CH8900 6 70

PSI-50

PSI-Bericht Nr. 50
November 1989

Paul Scherrer Institut

Labor für Reaktorphysik und Systemtechnik

Detrending of Non-Stationary Noise Data by Spline Techniques

K. Behringer

Paul Scherrer Institut
Würenlingen und Villigen
CH-5232 Villigen PSI

Telefon 056/99 21 11
Telex 82 74 14 psi ch
Telefax 056/98 23 27 PSI CH

PSI-Bericht Nr. 50

Detrending of Non-Stationary Noise Data by Spline Techniques

K. Behringer

November 1989

Detrending of Non-Stationary Noise Data by Spline Techniques

K. Behringer

Paul Scherrer Institute
CH-5232 Villigen PSI, Switzerland

Abstract

An off-line method for detrending non-stationary noise data has been investigated. It uses a least squares spline approximation of the noise data with equally spaced breakpoints. Subtraction of the spline approximation from the noise signal at each data point gives a residual noise signal. The method acts as a high-pass filter with very sharp frequency cutoff. The cutoff frequency is determined by the breakpoint distance. The steepness of the cutoff is controlled by the spline order.

Paper presented at the:
21st Informal Meeting on Reactor Noise (IMORN 21)
Villigen PSI, September 20-22, 1989

I. Introduction

Usual noise analysis methods are based on the assumption of stationary noise. As far as possible, non-stationary noise signals should at first be detrended in order to obtain more stationary data. Detrending, however, is always associated with a loss of information. It is desirable to have rules for determining this loss. We consider here the class of non-stationary noise data, where the noise of interest consists of stationary higher frequency components riding on an uncontrolled trend of low frequency components. In this case, the classical method of detrending is high-pass filtering. A well-known detrending procedure is the method of differencing. In parametric signal modelling methods, this procedure is contained in the ARIMA model (Box and Jenkins, 1976). The differencing type filter applies to a special linear digital filter of the non-recursive type and belongs to the class of finite impulse response (FIR) linear phase filters. When one calculates the filter transfer function, one can recognize that the filter is inflexible with respect to the choice of the frequency cutoff and the steepness at desired frequency points. There is the principal question whether the problem of detrending can be solved by any linear filter type.

It is known that spline techniques can be used for filtering (e.g. Friedrich, 1984). But we did not find indications about the spectral characteristics of this filtering approach. In the present study the detrending possibility of using a least squares spline approximation of the noise signal with equally spaced breakpoints is investigated. Subtraction of the spline approximation from the noise signal at each data point gives a residual noise signal. It will be shown that the low frequency cutoff as well as the steepness of this filter are well controllable quantities. Extremely steep filter characteristics can be obtained without any distortion by phase shifts, because the filter has zero phase shift. There is also no loss of data to be filtered.

II. Method

There exists a one-path detrending method proposed by Kuroda et al. (1985) using least squares cubic spline approximation with variable knots. It is partially based on the one-path method given by Ishida et al. (1977), but it uses different control criteria. For stability reasons, smoothness of the spline approximation can be achieved only up to and including the first derivative across each interior breakpoint. In the paper of Kuroda et al. it is remarked that the applied control criteria are rather heuristic and the problem should, in principle, be attacked from the viewpoint of the Bayesian approach.

The detrending method studied here is a simple à posteriori analysis with equally spaced breakpoints in the least squares spline approximation. All data in the data set selected from a given noise record are simultaneously included in the fit procedure. The method is applicable to relatively short data sets. The maximum allowance for the length of a data set depends on the memory size of the available computer. If x_n, n = 1, 2, ..., are the data points in the data set of the original signal, and $\bar{x}_n$ are the smoothed data points from the spline approximation which exhibits the approximated trend, then the residual or detrended data points $\hat{x}_n$ follow simply from

$$\hat{x}_n = x_n - \bar{x}_n \qquad (1)$$

The method makes available both the trend signal and the detrended signal.

An interactive FORTRAN programme has been written for the VAX 11/785. The spline function is represented in the form of B-splines. For the least squares B-spline approximation to a given data set, the IMSL routine BSLSQ is used. The B-spline representation is evaluated at each data point of the data set, using the IMSL routine BSVAL. All data points to be processed have equal weights (= 1). The main programme provides for the data handling and setting up the knot sequence. As data abscissa the current point numbers are used, starting with number 1 of the data set to be processed. The point numbers are internally handled as real quantities. The determination of the knot sequence is based on an equidistant segmentation of the noise record with adjacent segments. The segment length (number of data points in a segment) is an input parameter. The last segment must be completely filled with signal data. The boundaries of the segments are midpoints on the data abscissa and are chosen as breakpoints. A maximally possible smoothness of the spline function across each of the interior breakpoints has been adopted. If K is the spline order (K = 4 for a cubic spline), then all derivatives up to and including the (K-2)th derivative are continuous across each interior breakpoint.

There are essentially two control parameters which affect the filtering action. They are the segment length (breakpoint distance) and the spline order. They have been found to be independent. The segment length determines the low frequency cutoff in the detrended signal, while the spline order is responsible for the steepness of the filter.

III. Detrending by 4th-Order (Cubic) Spline Approximation

In this section the spline order has been fixed to a cubic spline. Results from 3 test examples are represented. The aim of the investigation was to study the influence of the segment length to the detrending efficiency.

A computer-generated stationary Gaussian white noise record with deviate (0,1), consisting of 5000 data points, was additively mixed with each of 3 different trend signals. A sampling frequency of 25.6 Hz has been assumed. For each case the white noise signal was the same or was reproduced with the same initial seed value.

The values of the detrend analysis parameters were systematically varied according to the scheme given in Table 1. Each subcase in the analysis is called hereafter a job. The number of interior breakpoints is one less than the number of segments.

Job Number	Number of Segments	Number of Data Points in a Segment	Number of Interior Breakpoints	Cutoff Frequency (Hz)
1	1	5000	0	2.56×10^{-3}
2	5	1000	4	1.28×10^{-2}
3	10	500	9	2.56×10^{-2}
4	50	100	49	1.28×10^{-1}
5	100	50	99	2.56×10^{-1}
6	200	25	199	5.12×10^{-1}
7	500	10	499	1.28
8	1000	5	999	2.56
9	1250	4	1249	3.20
10	2500	2	2499	6.40

Table 1: Detrend Analysis Parameter Values

The meaning of the last column in Table 1 is explained in Section III.1.

