Robust Estimation of Power Spectra via the Autocovariance Function

B. Spangl und R. Dutter

University of Natural Resources and Applied Life Sciences

and

Vienna University of Technology

RMED'03, July 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.

Autocovariance Function

The classical estimator for the autocovariance function, based on the method of moments, on a sample $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$, is

$$\widehat{\gamma}_M(h, \boldsymbol{x}) = \frac{1}{n-h} \sum_{i=1}^{n-h} (x_{i+h} - \bar{x})(x_i - \bar{x}) \quad 0 \leq h \leq n-1 \tag{1}$$

where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \ .$$

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

$$x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q} \ , \qquad (2)$$

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

Innovational Outliers (IO)

The ARMA(p, q) process $\{x_t : t \in \mathbb{Z}\}$ is said to have innovational outliers (IO) if the innovations w_t have a heavy-tailed distribution, for instance an ε-contaminated normal distribution

$$CN(\varepsilon, \sigma_1, \sigma_2) = (1 - \varepsilon)N(0, \sigma_1^2) + \varepsilon N(0, \sigma_2^2) \tag{3}$$

where $N(0, \sigma^2)$ denotes the normal distribution with mean zero and variance σ^2 and $\sigma_1^2 \ll \sigma_2^2$ and ε is small.

Additive Outliers (AO)

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

$$x_t = v_t + w_t \tag{4}$$

where v_t is an ARMA(p, q) process and the w_t are independent identically distributed (iid) with common distribution $F_w = (1 - \varepsilon)\delta_0 + \varepsilon H$ 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 v_t is observed with probability $1 - \varepsilon$ whereas the ARMA(p, q) process plus an error w_t is observed with probability ε. We shall also assume that v_t and w_t are independent.

Spectral Representation Theorem

Any second-order stationary process x_t has the spectral representation

$$x_t = \int_{-1/2}^{1/2} \exp(-2\pi i t f)\, dz(f) \tag{5}$$

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

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty \tag{6}$$

has the representation

$$\gamma(h) = \int_{-1/2}^{1/2} \exp(2\pi i h f) S(f)\, df, \quad h \in \mathbb{Z} \ . \tag{7}$$

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

$$S(f) = \sum_{h=-\infty}^{\infty} \gamma(h) \exp(-2\pi i h f) \ . \tag{8}$$

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

$$X(f_k) = n^{-1/2} \sum_{t=1}^{n} x_t \exp(-\frac{2\pi i t k}{n}) \ , \quad f_k = \frac{k}{n} \ , \quad k = -\left[\frac{n-1}{2}\right], \ldots, \left[\frac{n}{2}\right] \tag{9}$$

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

$$I(f_k) = |X(f_k)|^2 \ . \tag{10}$$

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

$$\widehat{S}(f_k) = \sum_{j=-L}^{L} w_j I(f_{k+j}) \ , \quad w_j = w_{-j} \ , \quad \sum_{j=-L}^{L} w_j = 1 \ . \tag{11}$$

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

$$x_t - \sum_{k=1}^{p} \phi_k x_{t-k} = z_t \ , \tag{12}$$

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

$$\widehat{S}(f) = \frac{\widehat{\sigma}_z^2}{\left|1 - \sum_{k=1}^{p} \widehat{\phi}_k \exp(-2\pi i k f)\right|^2} \ . \tag{13}$$

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):

$$\widehat{S}(f) = \frac{\widehat{S}_r(f)}{|\widehat{H}_p(f)|^2} \ , \tag{14}$$

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

$$\widehat{H}_p(f) = 1 - \sum_{j=1}^{p} \widehat{\phi}_j \, \text{exp}(i2\pi j f) \ . \tag{15}$$

The basic proposal is to robustify both the numerator $\widehat{S}_r(f)$ and the denominator $\widehat{H}_p(f)$ in (14).

Estimation Procedure Based on Robust Filtering

(i) We start by fitting (by least squares) the autoregression model

$$y_k = \sum_{j=1}^{p} \widehat{\phi}_j y_{k-j} + e_k \tag{16}$$

to the data $\{y_k, k = 1, \ldots, n\}$ obtaining the estimate $\widehat{\boldsymbol{\phi}} = (\widehat{\phi}_1, \ldots, \widehat{\phi}_p)^\top$.

