

The Accuracy of Time Series Analysis for Laser-Doppler Velocimetry

P.M.T. Broersen, S. de Waele and R. Bos

Department of Applied Physics,
Delft University of Technology
P.O.Box 5046, 2600 GA Delft
The Netherlands

tel. +31 15 2786419,
fax +31 15 2784263,
email broersen@tn.tudelft.nl

ABSTRACT

Four approaches for the processing of unequally sampled data are known, each with many variants. The Lomb-Scargle approach uses the precise time information in computing spectra and is a useful method for the detection of harmonic peaks and much less for the spectral slope or other details. The second approach consists of several different slotting methods which determine an equidistant covariance estimate of the data. Unfortunately, none of the slotting covariance estimators has the important positive semi-definite property that is required for a valid covariance estimate. Hence, spectral estimates become negative at some frequencies and the logarithm of the spectrum is not defined there. Theoretically, no useful interpretation can be given to the estimates, because they no longer give the distribution of the total power over the frequencies. Many local peaks appear in the log spectral density between the negative spectral estimates if the negative estimated values are replaced by zero. The third category fits parametric spectral models to raw data. It imposes a spectral shape, independent of the data, and it can only be useful if that shape is known a priori.

The fourth method first resamples the irregular data on an equidistant time base. Nearest Neighbor takes the nearest original observation for each resampled value. Afterwards, it uses the familiar and accurate equidistant signal processing algorithms for the evaluation. Time series models can give an accurate spectral representation for turbulence data, if the model type and the model order are selected properly. This is demonstrated by the fact that the spectral density is retrieved accurately with the spectrum of time series models in simulations. Resampling causes a distortion of the estimated spectrum, that increases strongly with frequency. Therefore, nearest neighbor resampling is limited in disclosing spectral details at higher frequencies. Although strongly filtered, peaks can be retrieved at frequencies up to the average sampling rate, so two times the Nyquist frequency that would belong to equidistant sampling. For still higher frequencies, even very strong peaks in the original irregular signal do not give rise to any ripple in the estimated spectral density after resampling.

For frequencies above the average sampling rate, the Lomb-Scargle method can detect the presence of narrow large peaks, if the power of the noise is not too high.

1. INTRODUCTION

The problem of irregularly spaced observations in time has two important areas of application: turbulence data observed with laser-doppler instruments (Tropea, 1995; Britz and Antonia, 1996) and astronomical data (Lomb, 1976; Scargle 1983). In astronomical data, the most important objective is generally to detect a periodicity in the data. The signals consist of sharp peaks and some more or less white background noise. Turbulence data are characterized by spectral slopes and possibly some broader spectral peaks. In principle, irregularly spaced data are not subject to aliasing and spectra can be estimated up to frequencies much higher than the limiting Nyquist frequency that applies for equidistantly observed data. However, most methods exhibit undesirable properties at higher frequencies. The main interest in this paper is the study of methods that can estimate or approximate spectra at high frequencies. The performance of four different approaches is discussed.

The Lomb-Scargle method (Lomb, 1976 and Scargle, 1989) has been introduced by Lomb for astronomical signal processing and a more complete description has been given by Scargle; it is sometimes denoted Lomb only. It uses the exact time information, a property that it shares with the direct Fourier transform of Gaster and Roberts (1977). The method is absolutely free of aliasing and it can find spectral peaks at frequencies that are as much as fifty times higher than the Nyquist frequency, the maximum frequency at which peaks can be found in equidistant data with the average irregular time spacing. Lomb-Scargle is powerful in finding narrow peaks with a low noise level. However, it loses its strength with smaller peaks and more noise. Moreover, the spectral background found in Lomb-Scargle spectra is often flat, independent of the true spectral slopes, like it is expected in white noise. The method is especially developed for the detection of peaks. The integral of the estimated spectrum depends strongly on the highest frequency for which the estimate is calculated. That maximum frequency is not necessarily limited to some given value. Therefore, the Lomb spectrum has no direct interpretation as a distribution of power over the frequencies. The computed spectral density is always positive. The properties of its inverse Fourier transform, that should be the covariance function, are not easily established. Subjectively choosing a different number of spectral frequency points for evaluation, with some chosen frequency interval and a chosen maximum frequency, would yield different covariance estimates from the same data. Its inverse transform does not necessarily yield a reliable estimate for the covariance function. Suppose a high frequency sine wave is not detected because the highest computation frequency was chosen too low. The true covariance is a high frequency sine wave, the estimated covariance is approximately a delta function, belonging to white noise.