The power spectral density (PSD) of the residual signals was estimated by the Welch method (Welch, 1967) using fast Fourier transform (FFT) techniques. All plots of the PSD shown in this paper were made with a sample size of 256 data points and the application of the Hanning signal window with 62.5 % sample overlap. The number of averages was 50.

III.1 Case of Linearly Increasing Trend Signal

The original signal is shown in Fig. 1 [1]. The given trend line has zero value at n = 1, and the value 25 at n = 5000. It is obvious that a cubic spline function can remove a linear trend without the need of any interior breakpoint. The retrieved trend line is shown in Fig. 2.1 (Job = 1). The PSD estimated on the detrended signal is represented in Fig. 3.1. There is no peak at zero frequency. If, however, the PSD is estimated on the original signal, an extremely large peak appears at zero frequency. If the number of segments is increased whereby the segment length is reduced to keep the total number of data points constant, the trend approximation becomes more and more noisy. An intermediate case (Job = 6) is shown in Fig. 2.2, and the extreme case (Job = 10) in Fig. 2.3. The residual noise signal behaves in the opposite direction, exhibiting an increasing loss of low frequency components. The PSDs of the corresponding noise signals are represented in Figs. 3.2 and 3.3.

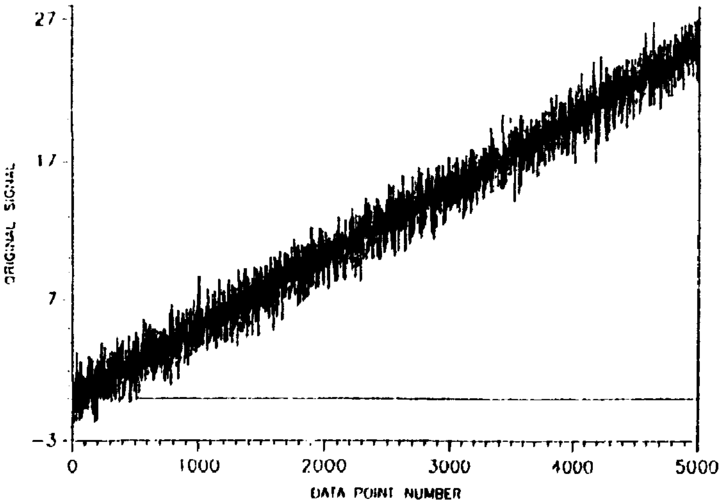

Fig. 1
Case III.1

One can observe that the spline approximation acts as a low-pass filter for obtaining the trend signal. The bandwidth increases with decreasing segment length. The method acts as a high-pass filter to separate the trend from the noise. The low frequency cutoff is shifted upwards with decreasing segment length. Considering that the spline approximation can follow frequency components with periods larger than the double segment length, we can define a low cutoff frequency f_c for the signal to be detrended by

$$f_c = (2N_s \Delta t)^{-1} \qquad (2)$$

where N_s is the number of data points in a segment, and Δt is the sampling

[1] In this and all following signal plots, the signal always starts at data point number 1, which is not resolved in the representations. For reasons due to computer graphics, the abscissa starts at zero point number.

interval. Values of f_c obtained from equation (2) are listed in the last column of Table 1. Equation (2) gives values which seem to be closely related to the 3 db cutoff frequency point used in linear filter theory. The cutoff is very sharp. There is no distortion by any overshoot in the vicinity of f_c. The coherence function between signals, which have been detrended with different values of f_c, has been found to be always equal to 1 over the frequency range above the higher value of f_c.

DETREND ANALYSIS

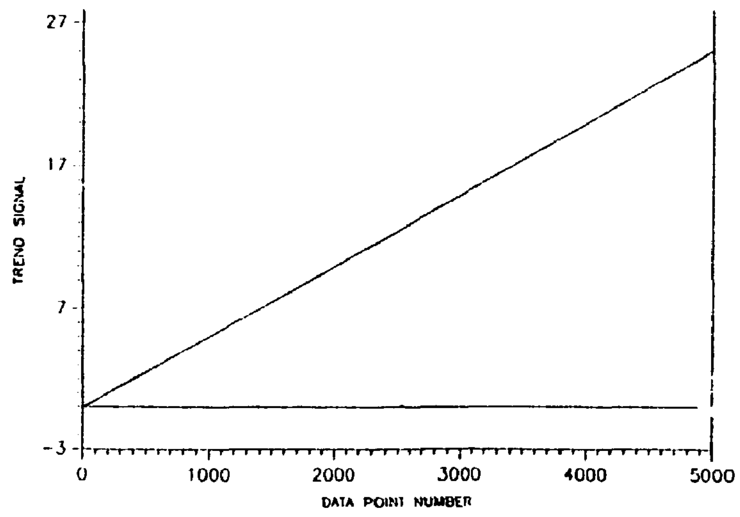

Fig. 2.1
Case III.1
Job = 1
(1 segment)

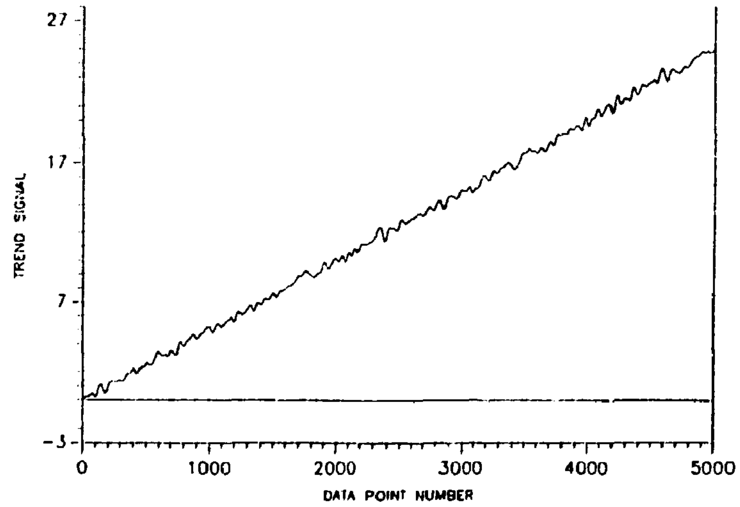

Fig. 2.2
Case III.1
Job = 6
(200 segments)

