A test for a difference between spectral peak frequencies

J. Timmer [1,2], M. Lauk [1,2,3], W. Vach [1,4] and C.H. Lücking [5]

[1] Zentrum für Datenanalyse und Modellbildung, Eckerstr. 1, 79104 Freiburg, Germany
[2] Fakultät für Physik, Hermann-Herder-Str.3, 79104 Freiburg, Germany
[3] Center for BioDynamics, Boston University, 44 Cummington St, Boston, MA 02215
[4] Department of Statistics and Demography, Odense University, Hollufgaard, Hestehaven 201, 5220 Odense SO, Denmark
[5] Neurologische Universitätsklinik Freiburg Breisacher Str. 64, 79110 Freiburg, Germany

(submitted for publication 3/5/98)

A procedure which allows to test the hypothesis that the spectral peak frequencies of two stochastic processes are equal is proposed. The method is based on drawing new realizations of the periodograms from the two estimated spectra in order to reestimate the spectra and to obtain the distribution of the peak frequency difference. We investigate the size and the power of the test in simulation studies and apply the new method to human tremor data.

1 Introduction

Spectral analysis is frequently applied to investigate oscillatory phenomena in basic medical and biological research (Marsden et al., 1969a; Marsden et al., 1969b; Nelson et al., 1979; Günther, Brunner and Klußmann, 1983; Cleeves and Findley, 1987; Hefter et al., 1987; Clayton et al. 1995; Beuter, 1996; Pinna, Maestri, Di Cesare, 1996; Hutchinson et al., 1997) and medical treatment monitoring (Herrmann, 1982; Khutorskaya and Fjodorova, 1996). Usually, the peak frequency and the power of the oscillation are of major interest. Often, it is desirable to decide whether differences in the observed spectral peak frequencies of two estimated spectra are significant, i.e. to test the hypothesis of a zero peak frequency difference.

We propose a procedure to test this hypothesis. It is inspired by the parametric bootstrap (Efron and Tibshirani, 1993) and the theory of the spectral estimation (Brockwell and Davis, 1987). General aspects of bootstrapping in the frequency domain are discussed in Franke and Härdle (1992), Janas and Dahlhaus (1994) and Dahlhaus and Janas (1995). A resampling procedure to obtain a confidence region for the spectral peak frequency is given in Timmer et al. (1997).

In the next section we briefly summarize parts of the theory of spectral estimation and describe the estimation procedure used in this study. In Section 3 we introduce the test procedure. In Section 4 simulation studies are reported which investigate the size and the power of the test. Finally, the procedure is applied to tremor data.

2 Spectral estimation

The spectrum $\mathrm{S}(f)$ of a stochastic process $X(t)_{t\in\mathbf{Z}}$ is defined as the Fourier transform of the autocovariance function $\mathrm{ACF}(\tau)$:

$$\mathrm{ACF}(\tau) \;=\; E(X(t)\,X(t+\tau)) \tag{1}$$

$$\mathrm{S}(f) \;=\; \frac{1}{2\pi}\sum_{\tau}\mathrm{ACF}(\tau)\exp\left(2\pi\, i\, f\, \tau\right), \tag{2}$$

where "$E(.)$" denotes expectations (Brockwell and Davis, 1987). One procedure to estimate the spectrum is based on the Fourier transform $\mathrm{d}(f)$ and the periodogram $\mathrm{Per}(f)$ of the data $x(t)_{t=1,\ldots,N}$ (which are usually tapered to reduce spectral leakage):

$$\mathrm{d}(f) \;=\; \frac{1}{\sqrt{2\pi\, N}}\sum_{t=1}^{N} x(t)\,\exp\left(2\pi\, i\, f\, t\right) \tag{3}$$

$$\mathrm{Per}(f) \;=\; |\,\mathrm{d}(f)\,|^2 \tag{4}$$

and is evaluated for the frequency bins:

$$f_k = \frac{k}{N}, \quad k = -\frac{N-1}{2},\ldots,\frac{N}{2} \quad . \tag{5}$$