The different slotting methods determine an equidistant covariance estimate of the data. The main attractive property of slotting is that the covariance function can have a small bias. It is believed that this will also produce accurate spectral estimates. However, an unbiased estimate for the spectrum does not guarantee that the log spectrum is unbiased: *the expectation of the logarithm of a spectral density is generally not equal to the logarithm of the expectation of the spectral density*. Unfortunately, none of the slotting covariance estimators has the important positive semi-definite property that is required for a valid covariance estimate. Only valid covariance estimates guarantee a positive spectrum for all frequencies. A negative spectral estimate may not seem important in linear spectral plots, but it has an enormous influence in logarithmic presentations of the results, which are common in LDA processing. Two possible reasons for violation of the positive semi-definite property are: taking only a truncated part of the covariance function, and the normalization of the estimated covariance. The first problem can be solved by using a standard window over the truncation length of the slotted covariance, but the second problem has not yet a useful solution. Therefore, a theoretical analysis of the spectral properties of slotting is impossible. Too many frequencies give a negative raw spectral estimate. As the integral of the Fourier transform of the slotted covariance is the total power $R(0)$, no interpretation as a distribution of the power over the frequencies is possible. The performance in the detection of strong spectral peaks is generally less than that of the Lomb-Scargle estimator. Further, spectral slopes in the estimates depend more on the windowing that is applied than on the true spectral shape of the irregular data. The visual appearance of logarithmic spectral plots depends on the arbitrary choice about what is done with *negative* spectral estimates, for which the logarithm does not exist. Finally, the equidistant covariance estimate is somewhat sensitive to aliasing if the slotting width becomes larger than a half period of the fastest phenomena.

Slotting can be studied by applying the slotting algorithm to equidistant data, with or without missing observations. The performance of slotting depends heavily on the spectral density of the data at hand. A theoretical study of slotting in equidistant observations shows: the unbiased covariance estimate is not positive semi-definite; see Priestley, 1981. Further, using a Parzen window on this unbiased estimated covariances still does not restore the positive semi-definite

property that among other things is necessary to ensure positive estimates for the spectral density. However, even with the best methods, slotting cannot be made positive semi-definite. The spectral quality cannot be judged by visual improvements in the estimated covariance, by looking at the variance of estimates close to zero shifts. It might seem to become less poor, but frequencies with weak power have no influence whatsoever on this “measure”; see Broersen and de Waele (1999). Arbitrary variations in covariances which are one or two orders of magnitude smaller than the standard deviation of the covariance in estimation can be sufficient to produce a non-positive semi-definite estimate. Techniques to transform arbitrary covariance estimates to positive semi-definite functions exist; see Rousseeuw and Molenberghs (1993). However, those techniques have not yet been applied to a large amount of covariance data and they use numerical rather than statistical properties for the transform.

The results of slotting have been improved with local normalization and variable windowing by Tummers and Passchier (1996). The variable window technique takes a wide covariance window at low frequencies and a small one at high frequencies. As such, it can sometimes give an integrated contribution over the spectral window width that is made positive everywhere. It can be considered as taking many different fixed length windows and selecting according to some subjective preference which one of those spectral estimates is used as a function of frequency.

The third category fits a parametric model, a prototype covariance or spectral model to the data; see van Maanen and Oldenziel (1998) and Müller et al. (1998). This method imposes a spectral shape, independent of the data, and it can only be useful if that shape is known a priori. Mathematically, minimization of errors in the covariance function is not attractive because the estimation errors will not at all be independent. Moreover, using estimated covariances to estimate parameters means that no attention is given to areas of the spectrum with low power, see Broersen and de Waele (1999). Minimization of differences between the logarithm of estimated and prototype spectra might be more attractive.

The fourth method first resamples the irregular data on an equidistant time base. Afterwards, it uses the familiar and accurate equidistant signal processing algorithms for the evaluation. Resampling introduces noise and filtering and is limited in disclosing details at frequencies higher than the average sampling rate. If very large peaks have no visible influence on the spectrum that is computed, it is hard to believe that the spectral shape at those frequencies can be reconstructed with undoing filtering and noise removal. Resampling largely removes velocity bias in irregular samples without making use of any model of the data, so it is very robust. Resampling will be followed by Fourier analysis with tapered and windowed periodograms or by time series analysis. Resampling has the disadvantage that it introduces a filtering operation that becomes evident if the resampling rate is greater than about 15 % of the average data rate. In equidistant spectral analysis, the accuracy of time series models is better than the accuracy of the usual windowed and tapered periodogram approach; see Broersen (1998c). In this paper, the maximum resampling rate with nearest neighbor resampling followed by time series analysis will be discussed. The advantages that can be obtained by using lower resampling rates are not discussed in this paper. They have been described by De Waele and Broersen (2000). An important advantage of a time series model at a larger time scale (and hence a smaller frequency range) can be the detection of smaller spectral peaks and details in that limited frequency range.

A promising new method could probably be the estimation of continuous-time autoregressive models from irregular data. However, the progress in that field is slow, even for equidistant observations; see Söderström and Mossberg (1998). Only very simple models can be estimated so far, not detailed enough for the usual turbulence data.

The paper discusses thoroughly the properties of resampling methods. Conclusions will be based on our own data and on data from the LDA benchmark generator II of Nobach (1999). Special attention is given to the maximum resampling frequency for Nearest Neighbor Resampling followed by the time series program ARMA_{sel}, and to what happens if the resampling rate is increased.