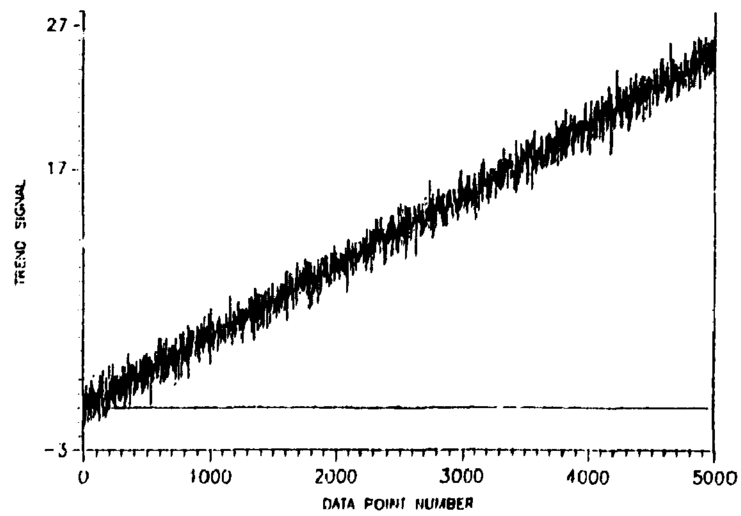

Fig. 2.3
Case III.1
Job = 10
(2500 segments)

SPECTRAL ANALYSIS

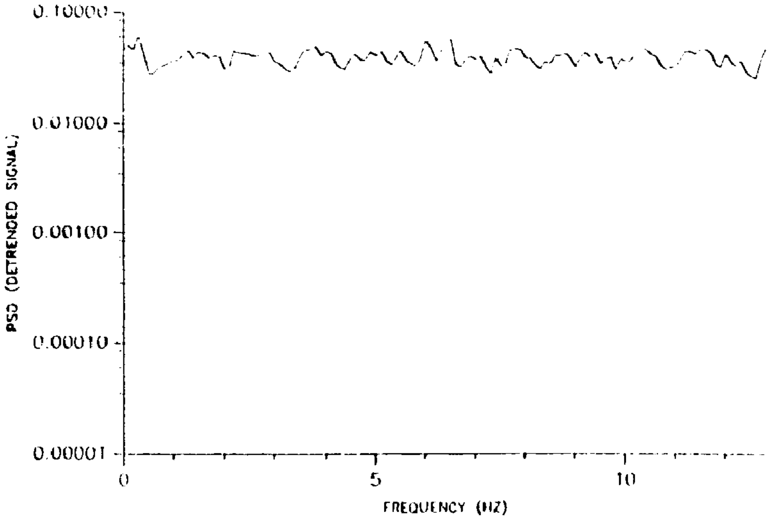

Fig. 3.1
Case III.1
Job = 1
(1 segment)

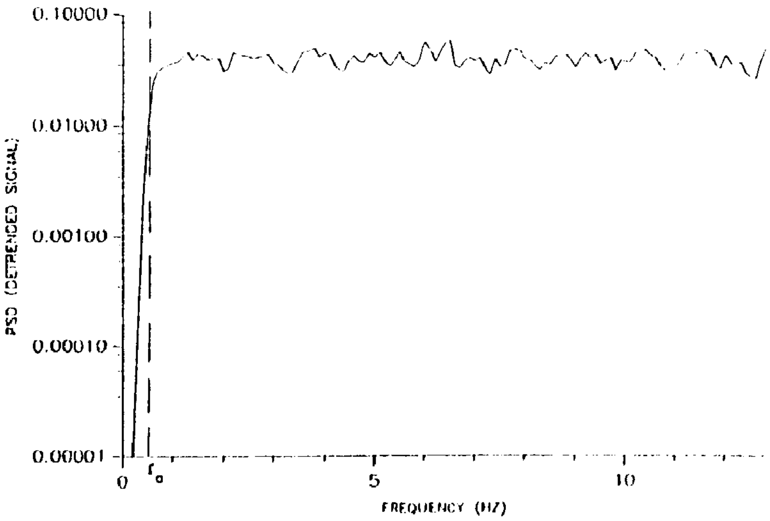

Fig. 3.2
Case III.1
Job = 6
(200 segments)

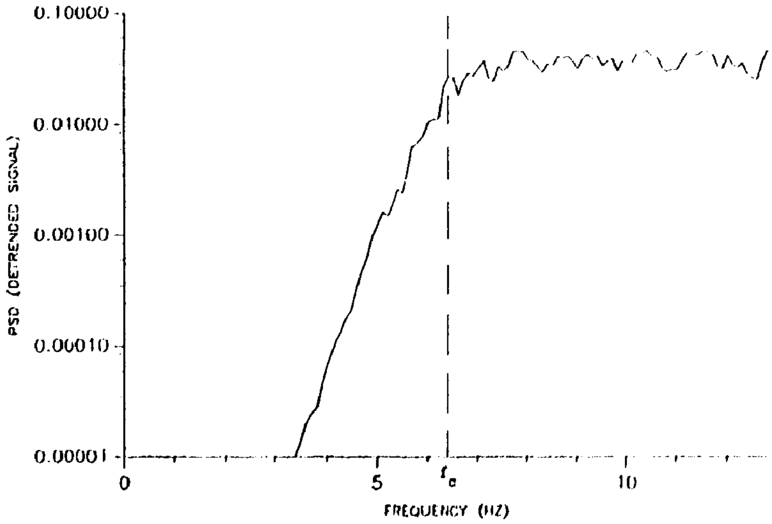

Fig. 3.3
Case III.1
Job =10
(2500 segments)

III.2 Case of Slowly Varying Trend Signal

The shape of the selected trend function is similar to that of the test example considered in the paper of Ishida et al. (1977). It is given by the broken rational function

$$F(t_n) = (0.04 + (t_n - 0.3)^2)^{-1} + (0.06 + (t_n - 1.2)^2)^{-1} \qquad (3)$$

$$\text{with } t_n = 4.0008 \times 10^{-4}(n-1); n = 1, ..., 5000$$

The original signal is shown in Fig. 4. The PSD estimated on this signal showed again a very large peak at zero frequency, but with a basic width larger than in the previous case. This width suggested choosing the detrending parameter values for a f_c value which is greater than 0.1 Hz. The conditions of Job = 3 are not yet sufficient for a complete detrending. Fig. 5.1 shows the trend signal approximation. In Fig. 6.1 one can observe some oscillatory influence in the residual noise signal from the trend, and in Fig. 7.1 a peak at zero frequency is still present in the PSD. With the conditions of Job = 5, however, a complete detrending is achieved. This is shown in the corresponding Figs. 5.2, 6.2 and 7.2. If the segment length is decreased further, the PSD of the residual noise behaves as in case III.1.

