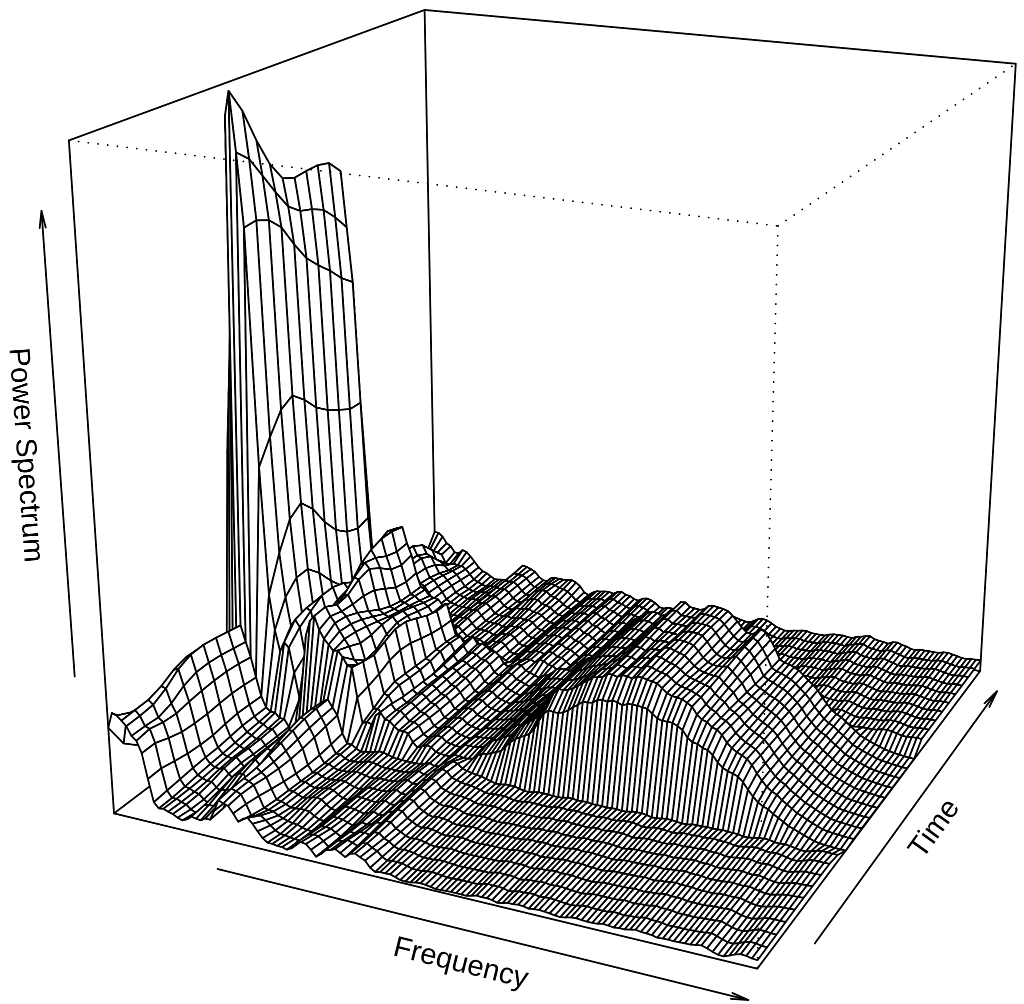
Heart Rate Variabiliy Data

- data: heart rate variability (HRV) recordings (tachogram of 1321 successive heart beats)

- provided by J. Pumprla and K. Howorka, Department of Biomedical Engineering and Physics, General Hospital of Vienna