For two different frequencies f and $\tilde{f}$, $\mathrm{Per}(f)$ and $\mathrm{Per}(\tilde{f})$ are asymptotically independent. For $f_k \notin \{0, 0.5\}$ the periodogram is asymptotically distributed as

$$\mathrm{Per}(f_k) \sim \frac{1}{2}\mathrm{S}(f_k)\,\chi^2_2 \quad . \tag{6}$$

Thus, the expectation of the periodogram is the spectrum but the periodogram is not a consistent estimator for the spectrum. To estimate the spectrum consistently, the periodogram is smoothed by a window functionW_j:

$$\widehat{\mathrm{S}}(f_k) = \sum_{j=-h}^{h} W_j\,\mathrm{Per}(f_{k+j}) \quad , \quad \sum_{j=-h}^{h} W_j = 1 \quad . \tag{7}$$

Other methods to estimate the spectrum as well as general aspects of spectral estimation concerning the choice of the window function W_j and its width $2h+1$ are given in Brockwell and Davis (1987).

The optimal width $2h+1$ of the smoothing window depends on the curvature of the true spectrum. For large curvature the width should be smaller than for a small a frequency dependent smoothing window can improve the spectral estimation. Such a frequency dependent smoothing procedure has been proposed in Timmer et al. (1996). This data driven algorithm was developed for an optimal estimation of the spectrum in the region of the main peak, i.e. the global maximum of the spectrum. It proceeds in five steps:

1. A preliminary spectrum is estimated by uniform smoothing ($h = h_0$). h_0 has to be chosen depending on sampling rate and the number of data.

2. The width w of the peak is determined by difference of the two frequencies left (f_l) and right (f_r) to the peak frequency (f_p) where the power decreased to half the value at the peak.

3. A larger width of the peak corresponds to a smaller curvature of the spectrum at the peak, calling for a larger bandwidth of the smoothing window. Therefore, the smoothing width at the peak is determined by:

$$h(f_p) = \frac{(f_l - f_r)^2}{b} \quad , \tag{8}$$

   where b is a parameter to be chosen depending on the sampling rate.

4. Moving away from the peak, the curvature decreases. Thus the width of the smoothing window should increase. We chose a linear increase determined by:

$$s_l = a\,\frac{f_l - f_0}{2h_0} \tag{9}$$

$$s_r = a\,\frac{f_r - f_0}{2h_0} \tag{10}$$

$$h(f_k) = \begin{cases} \min(h(f_0) + s_l(f_k - f_0), h_{max}) \,, & f_k \leq f_0 \\ \min(h(f_0) + s_r(f_k - f_0), h_{max}) \,, & f_k \geq f_0 \end{cases} \tag{11}$$

   a is a parameter to be chosen depending on the sampling rate. h_{max} represents a upper limit for the smoothing width to prevent oversmoothing.

5. The spectrum is estimated by smoothing with the frequency dependent kernel width. As smoothing kernel W_i we chose a triangular one:

$$W_i = \frac{1}{h(f_k) + 1} - \frac{1}{(h(f_k) + 1)^2}|i| \quad . \tag{12}$$

A more detailed description of the estimation procedure and applications to measured data are given in Timmer et al. (1996).

If two different processes are given, it is often of interest to assess the difference between the peak frequencies of the spectra of the two processes. The difference of the estimated peak frequencies $\widehat{f_1^p}$ and $\widehat{f_2^p}$ provides a natural estimate. However, further statistical inference is difficult, as the distribution of this estimate cannot be calculated analytically.

Therefore, we propose the following resampling procedure.

(1) Estimate the spectra of the two processes from the data.

(2) Estimate the difference Δf of the peak frequencies f_1^p and f_2^p of the two processes by $\widehat{\Delta f} := \widehat{f_1^p} - \widehat{f_2^p}$.

(3) Simulate numerous periodograms from the estimated spectra according to eq. (6).

(4) For each pair of simulated periodograms reestimate the spectra according to eq. (7-12) and the difference $\widehat{\Delta f}^*$ of the peak frequencies.

(5) Obtain the distribution of $(\widehat{\Delta f}^* - \widehat{\Delta f})$ and reject the null hypothesis if $\widehat{\Delta f}$ is smaller than the lower $\alpha/2$ quantile or larger than the upper $\alpha/2$ quantile.