2. TIME SERIES MODELS

Time series models are a general class of models for all stationary stochastic processes; see Priestley (1981). It uses a general class of models that does not impose any type of restriction on the shape of the spectra to be modeled. Time series models give a parametric description of signal properties with the parameters of an Autoregressive (AR), a Moving Average (MA) or a combined ARMA model. Together, those three model types are an excellent collection for the modeling of arbitrary stationary stochastic data with unknown properties. At least one of the three types with a properly selected model order gives a good time series model for measured data. The power spectrum and the covariance of the data can be expressed directly with the parameters of that model. A recent result in the application of time series is that the best model *type* (AR, MA or ARMA) and the best model *order* (the number of parameters in that

model) can be selected objectively and automatically, based on reliable statistical criteria; see Broersen (1998c), 2000. For equidistant data, Broersen (2000) demonstrated in many examples that the accuracy of the time series model with an automatically selected model type and model order is better than the best of all possible windowed periodogram estimates. De Waele and Broersen (1999b, 2000) showed the same for resampled LDA data. An important disadvantage of periodograms is that no objective or statistical rule exists to select the best among different estimates. Time series are also particularly suited for hot-wire analysis, obtained by sampling a continuous signal. Time series models have three different types, autoregressive or AR, moving average or MA and combined ARMA. An ARMA(P,Q) process can be written as (Priestley 1981):

$$\mathbf{x}_n + \mathbf{a}_1 \mathbf{x}_{n-1} + \dots + \mathbf{a}_p \mathbf{x}_{n-p} = \mathbf{e}_n + \mathbf{b}_1 \mathbf{e}_{n-1} + \dots + \mathbf{b}_q \mathbf{e}_{n-q}, \quad (1)$$

where \mathbf{e}_n is a purely random process, so a sequence of independent identically distributed stochastic variables with zero mean and variance σ_e^2 . This process is AR for Q=0 and MA for P=0. Almost any stationary stochastic process can be written as an unique AR(∞) or MA(∞) process (see Priestley 1981), independent of the origin of the process; it may be the sum of a true AR process and a colored noise signal or an ARMA(P,Q) process.

An estimated ARMA(p,q) model is given by

$$\mathbf{x}_n + \hat{\mathbf{a}}_1 \mathbf{x}_{n-1} + \dots + \hat{\mathbf{a}}_p \mathbf{x}_{n-p} = \hat{\mathbf{e}}_n + \hat{\mathbf{b}}_1 \hat{\mathbf{e}}_{n-1} + \dots + \hat{\mathbf{b}}_q \hat{\mathbf{e}}_{n-q}, \quad (2)$$

where p and q may be arbitrary and $\hat{\mathbf{e}}_n$ is some colored noise signal that is not purely random in general. Only if the parameters have the exact values of the true process in (1), $\hat{\mathbf{e}}_n$ will be purely random white noise. Parameters for different model orders are estimated by minimization of the residual variance, which is the variance of $\hat{\mathbf{e}}_n$. The theory shows that this minimum gives at the same time the maximal flatness for the spectral density of $\hat{\mathbf{e}}_n$. Afterwards, orders p and q are selected. Selection criteria add a penalty for each estimated parameter to the logarithm of the residual variance of each estimated model and look for the minimum of the criterion; see Broersen 1998c. Moreover, finite order models are quite well in describing true processes with infinite orders, because the true parameters are decreasing rapidly for many processes.

Practical problems prevented the routine application of time series analysis for a long time. The most important difficulty is that the model type and the model order are unknown for practical data. Moreover, the estimation is a non-linear minimization for MA or ARMA models, because both $\hat{\mathbf{b}}_i$ and $\hat{\mathbf{e}}_n$ must be estimated from the data \mathbf{x} . A satisfactory solution, using robust algorithms with only linear estimations has been given and is available from Broersen (1998c). It is called **ARMAset** and it automatically selects a single model order and model type from N equidistant observations. All measured data, if stationary and stochastic, can be analyzed. Automatically a single time series model type and model order is selected from about 250 estimated AR models, 100 MA models and 50 ARMA models. The quality of this selected model is, in numerous simulations, comparable with the quality that could be obtained if the true process type and order would be known in advance. In other words, not knowing the true process is no impediment for automatically finding the best model. Examples of computer algorithms that take care of reliable estimation and order selection have been described by Broersen (1998c).