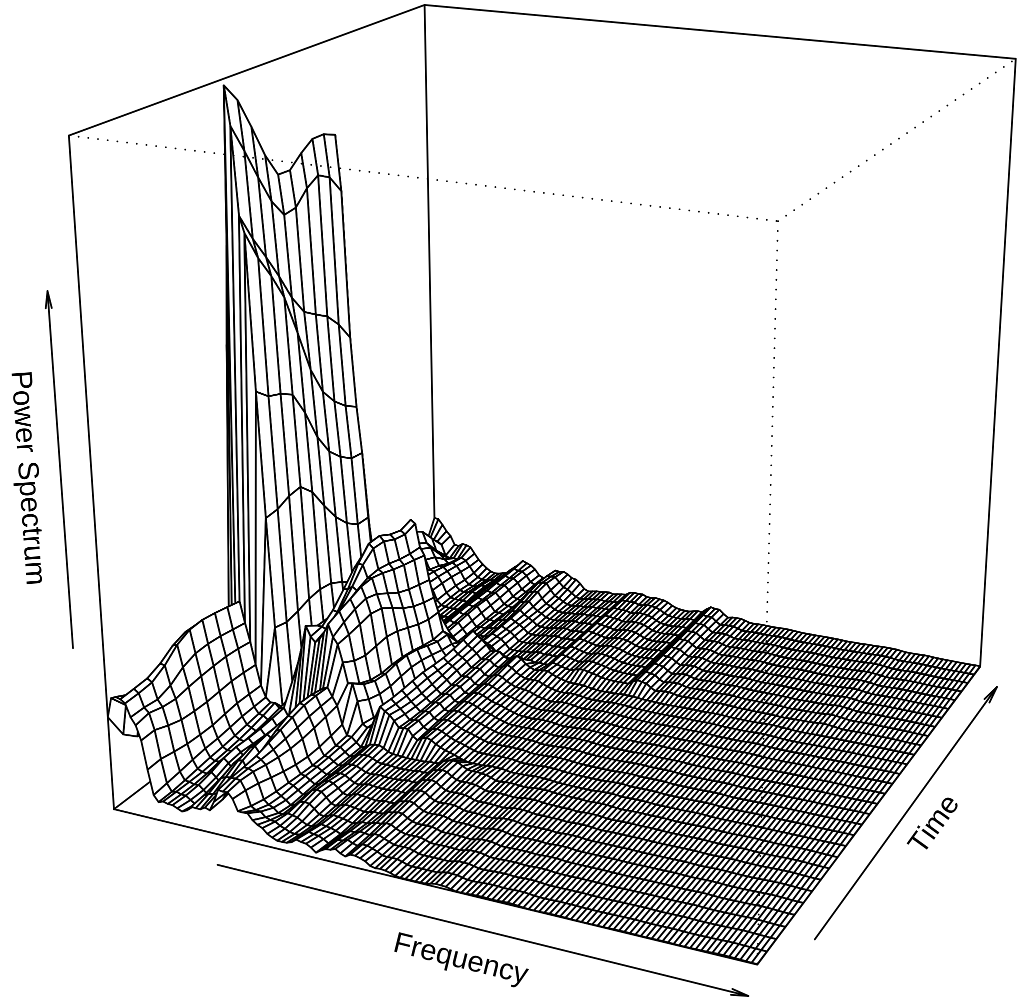

Original Tachogram
ms
1500
1000
500
0
200
400
600
800
1000
1200
beats

Classical Short−term Analysis of HRV with AO



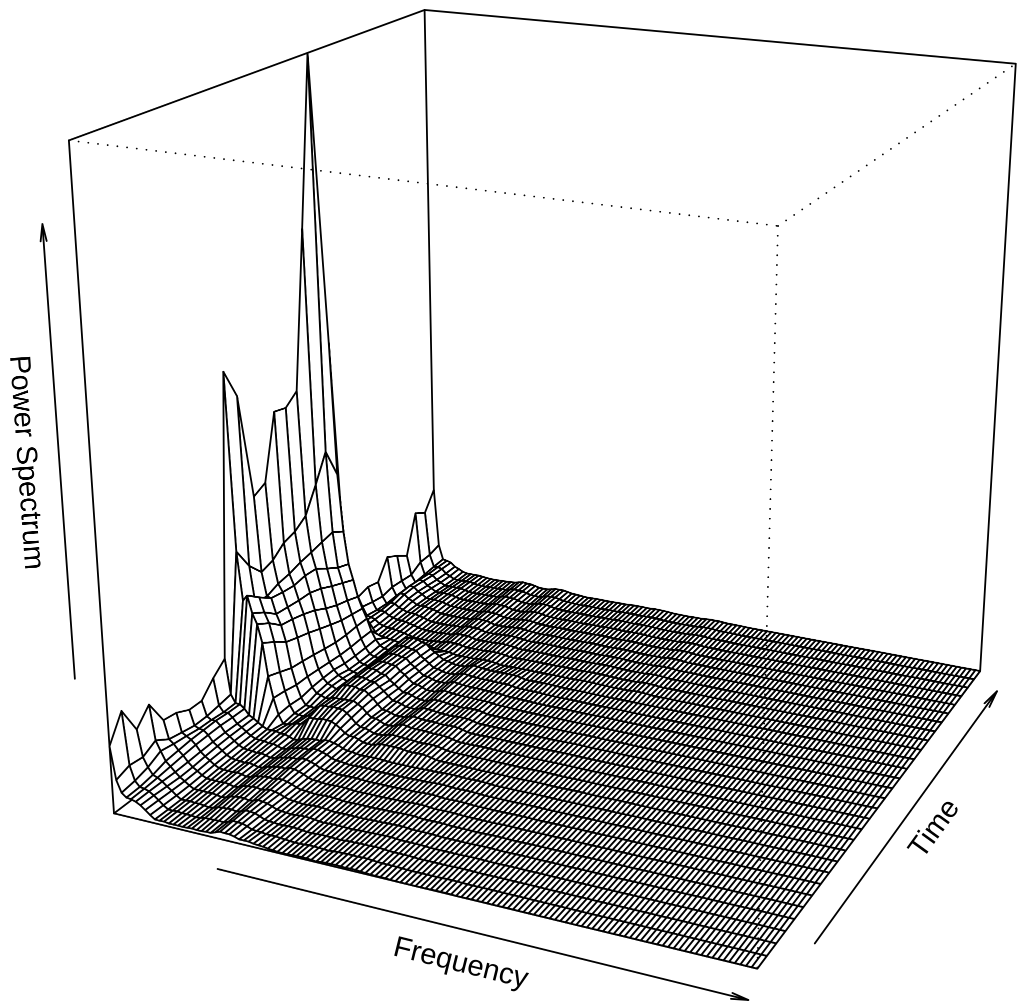
Maximum: 44.05 *10^3 [ms^2]