DETREND ANALYSIS

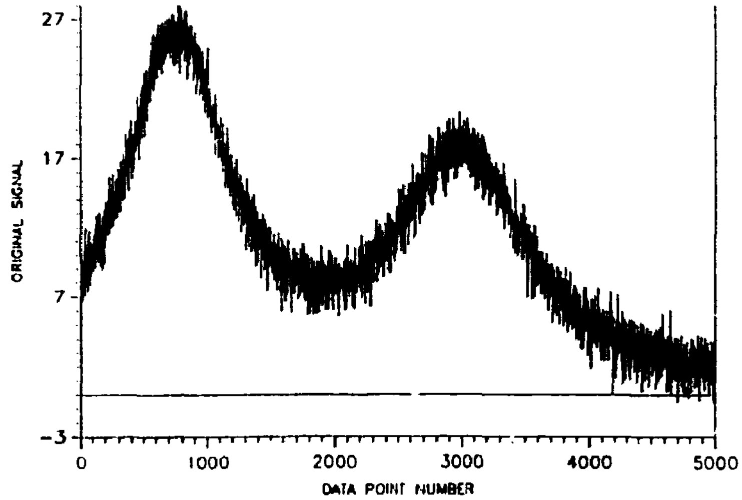

Fig. 4
Case III.2

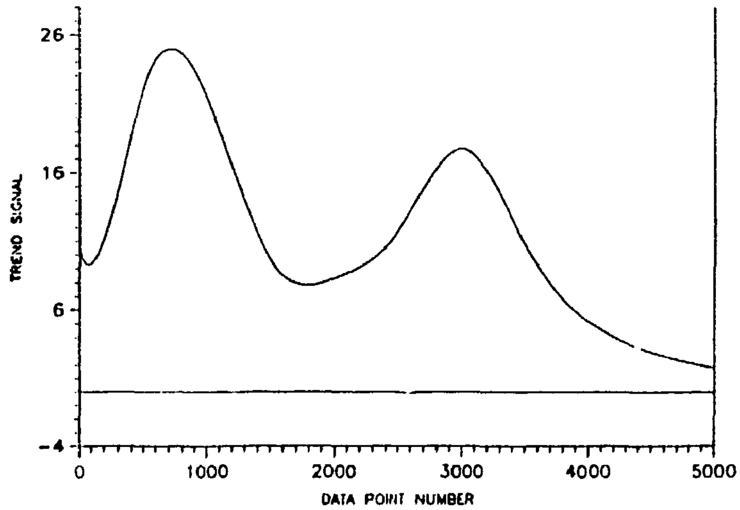

Fig. 5.1
Case III.2
Job = 3
(10 segments)

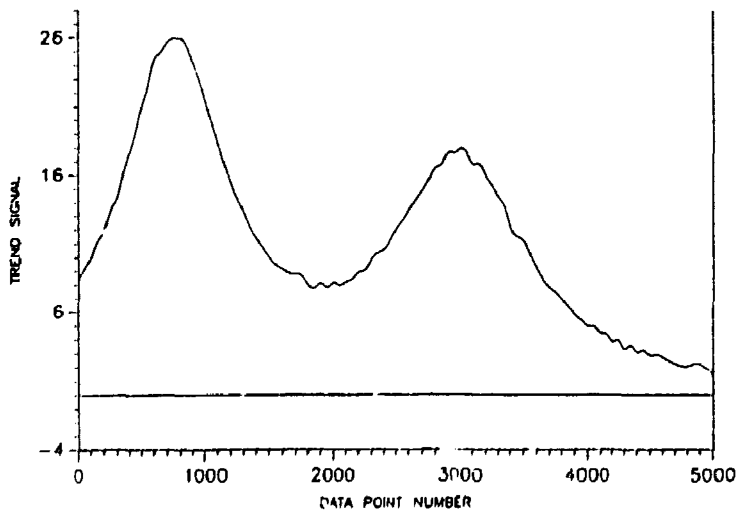

Fig. 5.2
Case III.2
Job = 5
(100 segments)

DETREND ANALYSIS

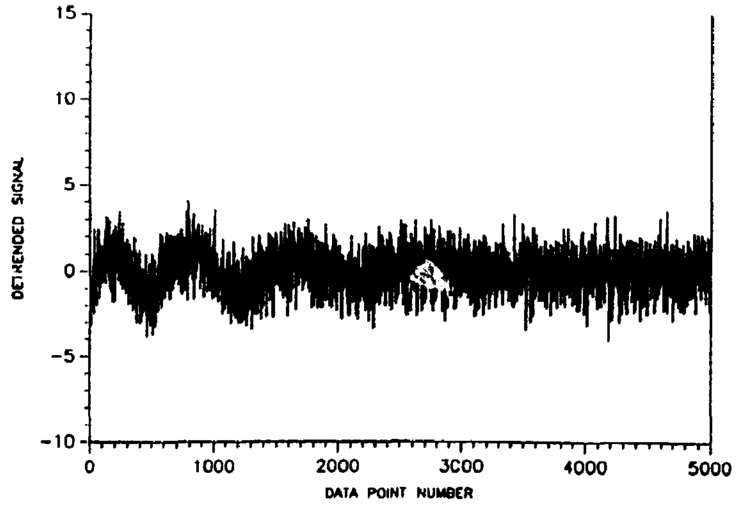

Fig. 6.1
Case III.2
Job = 3
(10 segments)

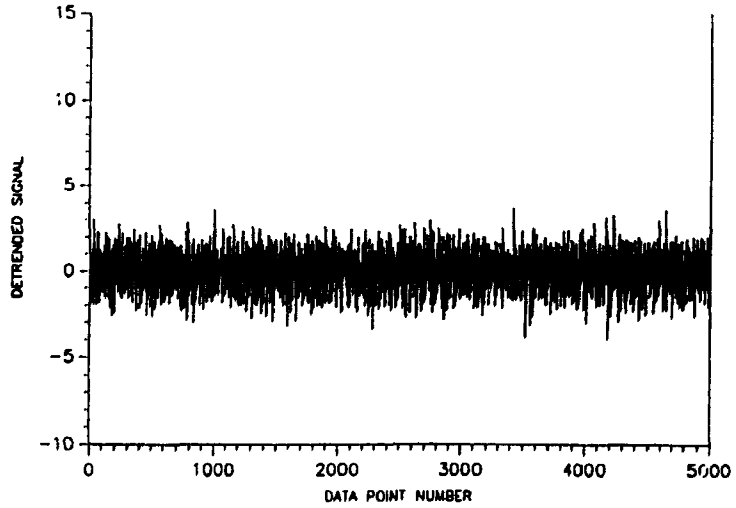

Fig. 6.2
Case III.2
Job = 5
(100 segments)