The covariance function and the power spectral density of the data can be computed from the parameters of the estimated model. The power spectrum $\hat{\mathbf{h}}(\omega)$ of the ARMA(p,q) model is given by (Priestley, 1981):

$$\hat{\mathbf{h}}(\omega) = \frac{S_e^2}{2p} \left| \mathbf{1} + \sum_{m=1}^q \hat{\mathbf{a}}_m \exp(-j\omega m) \right|^2 / \left| \mathbf{1} + \sum_{m=1}^p \hat{\mathbf{a}}_m \exp(-j\omega m) \right|^2. \quad (3)$$

The normalized spectrum $\hat{\mathbf{f}}(\omega)$ is defined as the result of (3) divided by the variance of the signal, S_x^2 . The covariance can be computed directly from (1) or (2), and it can be found approximately as the inverse Fourier transform of (3). Also specific time constants for turbulent flow can be expressed directly in the parameters of an estimated time series model. As an example, the integral time scale Γ is defined as the integral from zero to infinity of the normalized autocorrelation function. This can be found by substituting the time series estimates for the correlation in the formula:

$$G = \sum_{k=0}^{\infty} \hat{\mathbf{r}}(k) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \hat{\mathbf{r}}(k) + \frac{1}{2} \hat{\mathbf{r}}(0) = p\hat{\mathbf{f}}(0) + \frac{1}{2} = \frac{S_e^2}{2S_x^2} \frac{\sum_{i=1}^q (1 + \hat{\mathbf{a}}_i \hat{\mathbf{b}}_i)^2}{\sum_{i=1}^p (1 + \hat{\mathbf{a}}_i \hat{\mathbf{a}}_i)^2} + \frac{1}{2}. \quad (4)$$

This expression is similar to an estimator for the variance of the mean, given by Broersen (1998a). For that application, it has been shown that the accuracy of the time series estimator is close to the Cramér-Rao bound for the best attainable statistical accuracy in some examples.

At first sight, time series modeling and periodogram estimation are different. However, the inverse Fourier transform of an estimated (positive) windowed and tapered periodogram is based on a positive semi-definite estimate for the covariance. It is possible to recover a covariance estimate from the periodogram if zeros had been added before the Fourier transformation of the data; see Priestley (1981). As the covariance is of finite length, say M , it can be proved that a signal with this covariance can be generated by an MA process of order M . As such, periodogram estimates can be considered as MA models, a part of the more general class of time series models. However, MA models are least preferred in the automatic selection of the model type of time series models for practical data. Hence, it is not surprising that the quality of a selected time series model is often better than the best possible periodogram estimate.

An important class of accuracy measures for time series models contains the prediction error and also the spectral distortion and the likelihood ratio. The prediction error PE in the time domain is defined as the mathematical expectation of the squared error of prediction when applying the estimated model to another independent realization of that process. This is the variance of $\hat{\epsilon}_n$ in (2), provided that the estimated parameters are not found with this realization of x_n . Broersen (1998b) showed that it is equivalent with a *relative* error in the frequency domain:

$$PE = S_{\hat{\epsilon}}^2 = \frac{S_{\epsilon}^2}{2p} \int_{-p}^p \left| \frac{\hat{A}(e^{j\omega})\mathbf{B}(e^{j\omega})}{A(e^{j\omega})\hat{B}(e^{j\omega})} \right|^2 d\omega . \quad (5)$$

Here, $A(e^{j\omega})$ is defined as $1+a_1e^{-j\omega}+\dots+a_p e^{-jp\omega}$ with the parameters a_i as defined in (1) and similar definitions for the other polynomials in (5) with true and estimated AR and MA parameters in (1) and (2) respectively. So a computation of PE requires only knowledge of true and estimated parameters and the variance of ϵ_n . It is easily seen that for estimated parameters close to the true values, the integrand in (5) becomes close to 1, which means that the spectrum of $\hat{\epsilon}_n$ becomes almost white. A scaled value of PE is the model Error ME (Broersen, 1998b):

$$ME = N(PE / S_{\epsilon}^2 - 1) . \quad (6)$$

The expectation of ME equals $p+q$ if $p > P$ and $q > Q$, so for models with at least all non zero parameters included. This measure ME is corrected for the influence of N , the number of observations used to estimate the model. Moreover, if $p > P$ and $q > Q$ it is independent of the true shape of the process. That means that statistical significance of extra parameters can be evaluated independent of the true shape of the spectrum.

3. COVARIANCE OF RESAMPLED DATA

The irregular signal x is measured at N irregular time instants $t_1 \dots t_N$. The average distance between samples is denoted T_0 and is given by $T_0 = (t_N - t_1)/(N-1) = 1/f_0$. The signal is resampled at αN equidistant time instants at distance $T_R = T_0/\alpha$ and the resampled signal exists only for $t = nT_R$, with n integer. The spectrum can be calculated up to frequency $\alpha f_0/2$. It has been shown by de Waele and Broersen (2000) that the most accurate reconstruction of the signal does not necessarily yield the best spectral estimate. Linear and cubic interpolation use more than one original observation for each resampled observation; as such, they influence the variance of the resampled signal. In contrast, Sample and Hold and Nearest Neighbor resampling are methods which use only one original irregular observation for a resampled observation, thus conserving the variance of the data. The autocovariance function of the resampled signal $x_R(nT_R)$ is given by:

$$\begin{aligned} \mathbf{R}_{x_R}(kT_R) &= \mathbf{E}[\mathbf{x}_R(t)\mathbf{x}_R(t+kT_R)] & k \geq 0 \\ \mathbf{R}_{x_R}(kT_R) &= \mathbf{R}_{x_R}(-kT_R) & k < 0 . \end{aligned} \quad (7)$$