Classical Short−term Analysis of HRV without Outliers



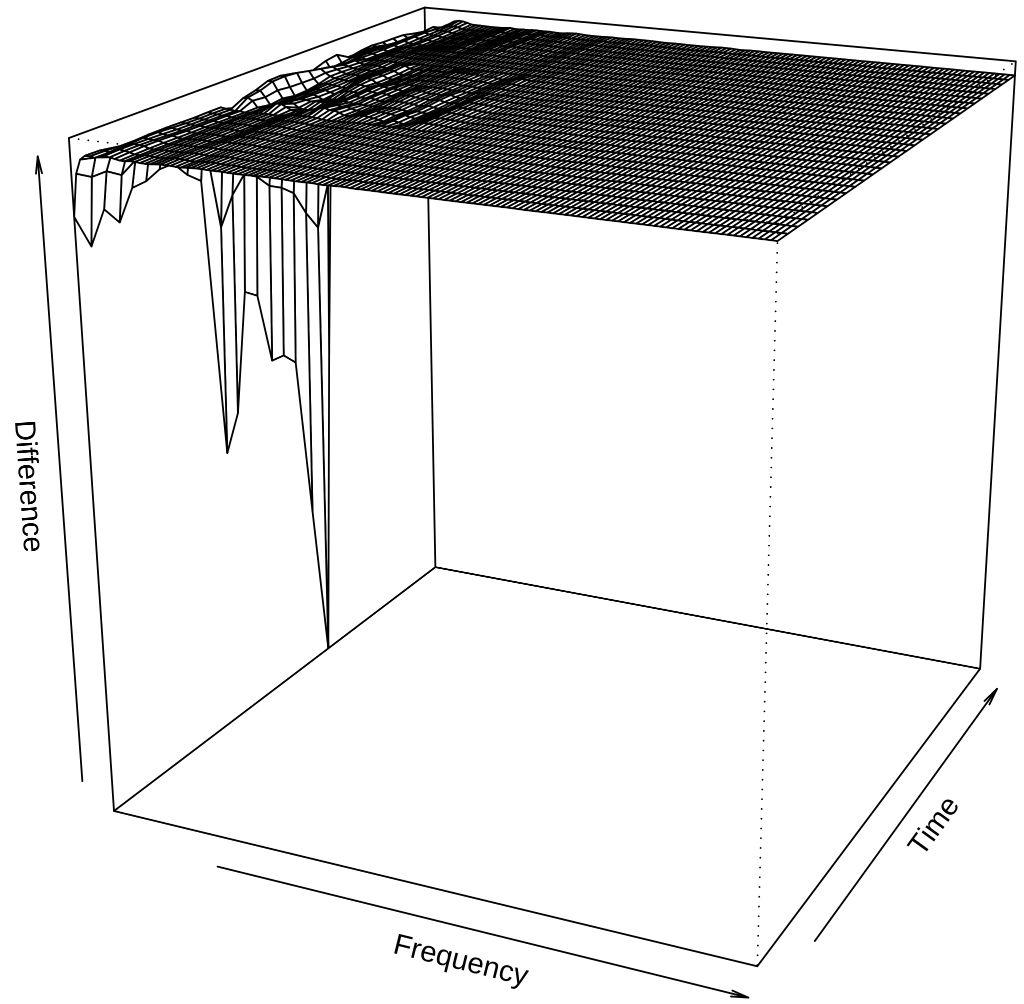
Maximum: 37.979 *10^3 [ms^2]

Robust Short−term Analysis of HRV with AO



Maximum: 224.82 *10^3 [ms^2]

Classical Short−term Analysis without Outliers vs. Robust Short−term Analysis with AO



Minimum: −212.139 *10^3 [ms^2], Maximum: 4.6 *10^3 [ms^2]

Estimation of Power Spectra via Robust Filtering

B. Spangl and R. Dutter

BOKU - University of Natural Resources and Applied Life Sciences

and

Vienna University of Technology

(joint work with J. Pumprla and K. Howorka,
General Hospital of Vienna)

Österreichische Statistiktage 2003, Oct. 2003