(ii) The robust filtering algorithm is

$$\widehat{x}_k = \widehat{\boldsymbol{x}}_{k-1}^\top \widehat{\boldsymbol{\phi}} + c\widehat{s}\psi\left(\frac{y_k - \widehat{\boldsymbol{x}}_{k-1}^\top \widehat{\boldsymbol{\phi}}}{c\widehat{s}}\right), \tag{17}$$

where $\widehat{\boldsymbol{x}}_{k-1} = (\widehat{x}_{k-1}, \ldots, \widehat{x}_{k-p})^\top$ and $\widehat{s}^2$ is an estimate of the innovations variance of the underlying process x_k.

(iii) Robustly prewhitened residuals $\{r_k, k = p+1, \ldots, n\}$ are obtained by

$$r_k = \widehat{x}_k - \widehat{\boldsymbol{x}}_{k-1}^\top \widehat{\boldsymbol{\phi}} = c\widehat{s}\psi\left(\frac{y_k - \widehat{\boldsymbol{x}}_{k-1}^\top \widehat{\boldsymbol{\phi}}}{c\widehat{s}}\right), \tag{18}$$

and these residuals are used to compute a new robust estimate $\widehat{s}$.

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:

$$\widehat{\gamma}_Q(h,\boldsymbol{x}) = \frac{1}{4}[Q^2_{n-h}(\boldsymbol{u}+\boldsymbol{v}) - Q^2_{n-h}(\boldsymbol{u}-\boldsymbol{v})] \ . \tag{19}$$

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

$$Q_n(\boldsymbol{z}) = c\{|z_i - z_j| : i < j\}_{(k)} \ , \tag{20}$$

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.

Estimates via a Robust Autocovariance Function

We note that the classical estimator of the autocovariance function (1) is not modified to ensure non-negative definiteness of the autocovariance matrix.

However, we should ensure non-negative definiteness of the autocovariance matrix obtained by the highly robust autocovariance estimator (19). This can be done by shrinking, the eigenvalue or the scaling method (cf. Rousseeuw & Molenberghs, 1993).

Now we use the highly robust autocovariance estimator in the Yule-Walker equations to estimate the parameters of an AR(p) model robustly.

These robustly estimated parameters, $\widehat{\phi}$ and $\widehat{s}$, are directly used in the robust filtering algorithm (17) to compute a cleaned process $\widehat{x}_k$ as well as robustly prewhitened residuals r_k.

Again $\widehat{S}_r(f)$ is computed from the $\{r_k, k = p+1, \ldots, n\}$ and $\widehat{\phi}$ is inserted in (15) to get the robust spectrum estimate (14).

Examples

To compute the power spectrum we iterate the robust filtering algorithm four times.

The influence function used in (17) has the redescending form

$$\psi(x) = x \,\mathsf{exp}(-\,\mathsf{exp}(q(|x| - q))) \ , \tag{21}$$

with a value of $q = 3.1$.