SPECTRAL ANALYSIS

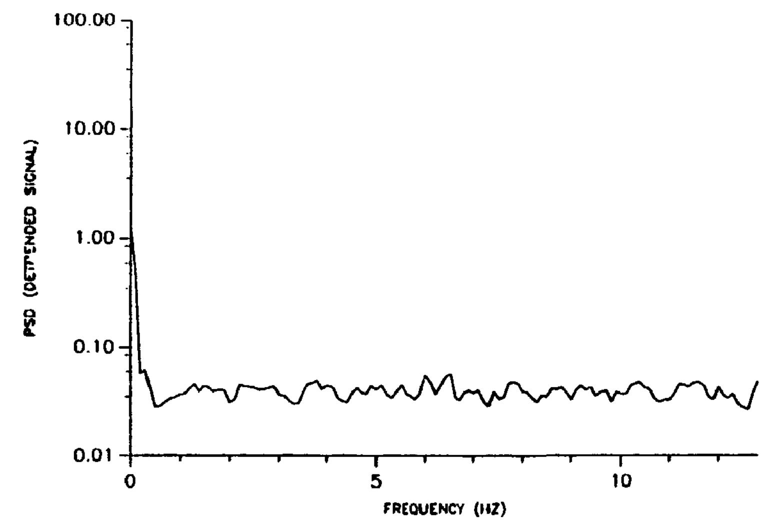

Fig. 7.1
Case III.2
Job = 3
(10 segments)

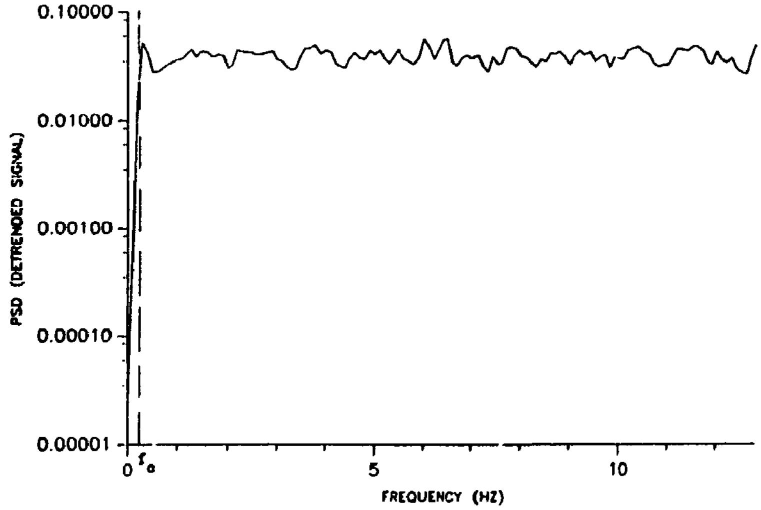

Fig. 7.2
Case III.2
Job = 5
(100 segments)

III.3 Case of an Exponentially Increasing Sinusoidal Trend Signal

This case does not exactly fit into the considered class of non-stationary noise data with smooth low frequency trend components. But it was interesting to investigate how the fit procedure reacts to this type of signal. The given trend signal consists of a sinusoid of frequency f_0 = 1 Hz with an exponentially increasing envelop having the value zero at n = 1 and the value 25 at n = 5000. The signal shown in Fig. 9.1 is practically identical with the original signal. For the conditions of Job = 6,7 and 8, the trend signal approximations are shown in Figs. 8.1, 8.2 and 8.3, the residual signals are represented in Figs. 9.1, 9.2 and 9.3, and the PSDs of the residual signals are given in Figs. 10.1, 10.2 and 10.3. As long as f_c is below f_o, the trend signal approximation is a noisy zero line. The sinusoidal peak in the PSD is fully developed. If f_c is sufficiently above f_o, the spline approximation follows the given trend signal. The peak in the PSD vanishes. There is the intermediate case of Job = 7. The given trend signal is not completely removed. In the PSD (Fig. 10.2) a small remaining part of the 1 Hz sinusoidal peak is still visible, but the fit procedure creates another peak at about 1.6 Hz. This test example demonstrates that the detrending method works, only if f_c is chosen sufficiently above f_o, i.e. about a factor 2 greater.

DETREND ANALYSIS

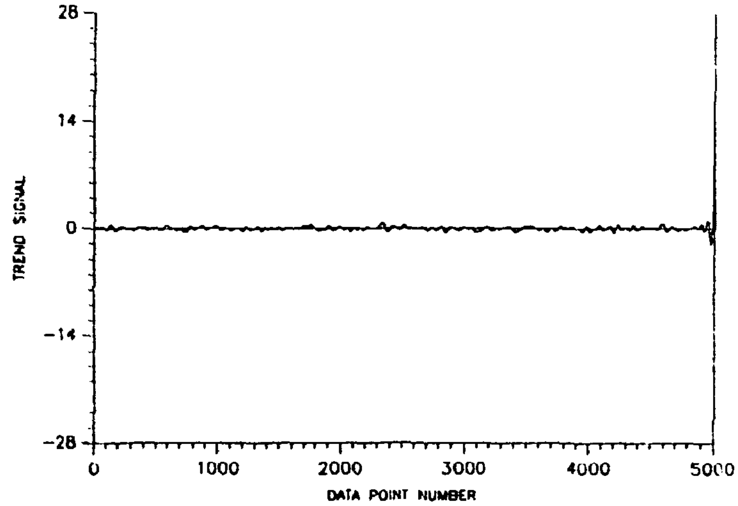

Fig. 8.1
Case III.3
Job = 6
(200 segments)