By (5) we take into account the first guideline for bootstrapping hypothesis tests by Hall and Wilson (1991). This version of the test is called variant 1 in the following. The second guideline of Hall and Wilson suggests to divide the test statistic by an estimate of the scale of the test statistic, i.e. to perform some kind of pivoting. Now the variance of the peak frequency estimator is inverse proportional to the curvature at the peak and the curvature is inverse proportional to the square of the width of the peak assuming that a parabolic approximation is valid. As we can hope to estimate the widths of the peaks rather reliable by the empirical widths $w_{i,\, i=1,2}$ we have a good candidate for pivoting. The width w is again defined as the difference of the frequencies of half power right and left to the peak.

Hence in variant 2 we consider the bootstrap distribution:

$$\frac{\widehat{\Delta f}^* - \widehat{\Delta f}}{\sqrt{w_1^{*2} + w_2^{*2}}} \tag{13}$$

and reject the null hypothesis if $\widehat{\Delta f}/\sqrt{w_1^2 + w_2^2}$ is smaller than the lower $\alpha/2$ quantile or larger than the upper $\alpha/2$ quantile.

In bootrapping the periodograms based on eq. (6), we assume independence of the periodogram among all frequency bins. Asymptotically, for two fixed frequencies the values of the periodogram are indeed independent.

First, however, two adjacent values of the periodogram might show correlations independent of the and the tapering window applied (Brockwell and Davis, 1987). The proposed procedure does not take these correlations into account but this can be justified by the fact that they decay rapidly within some frequency bins and the smoothing usually involves more frequency bins than that correlation length.

Second, in the case of nonlinear or non-Gaussian processes, there might be higher order correlations, e.g. triple correlations described by the bispectrum (Nikias and Mysore, 1987). For oscillatory, i.e. non chaotic, processes these higher order correlations usually appear for harmonically related peaks. Since the proposed procedure only considers the distribution of the periodogram in the neighbourhood of the global maximum, it can be justified that these correlations are not taken into account.

4 Simulation Studies

For the simulation studies we choose linear autoregressive (AR) processes since for these processes the parameters can be calculated analytically for a desired spectral peak frequency. The parameters chosen for the simulation are inspired by observed properties of spectral peaks of time series from human hand tremor, for examples see Timmer et al. (1996).

4.1 Design of the simulation study

An AR process of order p is given by :

$$x(t) = \sum_{i=1}^{p} a_i \, x(t-i) \, + \, \epsilon(t) \quad , \tag{14}$$

where $\epsilon(t)$ denotes i.i.d. Gaussian random variables with zero mean and variance σ^2. Such processes can be interpreted in terms of physics depending on the chosen parameters as a combination of relaxators and damped oscillators (Honerkamp, 1993). For example, in the case of an AR process of order 2 (AR[2]) with appropriate chosen parameters a_1 and a_2 a damped oscillator is given. The characteristic period T, respectively the frequency $f = 1/T$, and relaxation time τ are related to the parameters by:

$$a_1 = 2\cos\left(\frac{2\pi}{T}\right) \exp\left(-1/\tau\right) \tag{15}$$

$$a_2 = -\exp\left(-2/\tau\right) \quad . \tag{16}$$

The spectrum of an AR[2] process is given by:

$$\mathrm{S}(f) = \frac{1}{2\pi} \frac{\sigma^2}{|1 - a_1 e^{-2\pi i f} - a_2 e^{-4\pi i f}|^2} \tag{17}$$

and the peak frequency is located at:

$$f_{peak} = \arccos(\cos(2\pi/T)\,\cosh(1/\tau)) \quad . \tag{18}$$

To examine the behavior of the test in dependence on the width of the peak we choose two AR[2] processes and investigated the three possible combinations of these processes. With respect to the null hypothesis we consider process 1 associated with a broad spectral peak (characteristic times $T_1 = 50$ and $\tau_1 = 100$, corresponding to $a_1 = 1.964486$ and $a_2 = -0.980199$) and process 2 showing a sharp spectral peak at the same frequency (characteristic times $T_1 = 50.15$ and $\tau_1 = 500$, corresponding to $a_1 = 1.980359$ and $a_2 = -0.996008$).