The resampled signal $x_R(nT_R)$ is equal to the original signal $x(nT_R - \tau_1)$ and $x_R(nT_R + kT_R)$ is equal to $x(nT_R + kT_R - \tau_2)$. An approximate solution for small α , where all covariance products except the covariance at shift zero are based on two different original observations has been given by de Waele and Broersen (2000). Here the derivation is improved to cover a dense resampling scheme. There are two distinct possible situations: $P_{1point}(kT_R)$ is the probability that $x_R(t)$ and $x_R(t+kT_R)$ are both given by the same original observation and $1 - P_{1point}(kT_R)$ is the probability that two different original observations are used for the two resampled observations. Only values $k \geq 0$ are to be considered. For $k=0$, $P_{1point}(kT_R)$ equals one. The total covariance is the sum of two conditional expectations. With theorem 6 of Mood, Graybill and Boes (1974, p. 158), it follows that:

$$\mathbf{R}_{x_R}(kT_R) = \mathbf{E}[\mathbf{x}_R(nT_R)\mathbf{x}_R(nT_R+kT_R) | 1point] P_{1point}(kT_R) + \mathbf{E}[\mathbf{x}_R(nT_R)\mathbf{x}_R(nT_R+kT_R) | 2points] \{1 - P_{1point}(kT_R)\} . \quad (8)$$

$\mathbf{E}[\mathbf{x}(nT_R)\mathbf{x}(nT_R+kT_R) | 1point]$ denotes the conditional expectation of the product for the case where $x_R(t)$ and $x_R(t+kT_R)$ are both given by the same original observation. The first expectation in (8) is $\mathbf{R}_x(0)$, the variance of the signal because both $x_R(nT_R)$ and $x_R(nT_R+kT_R)$ are given by the same $x(nT_R - \tau_1)$. The second expectation becomes

$$\begin{aligned}
 \mathbf{E}[\mathbf{x}_R(\mathbf{nT}_R)\mathbf{x}_R(\mathbf{nT}_R + \mathbf{kT}_R) | 2 \text{ points}] &= \mathbf{E}[\mathbf{x}(\mathbf{nT}_R - \tau_1)\mathbf{x}(\mathbf{nT}_R + \mathbf{kT}_R - \tau_2)] \\
 &= \mathbf{E}[\mathbf{x}(\mathbf{nT}_R)\mathbf{x}(\mathbf{nT}_R + \mathbf{kT}_R - \Delta\tau)] \\
 &= \mathbf{E}[\mathbf{E}[\mathbf{x}(\mathbf{nT}_R)\mathbf{x}(\mathbf{nT}_R + \mathbf{kT}_R - \Delta\tau) | \Delta\tau]] \\
 &= \int_{-\infty}^{\infty} \mathbf{E}[\mathbf{x}(\mathbf{nT}_R)\mathbf{x}(\mathbf{nT}_R + \mathbf{kT}_R - \Delta\tau) | \Delta\tau] f_{\Delta\tau}(\Delta\tau) d\Delta\tau \\
 &= \int_{-\infty}^{\infty} \mathbf{R}_x(\mathbf{kT}_R - \Delta\tau) f_{\Delta\tau}(\Delta\tau) d\Delta\tau.
 \end{aligned} \tag{9}$$

The second line in this derivation replaces $\tau_2 - \tau_1$ by $\Delta\tau$. The third line is a second application of the above mentioned theorem of conditional expectations; $f_{\Delta\tau}$ denotes the probability density function of the interval $\Delta\tau$ and $f_{\Delta\tau}$ depends on the resampling method. Together, the resampled covariance becomes:

$$\mathbf{R}_{x_R}(\mathbf{kT}_R) = \mathbf{P}_{\text{Ipoint}}(\mathbf{kT}_R)\mathbf{R}_x(\mathbf{0}) + \left\{ \mathbf{1} - \mathbf{P}_{\text{Ipoint}}(\mathbf{kT}_R) \right\} \int_{-\infty}^{\infty} \mathbf{R}_x(\mathbf{kT}_R - \Delta\tau) f_{\Delta\tau}(\Delta\tau) d\Delta\tau. \tag{10}$$

The integral can also be rewritten as a convolution:

$$\mathbf{R}_{x_R}(\mathbf{kT}_R) = \mathbf{P}_{\text{Ipoint}}(\mathbf{kT}_R)\mathbf{R}_x(\mathbf{0}) + \left\{ \mathbf{1} - \mathbf{P}_{\text{Ipoint}}(\mathbf{kT}_R) \right\} (f_{\Delta\tau} * \mathbf{R}_x)(\mathbf{kT}_R). \tag{11}$$