Simulated Data

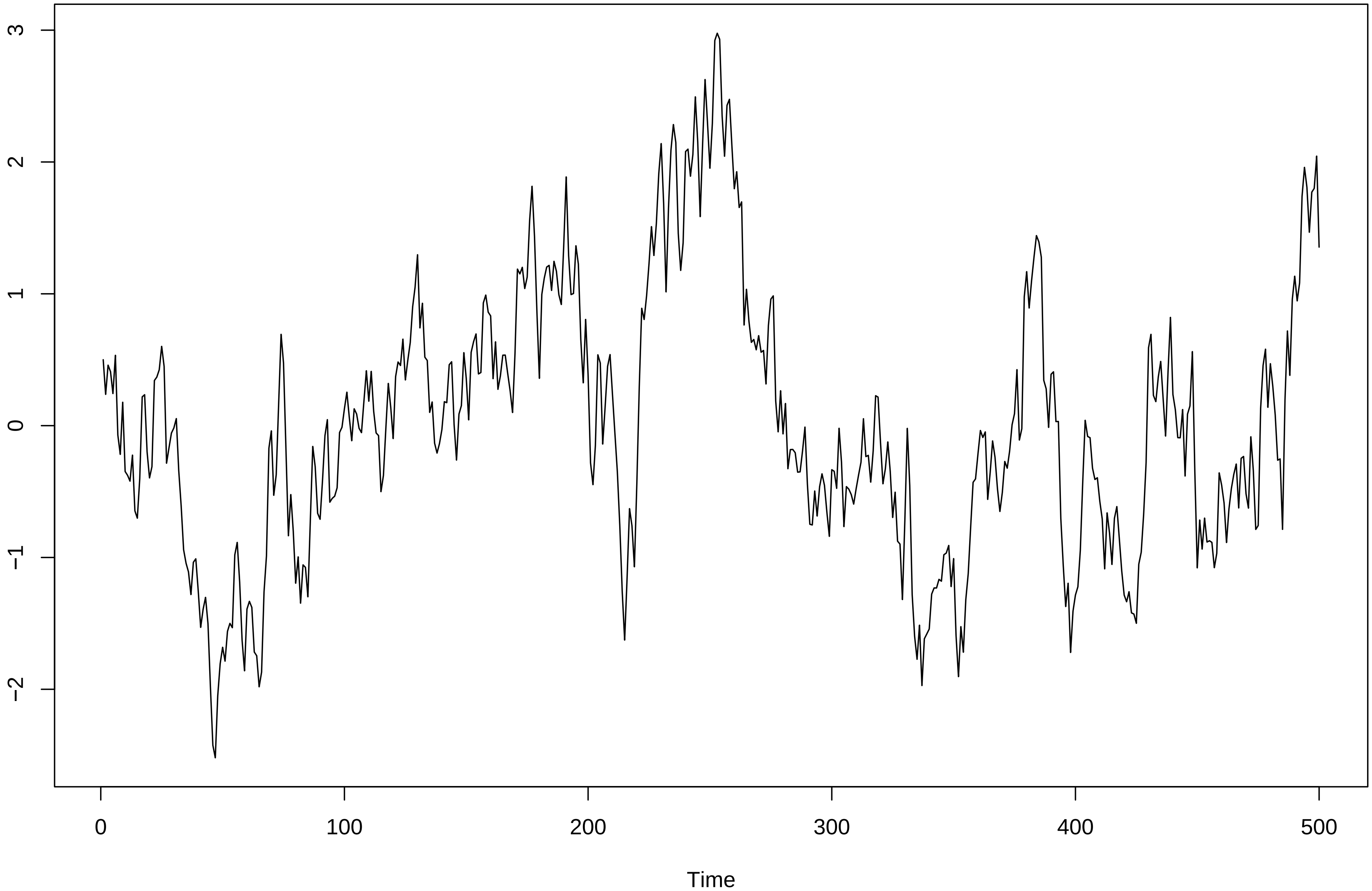
This example is a simulated process with AO composed of the following three autoregressive processes:

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

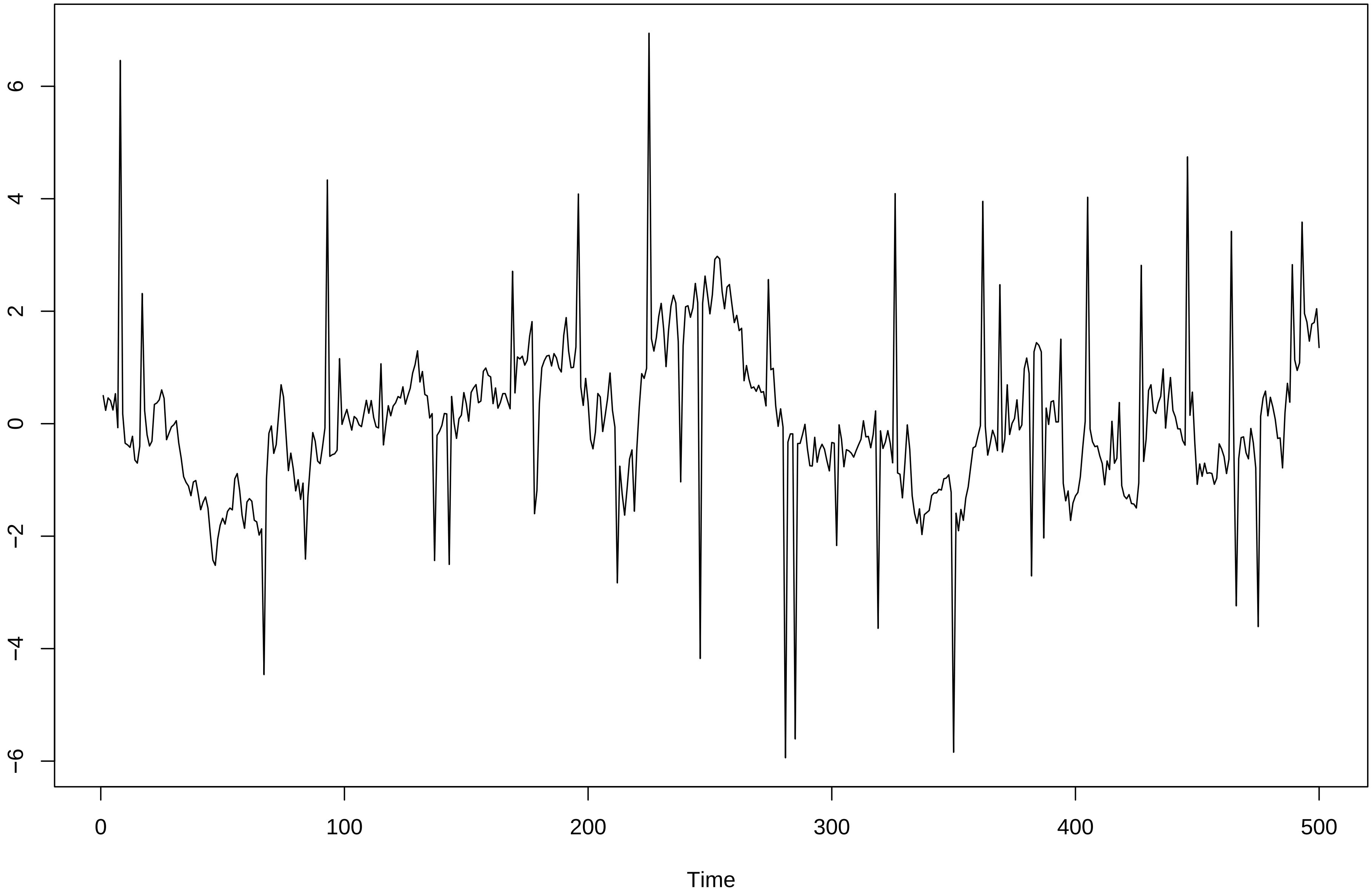
with $\varepsilon_k, \eta_k, \zeta_k \sim N(0,1)$. We standardize u_k, w_k and z_k, and compute