To give a more vivid impression of the results and to faciliate the comparison with the applications in Section 5 we regard the data as sampled with 300 Hz again inspired by common setting in tremor recordings. By this transformation the peaks are located at approximately 6 Hz. Fig. 1 displays the spectra of the two processes according to (17) for the frequency range 0-30 Hz and their widths of half power (0.98 Hz for process 1, 0.19 Hz for process 2). To investigate the power of the test, we consider alternatives where the peak frequency of one process is shifted to higher frequencies keeping its width constant.

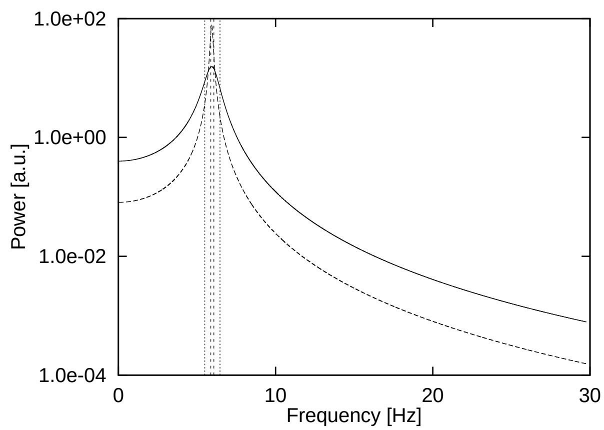

Figure 1: Spectra of the two AR[2] processes used in the simulation study (solid line: process 1; dashed line: process 2). The widths of half power are marked with vertical dotted and dashed lines, respectively.

We investigate the behavior of the test for time series of length 2000, 10000, and 50000 data points. 10000 data points is a typical amount of data in tremor measurements as well as in other electrophysiological recordings like EEG, MEG and EMG. The frequency binning is 0.15 Hz for the time series of length 2000, 0.03 Hz for those of length 10000 and 0.006 Hz for time series consisting of 50000 data points.

We use the pseudo-random-number generator GASDEV and RAN2 from Press et al. (1992) to generate the driving noise $\epsilon(t)$ of the AR processes and to resample the periodograms. All programs are written in C, the simulations were done on unix workstations. The parameters specifying the spectral estimation in (8-11)

For each choice of the parameters we apply 500 independent repetitions of the simulation study. In each repetition of the simulation we use 500 bootstrap samples. In rejecting the null hypothesis we use the 25th and the 476th of the ordered values as quantiles, corresponding to a nominal significance level of 10 %. As the spectrum is evaluated at the frequency bins all test statistics are discrete and it arises the question of appropriate boundaries in presence of ties. In variant 2 due to pivoting (eq.(13)) the problem is negligible. In variant 1 we consider the anticonservative version where both quantiles are within the region of rejection, but as we will show even for this version variant 2 has asymptotically more power.

4.2 Results of the simulation studies

Fig. 2 shows the results of the simulation study for the three combinations of the processes, the different lengths of the time series and the two variants of the test. In any case we observe an extreme conservative behavior: The actual niveau of the test is close to 0, although the nominal niveau of 10%. In spite of this extreme conservativism the test shows a large power. For example, see Fig. 2a, we can expect a power of 0.9 for time series of length 10000 of process 1, if the difference of the peak frequencies is 0.5 Hz, corresponding to approximately half the width of the peak. For the time series of length 10000 and 50000 variant 2 is always more powerful than variant 1, in spite of the anticonservative choice of the critical region for variant 1. As expected the power increases with the length of the time series. Comparing Fig. 2a-c confirms the expectation that a violation of the null hypothesis is the faster detected the sharper the two peaks under investigation are.