For irregular white noise observations, the true covariance has only $\mathbf{R}_x(\mathbf{0})$ non-zero and the resampled covariance is given by:

$$\mathbf{R}_{x_R}(\mathbf{kT}_R) = \mathbf{P}_{\text{Ipoint}}(\mathbf{kT}_R)\mathbf{R}_x(\mathbf{0}) + \left\{ \mathbf{1} - \mathbf{P}_{\text{Ipoint}}(\mathbf{kT}_R) \right\} f_{\Delta\tau}(-\mathbf{kT}_R)\mathbf{R}_x(\mathbf{0}). \tag{12}$$

3.1 Distribution of time intervals

Suppose that the irregular time intervals have a Poisson distribution with average $1/T_0$. The distribution function for the time intervals τ_1 and τ_2 is given by de Waele and Broersen (2000) for Sample and Hold (SH) and Nearest Neighbor (NN) as

$$\begin{aligned}
 f_{t,\text{SH}} &= 0, \quad t < 0 \quad ; \quad f_{t,\text{SH}} = \frac{1}{T_0} \exp(-t/T_0), \quad t \geq 0. \\
 f_{t,\text{NN}} &= \frac{1}{T_0} \exp(-2|t|/T_0)
 \end{aligned} \tag{13}$$

If resampled observations come from two different original observations, τ_1 and τ_2 are independent and the distribution function $f_{\Delta\tau}$ of $\Delta\tau = \tau_2 - \tau_1$ is given by the convolution of f_{τ}^+ with f_{τ} where f_{τ}^+ is the mirrored version of the distribution function f_{τ} , so $f_{\Delta\tau} = f_{\tau} * f_{\tau}^+$. Some exercise yields:

$$\begin{aligned}
 f_{\Delta\tau,\text{SH}}(t) &= \frac{1}{2T_0} \exp(-|t|/T_0) \\
 f_{\Delta\tau,\text{NN}}(t) &= \frac{1}{2T_0} + \frac{|t|}{T_0^2} \exp(-2|t|/T_0).
 \end{aligned} \tag{14}$$

3.2 Distribution of covariance estimates using only one irregular observation

$\mathbf{P}_{\text{Ipoint}}(\mathbf{kT}_R)$ will now be specified for sample and hold and nearest neighbor resampling. The calculation of $\mathbf{P}_{\text{Ipoint}}(\mathbf{kT}_R)$ assumes a Poisson distribution for the irregular time instants t_i .

3.2.a Sample and Hold

Suppose that $\mathbf{x}_R(\mathbf{nT}_R)$ and $\mathbf{x}_R(\mathbf{nT}_R + \mathbf{kT}_R)$ are both coming from a single irregular observation. This means that there was no irregular observation between \mathbf{nT}_R and $\mathbf{nT}_R + \mathbf{kT}_R$. The probability of no irregular observation in the interval of length \mathbf{kT}_R is given as

$$\mathbf{P}_{\text{Ipoint}}(\mathbf{kT}_R) = \exp\left\{ -\frac{\mathbf{kT}_R}{T_0} \right\}. \tag{15}$$

3.2.b Nearest Neighbor

The basic formula is an integral over the conditional probability that both resampled values use the same original observation $x(t_i)$, given the location $t = nT_R - \tau_i$ of the original observation $x(t_i)$ that is used for $x_R(nT_R)$:

$$P_{1point}(kT_R) = \int_0^{kT_R} P_{1point}(kT_R | t) f_{t,NN}(t) dt. \quad (16)$$

Suppose that the resampled $x_R(nT_R)$ and $x_R(nT_R + kT_R)$ are both coming from a single irregular observation t_i , with $\tau_i = nT_R - t_i$. For positive τ_i , so for $t_i < nT_R$, this means that there was no irregular observation between $nT_R + \tau_i$ and $nT_R + 2kT_R + \tau_i$. The interval between $nT_R - \tau_i$ and $nT_R + \tau_i$ contains no irregular observation because the nearest observation is supposed to be at distance τ_i . So if an additional interval of length $2kT_R$ contains no irregular points, the second resampled observation uses the same irregular observation $x(t_i)$. For negative τ_i , for $nT_R < t_i < nT_R + kT_R/2$, the minimum length of the empty interval is $2kT_R + 2\tau_i$, where τ_i is between $-kT_R/2$ and 0. This means that the irregular observation used is in the interval between the resampling points. The total probability is twice this contribution due to considerations of symmetry: the observation that determines the common resampled observation is either closer to the first resampling time nT_R or closer to the second $nT_R + kT_R$.

$$P_{1point}(kT_R) = 2 \int_0^{kT_R/2} \exp\left(-\frac{2kT_R + 2t}{T_0}\right) \frac{1}{T_0} \exp\left(-\frac{2t}{T_0}\right) dt + 2 \int_0^{kT_R/2} \exp\left(-\frac{2kT_R - 2t}{T_0}\right) \frac{1}{T_0} \exp\left(-\frac{2t}{T_0}\right) dt. \quad (17)$$