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

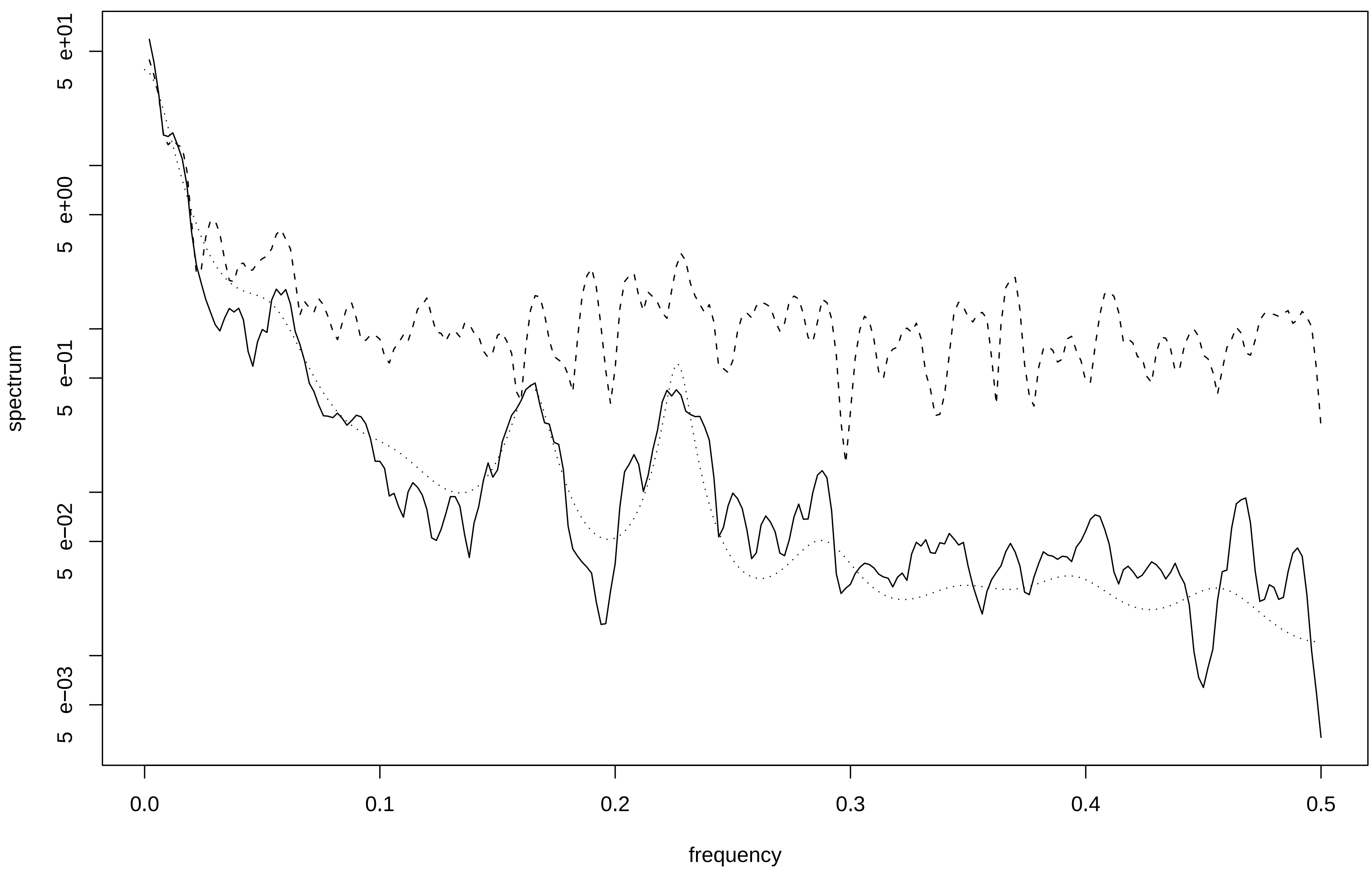
and then standardize y_k again. Additionally we add some noise from $0.9\delta_0 + 0.1 N(0,10)$.