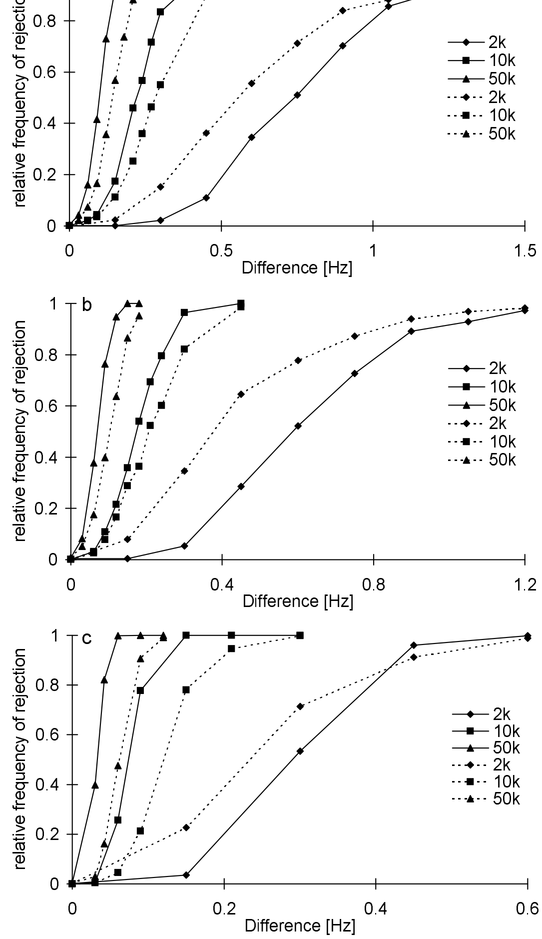

Figure 2: Power of the test in dependence of the peak frequency difference. Test variant 1 (solid lines), variant 2 (dashed lines). $N = 2000$ (♦), $N = 10000$ (■), $N = 50000$ (▲). a: Both processes are process 1. b: Combination of processes 1 and 2. c: Both processes are process 2.

5 Application

We apply the proposed test procedure to three pairs of human hand tremor time series recorded under different conditions from patients suffering from Parkinson's disease. All time series are of length 30000 data points and sampled with 300 Hz. We chose a significance level of 1 % which is determined by 5000 resampled periodograms.

The first example investigates the effect of an unilateral high frequency electrical stimulation in the ventral intermediate thalamic nucleus (Benabid et al.,1991) on the frequency of the postural tremor. Fig. 3a shows the estimated spectra with stimulation (solid line) and without stimulation (dashed line). The spectra were estimated from time series of the electromyogram, i.e. the muscle activity measured by surface electrodes, obtained from the left hand's extensor. The spectrum of the time series without stimulation exhibits higher harmonics typically for Parkinsonian tremor. The peak frequency is located at 4.9 Hz. Obviously, the stimulation is successful in reducing the power of the oscillation by a factor of ten. Under stimulation the peak frequency is located at 6.27 Hz. The difference between the peak frequencies is 1.37 Hz, the pivotized difference is 1.99. Variant 1 of the test yields a 0.5 % quantile of the peak frequency difference distribution assuming the null hypothesis of equal spectral peak frequencies at -0.30 Hz and a 99.5% quantile at 0.33 Hz, variant 2 gives respective values of -0.31 and 0.34. Thus, the hypothesis of equal frequencies has to be rejected at the 1% level.

The second example examines the effect of a medication by Botulinum Toxin (Botox). This drug is usually applied to treat dystonia. In this case the patient suffered from a hand dystonia and showed, additionally, a Parkinsonian hand tremor. The Botox was applied approximately 20 days prior to the second recording (musculus flexor ulnaris 15 units, musculus flexor digitorum profundus 10 units, musculus flexor pollicis longus 10