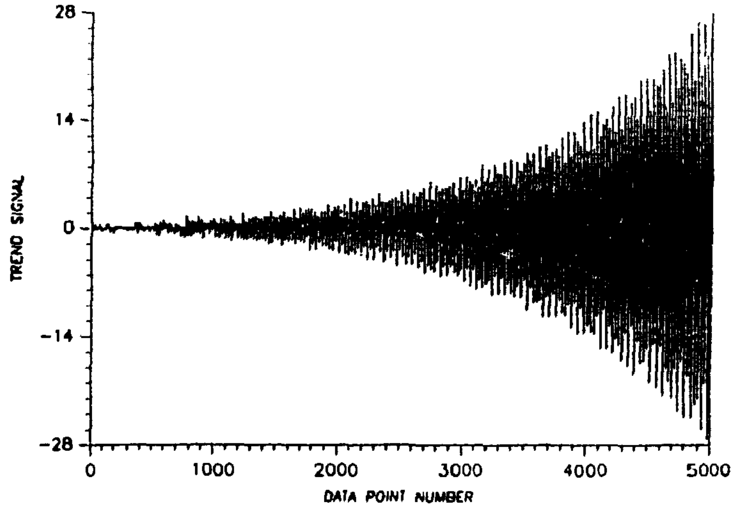

Fig. 8.2
Case III.3
Job = 7
(500 segments)

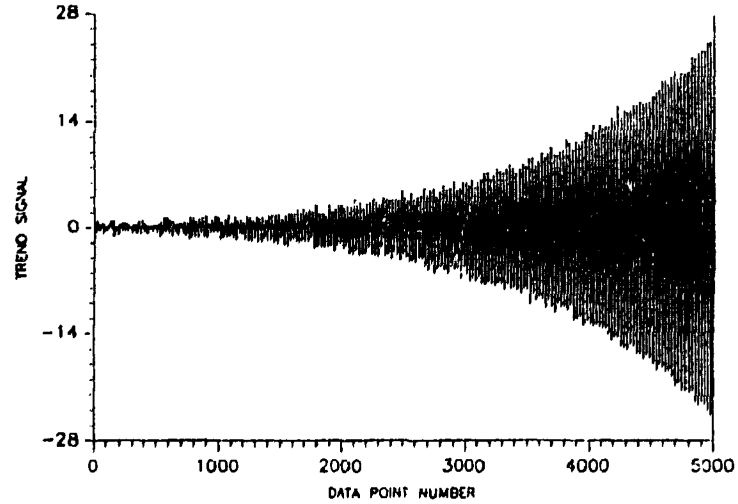

Fig. 8.3
Case III.3
Job = 8
(1000 segments)

DETREND ANALYSIS

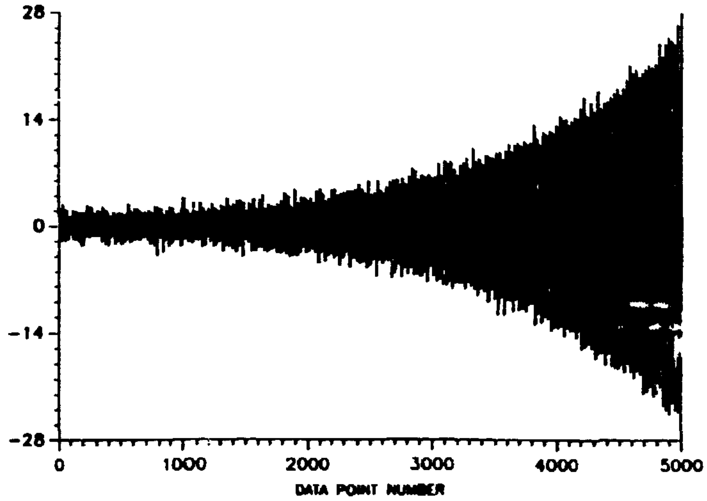

Fig. 9.1
Case III.3
Job = 6
(200 segments)

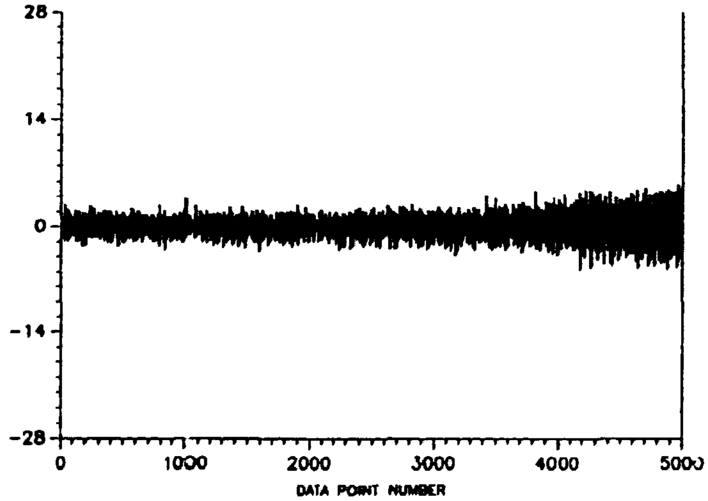

Fig. 9.2
Case III.3
Job = 7
(500 segments)

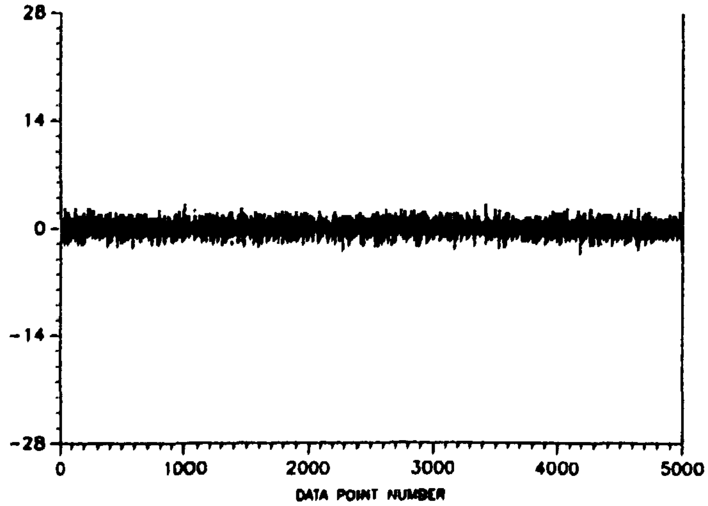

Fig. 9.3
Case III.3
Job = 8
(1000 segments)