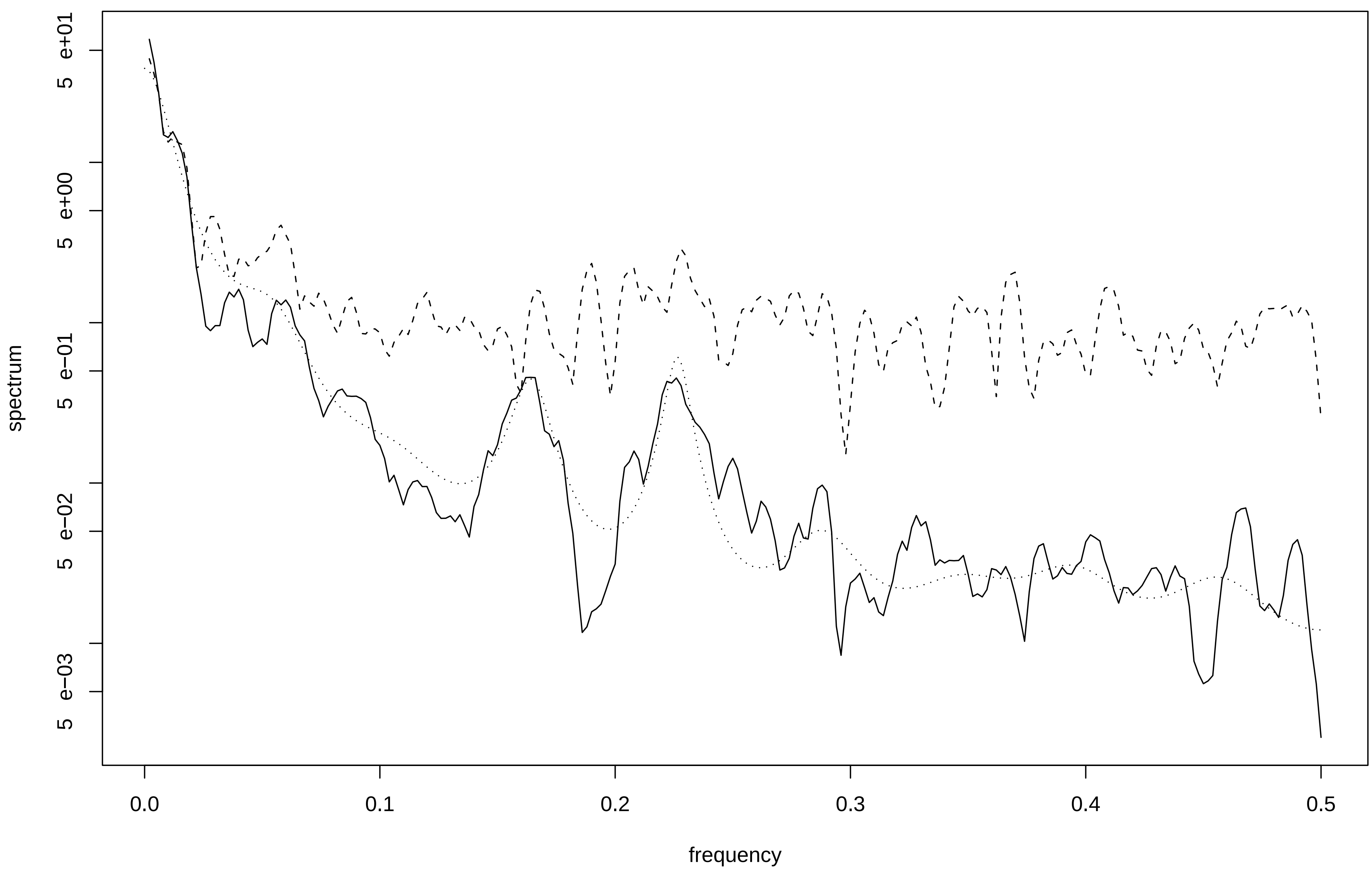
3
2
1
0
−1
−2
0
100
200
300
400
500
Time