The positive exponent in the first integral reflects the absolute value at negative value for τ . This gives the final result

$$P_{1point}(kT_R) = \exp\left(-\frac{kT_R}{T_0}\right) \left(1 + \frac{kT_R}{T_0}\right). \quad (18)$$

As $1 + \beta < \exp(\beta)$ for $\beta > 0$, it follows that $P_{1point}(kT_R)$ for nearest neighbor in (18) is always smaller than then the probability (15) that has been derived for sample and hold. Fig.1 shows that nearest neighbor has almost no 1-point covariance estimates beyond a shift kT_R of $3T_0$, while SH has a non-vanishing probability until $5T_0$. The probabilities can be substituted in (12) to find resampled white noise spectra. Fig.2 gives these theoretical spectra of resampled white noise. Sample and Hold is a factor 2 down at $0.16f_0$, in agreement with the result $f_0/2\pi$ of Adrian and Yao (1987). This 3 dB point is found for Nearest Neighbor Resampling at $0.22f_0$, at a frequency about $\sqrt{2}$ times higher. This shows that the intuitive attractive property of nearest neighbor resampling which has in the average a shorter distance to the true irregular observation becomes obvious in the later decay of the spectrum after resampling.

3.3 Approximation for a high irregular data rate or a low resampling frequency.

Resampling may give two conflicting contributions, especially at the high frequency end: filtering to lower the spectrum and additional noise to make it higher. Ad hoc compromises in signal processing algorithms can accidentally give seemingly perfect solutions for some specific examples. It turns out that the character of the resampled covariance (11) depends strongly on the resampling rate.

For irregular data rates much higher than the resampling rate, the probability $P_{1point}(kT_R)$ can be approximated with zero for k not equal to zero. De Waele et al (1999c,2000) showed that the resampled covariance is then given by:

$$R_{x_R}(kT_R) = S_n^2 d(kT_R) + f_{Dt} * R_x(kT_R) = S_n^2 d(kT_R) + (f_t^+ * R_x * f_t)(kT_R), \quad (19)$$

where the additional white noise term is a consequence of the fact that the variance of the observations is not influenced by resampling. Only terms coming from the same observation are used for $k=0$ and the probability that only one observation is used for covariance estimation is $P_{1point}(kT_0/\alpha)$. Theory for spectra of NN and SH, estimated from white noise

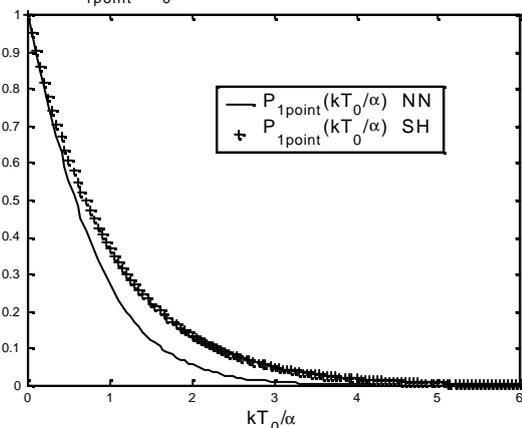


Fig.1 Probability of using the same irregular data point for both resampled points in covariance estimation, as a function of the time shift kT_R , with $T_0 = \alpha T_R$.

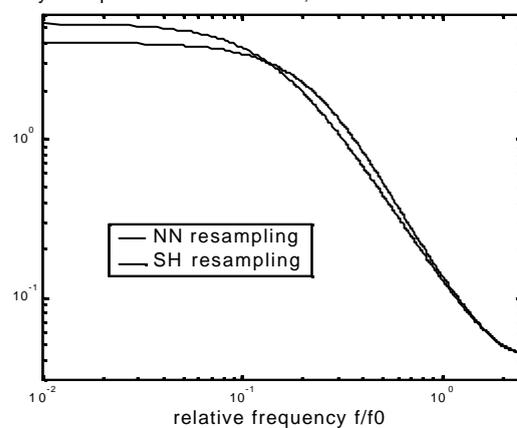


Fig.2 Theoretical spectra of resampled white noise, computed with Eq (12) for Nearest Neighbor and Sample and Hold, with $\alpha = 5$.

one original observation is used for a shift $k \neq 0$ is taken to be zero. In this case, the spectrum becomes:

$$\mathbf{h}_{x_r}(w) = s_n^2 / 2p + \hat{A}(f_t^+) \mathbf{h}_x(w) \hat{A}(f_t^-) = s_n^2 / 2p + \left| \hat{A}(f_t^-) \right|^2 \mathbf{h}_x(w) \quad (20)$$

which can be seen as a filtering of the original signal with a linear filter with impulse response f_t and an additional white noise term. De Waele and Broersen (2000) give an expression to estimate this noise level directly from the data if time intervals are Poisson distributed. It equals the average of $0.5\{x(t_{i+1})-x(t_i)\}^2$. This filtering can be compensated, as has been described by Nobach et al. (1998). Unfortunately, taking the sparsely irregularly sampled situation, the resampling period T_R will become of the same order of magnitude as T_0 or smaller and the approximation is no longer valid. The more complete formula (10) with $P_{1point}(kT_R)$ precludes the possibility to describe the influence of resampling as a simple filtering operation. Only for low resampling frequencies, the covariances of (11) and (19) are similar. For $f/f_0 > 0.5$, the spectrum keeps going down with a steep slope, also for higher relative frequencies than shown in Fig.2. The modeling of nearest neighbor and sample and hold as a combination of filtering and adding white noise is only valid for low resampling frequencies with $\alpha < 0.5$, not for the much more interesting situation with a high resampling rate.