SPECTRAL ANALYSIS

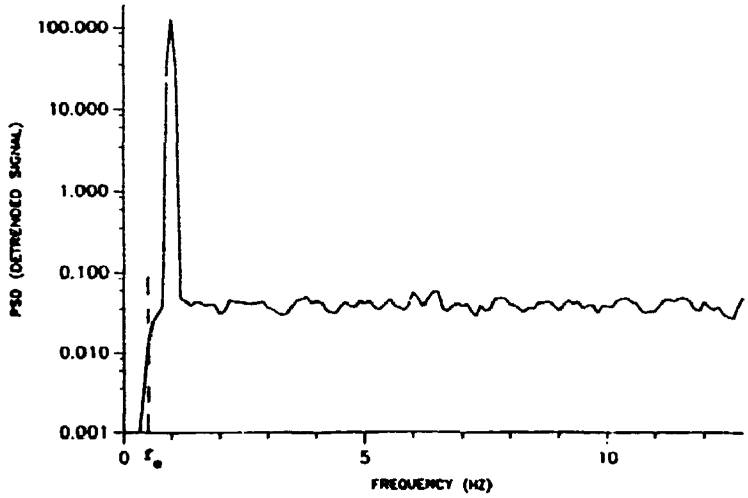

Fig. 10.1
Case III.3
Job = 6
(200 segments)

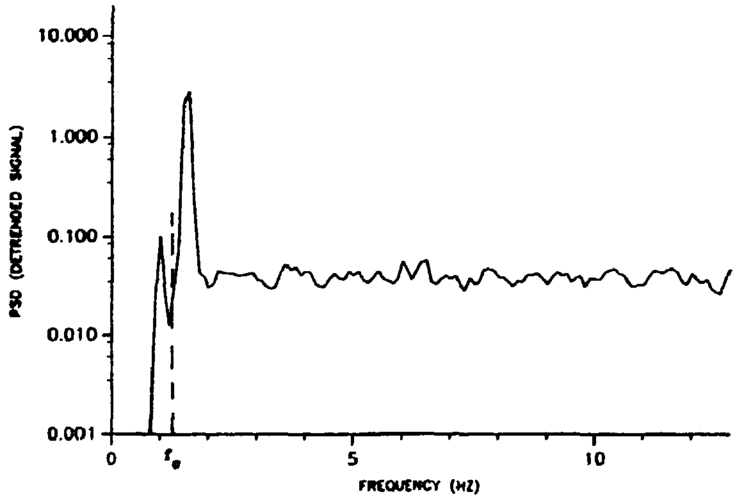

Fig. 10.2
Case III.3
Job = 7
(500 segments)

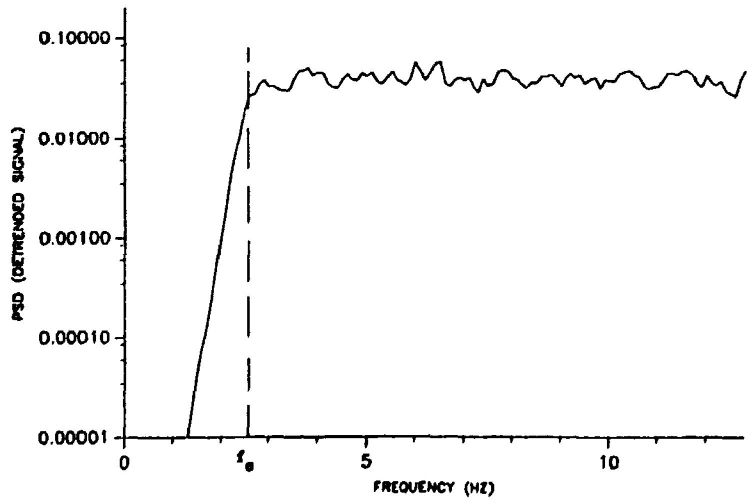

Fig. 10.3
Case III.3
Job = 8
(1000 segments)

IV. Detrending by Higher Order Spline Approximations

The signals of the test examples III.1 and III.2 were taken to study the filter action with spline functions of order higher than K = 4. K was varied up to K = 10. The cutoff frequency f_c defined by equation (2) is independent of the spline order. This was experimentally found to be true. For the conditions of Job = 10, the estimated PSDs are shown in Fig. 11.1 for K = 7, and in Fig. 11.2 for K = 10. These figures must be compared with Fig. 3.3, where K = 4. One can recognize that an increase of the spline order makes the filter steeper. There are more degrees of freedom for cutting the low frequency components. For the evaluation of the steepness, the PSD of the residual signals were estimated with sample sizes of up to 1024 data points in order to obtain sufficient frequency resolution. Recalling that the noise in the original signals is white, the steepness has been defined in db units simply by

$$Steepness = 10 log_{10}(PSD(0.5 f_c)/PSD(f_c)) \qquad (4)$$

The dependence of the steepness on the spline order is plotted in Fig. 12. Each data point is an average obtained from steepness determinations by Job = 6-10. The last point for K = 10 is somewhat unsafe, because it is influenced by the computer background noise. Fig. 12 shows that extremely sharp cutoffs can be obtained.