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



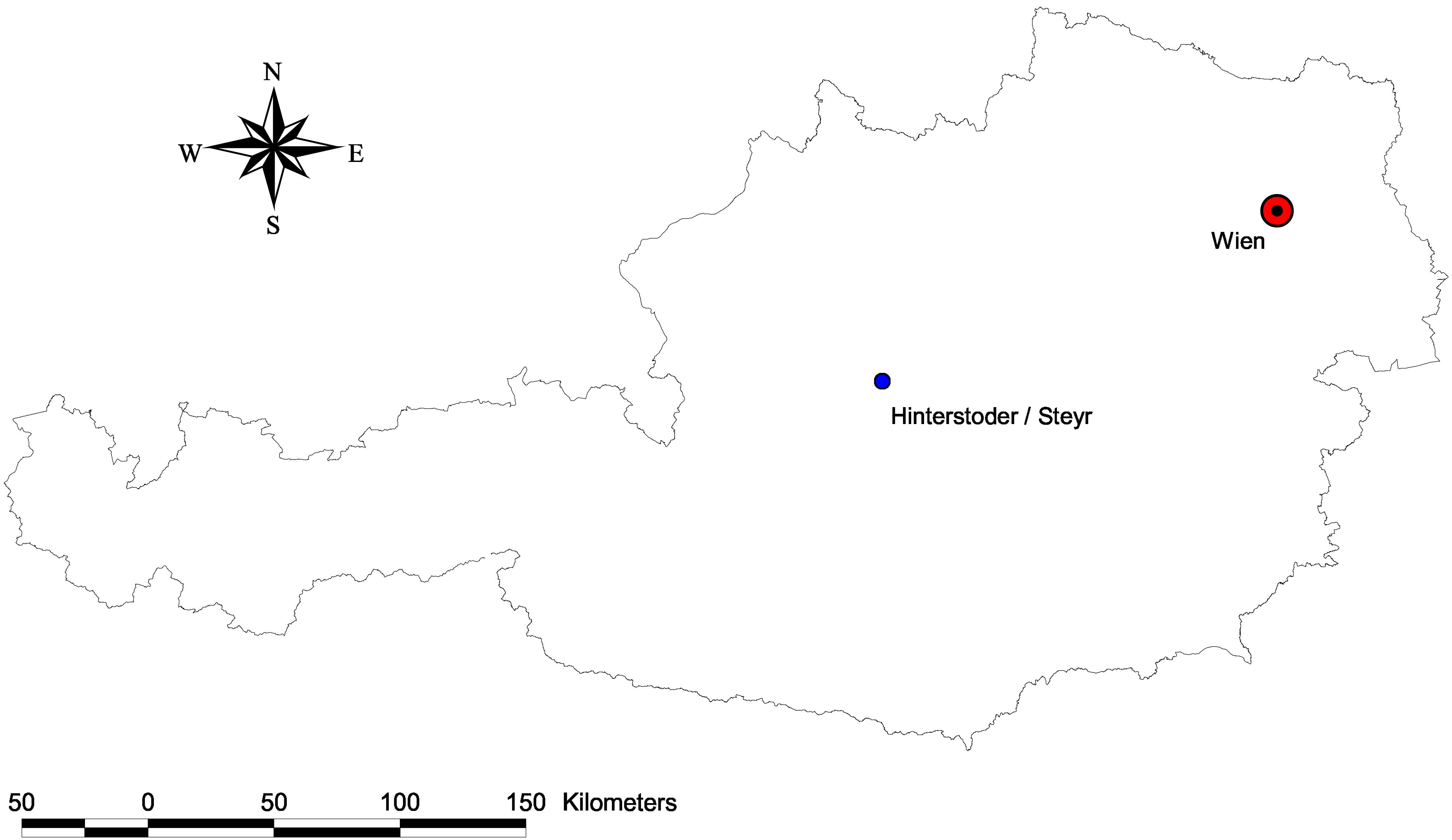
spectrum
frequency
0.0
0.1
0.2
0.3
0.4
0.5
5 e−03
e−03
5 e−02
e−02
5 e−01
e−01
5 e+00
e+00
5 e+01
e+01