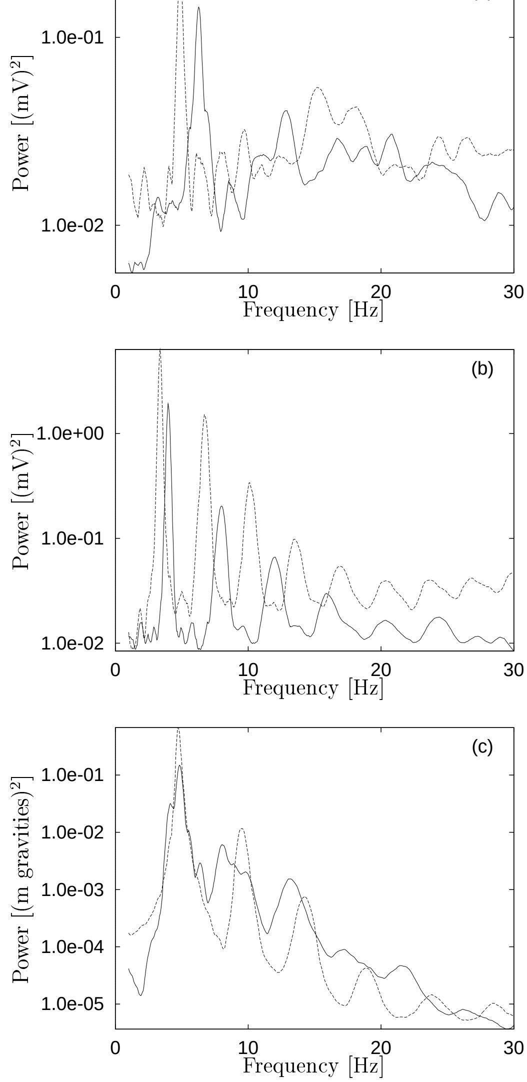

Figure 3: Estimated spectra of EMG and hand tremor time series from patients suffering from Parkinson's disease. a: Extensor EMG with (solid line) and without (dashed line) electrical stimulation of the thalamus. b: Extensor EMG with (solid line) and without (dashed line) treatment with Botulinum Toxin. c: Acceleration of the hand with medication by Propanolol (solid line) and by Metixen (dashed line).

Units). The estimated spectra, again based on the electromyogram, in Fig. 3b reveal no large effect either on the variance nor on the harmonic structure, as expected since Botox have not been applied to treat the tremor. Before treatment the peak frequency is located at 3.97 Hz. After treatment the peak occurs 0.6 Hz lower at 3.37 Hz. The pivotized difference amounts to 1.69. Variant 1 of the test gives a critical region for the null hypothesis of equal peak frequencies outside -0.17/0.20 Hz, variant 2 yields -0.35/0.44. Thus, while not affecting the amplitude and the nonlinear type of oscillation reflected by

gests that, in this case, Botox had a significant effect on the frequency of the Parkinsonian tremor.

In the third example the effect of two different drugs (Propanolol and Metixen) on the peak frequency of the tremor measured by the acceleration of the hand is investigated. The peak frequencies are located at 4.83, respectively 4.73 Hz, resulting in a difference of 0.1 Hz and 0.21 for the pivotized version. The critical values for the hypothesis of a zero spectral peak frequency difference of test variant 1 are -0.23/0.20 Hz, those of variant 2 are -0.38/0.35. Thus the null hypothesis can not be rejected at the 1 % level.

For all three examples the ratio of the test statistics and the critical values are larger for the pivotized variant 2 of the test. For the firtst two examples, variant 2 allows rejection at a lower significance level than variant 1. This confirms the results of the simulation study with respect to the larger power of variant 2 for this number of data points.

6 Conclusion

Testing for a difference between spectral peak frequencies allows to judge the variability of oscillatory phenomena, e.g. effects of treatment like medication, stereotactic surgery or electrical stimulation on Parkinsonian tremor or the significance of temporal, biological fluctuations of the tremor peak frequency.

The simulation studies revealed that the proposed test procedure is extremely conservative. This is a consequence of the smoothing of the periodogram to estimate the spectrum. Since the smoothing kernel is positive, the estimated spectra on which the resampling is based exhibit a width which is too large corresponding to an underestimation of the curvature at the peak. The conservativism is contrasted by the good power of the procedure. The simulation studies to investigate the power of the test were designed to match conditions which are met in typical clinical data recordings. A relative difference in peak frequencies in the order of 10 % can be detected reliably by the test.

We are confident that the proposed procedure will help to judge results of spectral analysis in a more quantitative way than possible up to now.

Acknowledgements.
M.L. acknowledges financial support by the German Federal Ministry of Education, Science, Research and Technology by the grant "Mathematik: Schlüsseltechnologie der Zukunft".

Address correspondence to: J. Timmer, Zentrum für Datenanalyse und Modellbildung, Eckerstr. 1, 79104 Freiburg, Germany.
Phone: +49/761/203-7710, FAX: +49/761/203-7700, e-mail: jeti@fdm.uni-freiburg.de

Benabid, A.L., Pollack, P., Gervason, C., Hoffmann, B., Gao, D.M., Hommel, M., Perret J.E. and J. de Rougement, Long-term suppression of tremor by chronic stimulation of the ventral intermediate thalamic nucleus. *Lancet* **337** (1991) 403-406.