SPECTRAL ANALYSIS

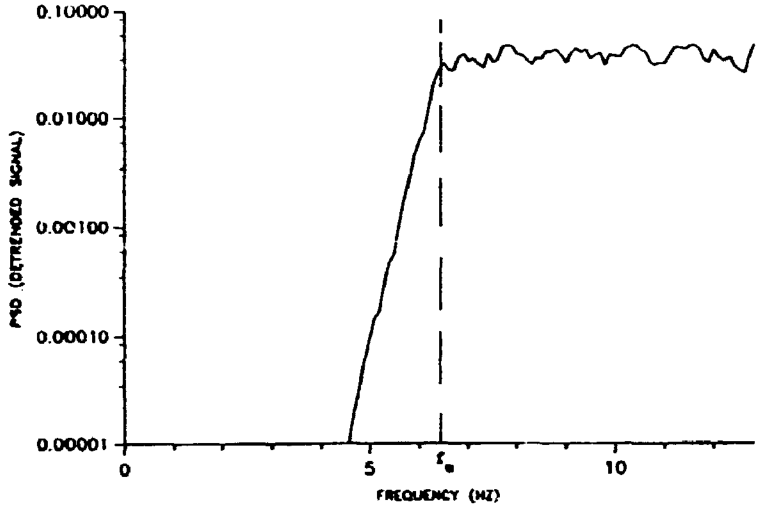

Fig. 11.1
Job = 10
(2500 segments)
K = 7

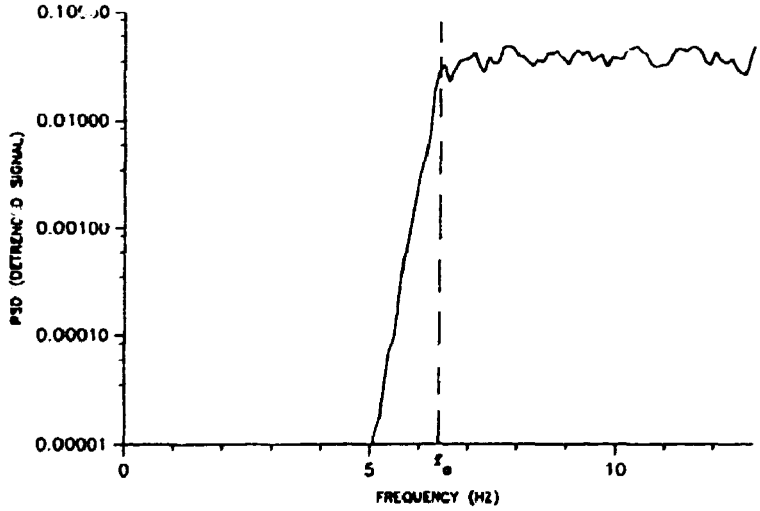

Fig. 11.2
Job = 10
(2500 segments)
K = 10

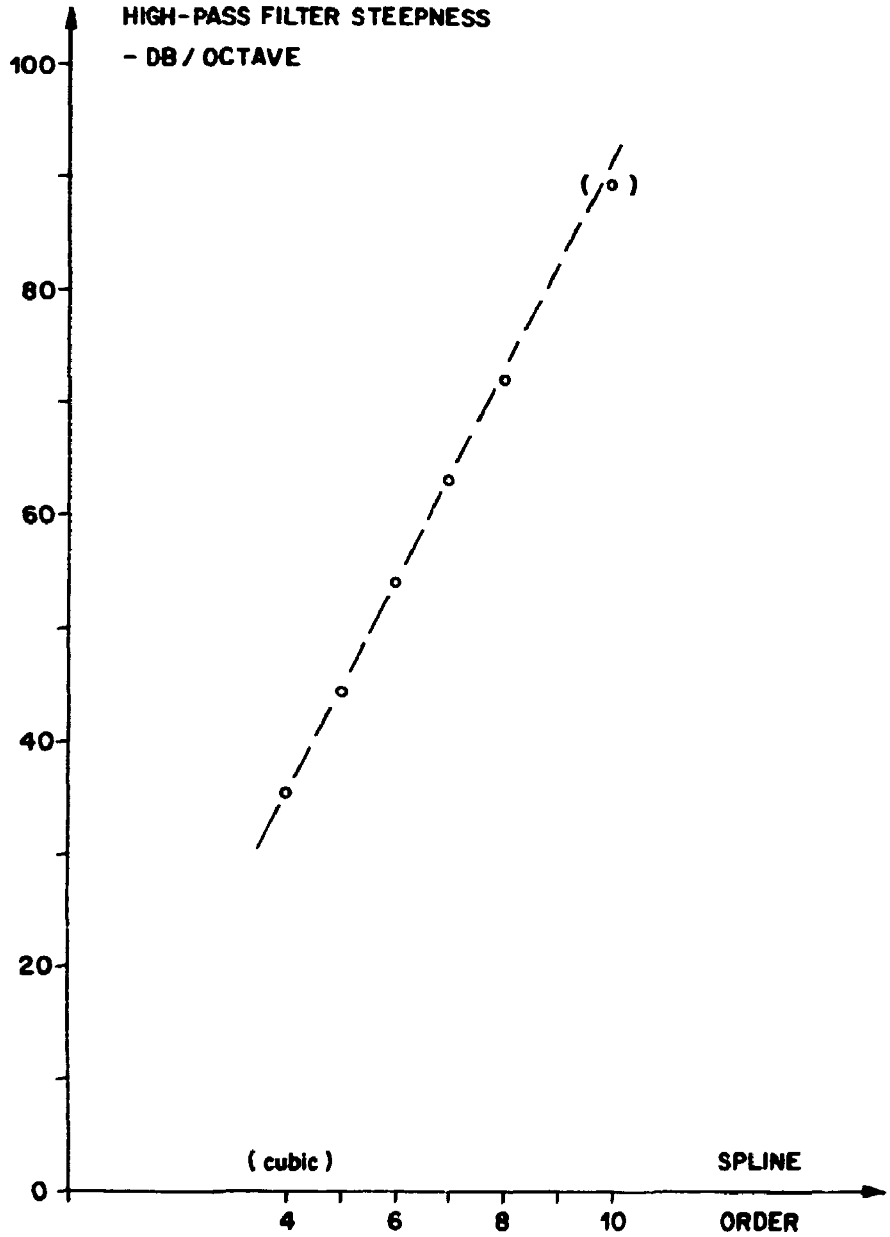

Fig.12 DEPENDENCE OF THE FILTER STEEPNESS ON THE SPLINE ORDER

V. Summary

A class of non-stationary noise data is considered where the noise of interest contains stationary higher frequency components riding on a trend function of low frequency components. A simple detrending method has been investigated which uses a least squares spline aproximation of the noise data with equally spaced break points. Cubic and higher order spline functions with a maximally achievable smoothness across the interior breakpoints have been considered. The method exhibits a sharp high-pass filter of zero phase shift and removes very effectively smooth low frequency trend functions. There are two independent control parameters. The breakpoint distance determines the cutoff frequency. The steepness of the cutoff is controlled by the spline order. Extremely steep cutoffs can be obtained. The detrending method is an à posteriori analysis applicable to relatively short noise records.

References

[1] Box G.E.P. and Jenkins G.M. (1976). Time Series Analysis, Forecasting and Control, Holden-Day.

[2] Friedrich D. (1984). Interpolation, Glättung and Saisonbereinigung von Zeitreihen mit Splinefunktionen, Statistische Hefte 25, 13, Springer Verlag.

[3] Ishida K., Kiyono T. and Yoshimoto F. (1977). ACM Trans. Math. Software 3, 164.

[4] Kuroda Y., Uchida H. and Ohmasa Y. (1985). Progr. Nucl. Energy 15, 849.

[5] Welch P.D. (1967). IEEE Trans. Audio Electroacoustics 15, 70.