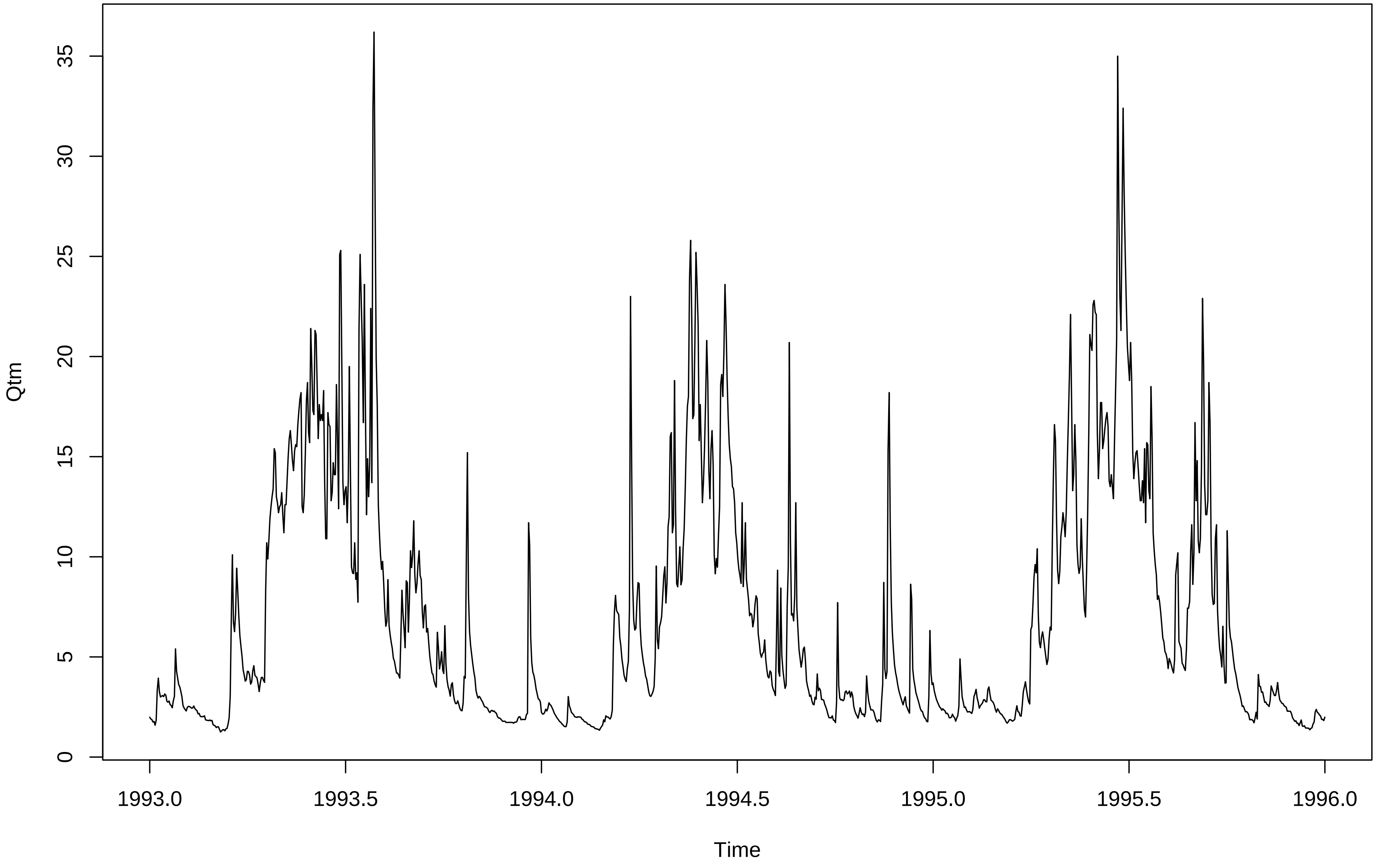


spectrum
e+01
5
e+00
5
e−01
5
e−02
5
e−03
5
0.0
0.1
0.2
0.3
0.4
0.5
frequency

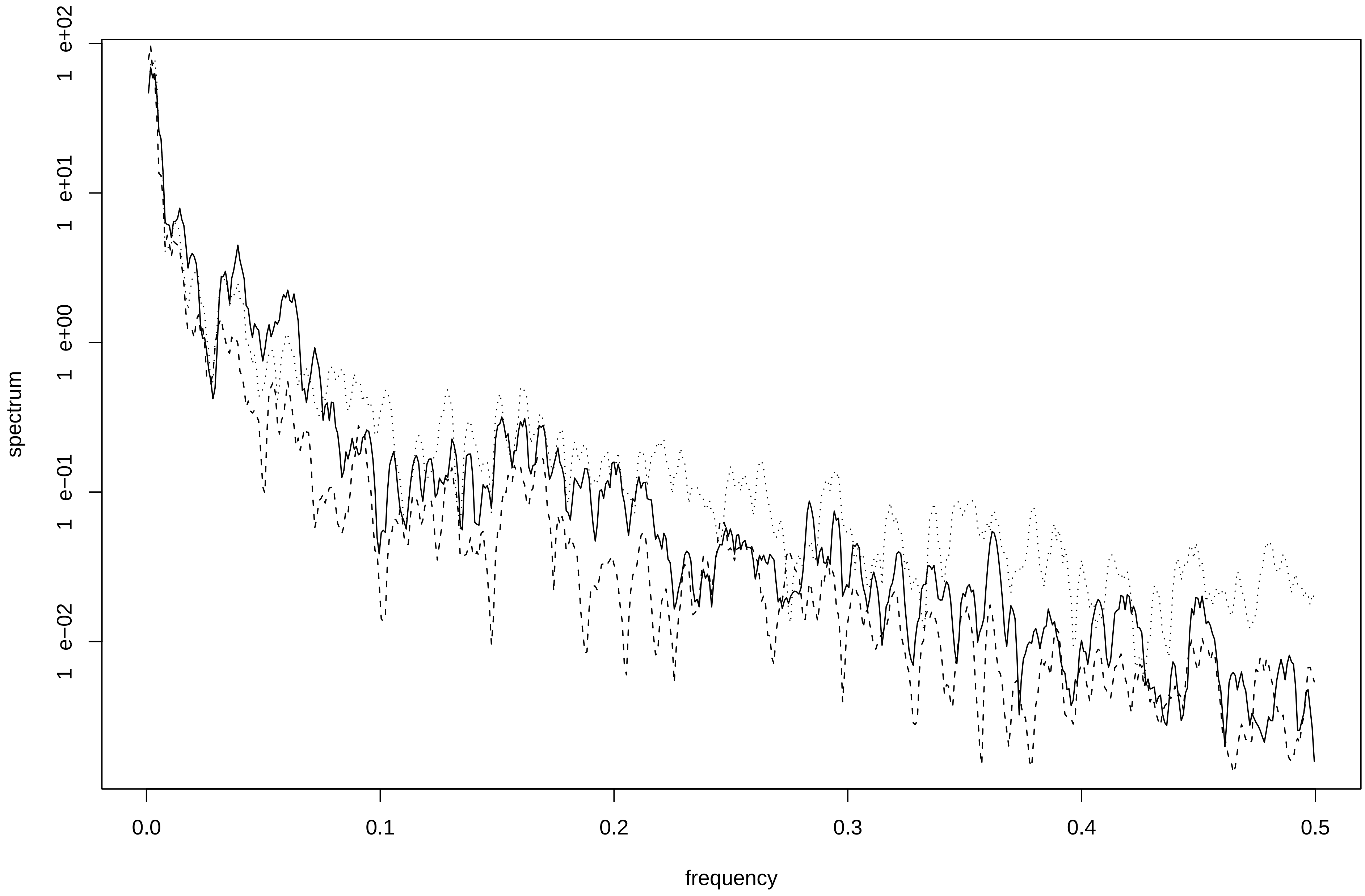
River Data

- river: Steyr
- gauging station: Hinterstoder (Upper Austria)
- observations: daily discharges ($[m^3/s]$) of three years (1993-1995)


N
W
E
S
Wien
Hinterstoder / Steyr
50
0
50
100
150
Kilometers



Qtm
0
5
10
15
20
25
30
35
1993.0
1993.5
1994.0
1994.5
1995.0
1995.5
1996.0
Time



1 e+02
1 e+01
1 e+00
1 e−01
1 e−02
spectrum
0.0
0.1
0.2
0.3
0.4
0.5
frequency