Beuter, A. and A. de Geoffroy, Can tremor be used to measure the effect of mercury exposure in human subjects. *Neurotoxicology* **17** (1996) 213-227.

Brockwell, P.J. and R.A. Davis, *Time Series: Theory and Methods* (Springer, New York, 1987).

Clayton, H.R., Murray A., Campbell, R.W.F., Analysis of the body surface ECG measured in independent leads during ventricular fibrillation in humans *PACE* **18** (1995) 1876-1882.

Cleeves, L. and L.J. Findley, Variability in amplitude of untreated tremor. *J. Neuro. Neurosurg. Psych.*, **50** (1987) 704-708.

Dahlhaus, R. and D. Janas, A frequency domain bootstrap for ratio statistics in time series analysis. *Annals of Statistics*, **24** (1995) 1934-1963.

Deuschl, G., Krack, P., Lauk, M. and Timmer, J., Clinical neurophysiology of tremor, *J. Clinical Neurophys.* **13** (1996) 110-121.

Efron, B. and R.J. Tibshirani, *An introduction to the Bootstrap.* (Chapmann and Hall, New York, 1993).

Franke, J. and W. Härdle, On bootstrapping kernel spectral estimates. *Ann. Stat.*, **20** (1992) 121-145.

Günther, H., Brunner, R. and F.W. Klußmann, Spectral analysis of tremorine and cold tremor electromyograms in animal species of different size. *Pflüg. Arch.*, **399** (1983) 180-185.

Hall, P. and S.R. Wilson, Two guidelines for bootstrap hypothesis testing, *Biometrics*, **47** (1991) 757-762.

Hefter, H., Hömberg, H., Reiners, K. and H.J. Freund, Stability of frequency during long-term recordings of hand tremor. *Electroenceph. clin. Neurophys.*, **67** (1987) 439-446.

Herrmann, W.M., *Electroencephalography in Drug Research* (Fischer, Stuttgart, 1982).

Honerkamp, J., *Stochastic dynamical systems* (VCH, New York, 1993).

Lang and J.O. Dostrovsky, Identification and characterization of neurons with tremor-frequency activity in human globus pallidus. Exp. Brain Res., **113** (1997) 557-563

Janas, D. and R. Dahlhaus, A frequency domain bootstrap for time series. In: *Computationally intensive statistical methods.* Proc. of the 26th Symp. on the Interface, J. Sall, A. Lehman, eds. (1994) 423-425.

Khutorskaya, O.Y and N.V. Fjodorova, Application of the method of spectral-statistical analysis EMG for evaluating the influence of medication on the treatment of Parkinson's disease. *Electromyo. clin. Neurophys.*, **36** (1996) 265-268.

Marsden, C.D., Meadows, J.C., Lange, G. W. and R.S. Watson, Variations in human physiological finger tremor, with particular reference to changes with age. *Electroenceph. clin. Neurophys.*, **27** (1969a) 169-178.

Marsden, C.D., Meadows J.C., Lange G.W. and R.S. Watson, The relation between physiological tremor of the two hands in healthy subjects. *Electroenceph. clin. Neurophys.* **27**, (1969b) 179-185.

Nelson, W., Tong, Y. L., Lee, J.K. and F. Halberg, Methods for cosinor rhythmometry. *Chronobiologia*, **6** (1979) 226-231.

Nikias, C.L. and R.R. Mysore, Bispectrum Estimation: A digital signal processing framework. *Proc. IEEE*, **75** (1987) 869-891.

Pinna, G.D., Maestri, R. and A. Di Cesare, Application of time series spectral analysis theory: analysis of cardiovascular variability signals. *Med. & Biol. Eng. & Comput.*, **34** (1996) 142-148.

Press, H.P., Teukolsky, A.S., Vetterling W.T. and B.P Flannery, *Numerical Recipes* (Cambrigde University Press, 1992).

Stuart, A., 1991: *Kendall's advanced theory of statistics.* Vol. 2, London

Timmer, J., Lauk, M. and G. Deuschl, Quantitative analysis of tremor time series, *Electroenceph. clin. Neurophys.*, **101** (1996) 461-468.

Timmer, J., Lauk, M. and C.H. Lücking, Confidence regions for spectral peak frequencies *Biometrical J.* **39** (1997) 849-861.