4. BENCHMARK TESTS

Benchmark tests have been used before to compare the performance of a number of algorithms; see Benedikt, Nobach and Tropea (1998). Fig.3 gives the estimated spectrum from Nobach's (1999) benchmark, type 1 data which is a band-limited white noise signal. All figures show spectra that have been normalized to have the same surface. A comparison of Fig.2 and Fig.3 shows that the estimated NN spectrum with ARMAse1 follows its theoretical white noise prediction. It is not possible to detect the bandwidth boundary at $2f_0$ in equidistantly *resampled* data. This includes different methods that have been tried, like ARMAse1 with selected time series models, high-order time series models and several windowed periodogram based estimates. In contrast, both Lomb-Scargle and slotting show a change of level in Fig.3 at $2f_0$, so those methods recognize the bandwidth limitation. Between $2f_0$ and $2.5f_0$, the upper dark band in Fig.3 is the continuation of the Lomb-Scargle spectrum, the lower very irregular part is the presentation of the slotting spectrum, with zero introduced for the logarithm of negative spectral estimates. In this figure, the rectangular spectral Daniell window (Priestley, 1981) has been used with width $N/100$. The slotting spectrum produced many negative raw spectral values above $2f_0$, also after taking the average over $N/100$ raw periodogram estimates. Both Lomb-Scargle and slotting are too wild to allow a study of details in spectral shape or slope. Different versions of the slotting algorithm have been tested, with local normalization of Tummers and Passchier (1996) and with different definitions of $R(0)$, using the sum of squares of the N irregular observations or using some slotted variable, by including all products within the first slot. The main difference is that local normalization is somewhat less wild. A variable window has advantages if the spectrum has low power at high frequencies. In that case, the spectral window is widest at the weakest part of the spectrum, which favors the visual appearance. This property turns out to a disadvantage if the spectral window is narrow at weak parts of the spectrum.

It is more interesting for LDA purposes to observe differences in spectra with slopes instead of white noise data. Two types of data will be used: benchmark data and other simulations. The first is Nobach's (1999) last benchmark signal that provides a two-peak spectrum with a slope. Fig.4 and 5 give the results with 10000 observations, Fig.6 with 100000. The first peak is at $0.1f_0$, the second at f_0 . Fig.4 shows that 10000 observations are not enough to detect the second peak with resampled NN data. The slotted spectrum with local normalization and variable windows shows a

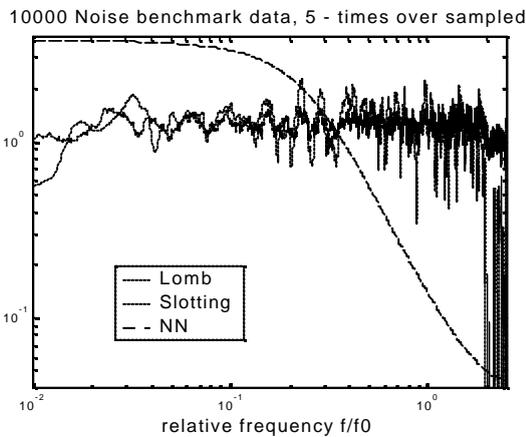


Fig.3. Spectrum of benchmark band-limited white noise with $\alpha=5$, limiting frequency of noise at $f/f_0=2$

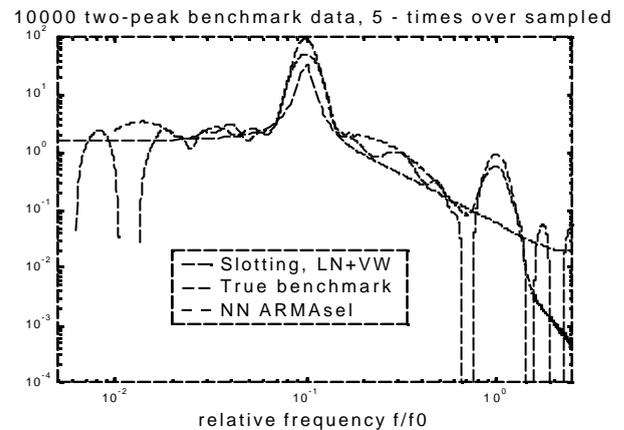


Fig.4. Two estimated spectra of two-peak benchmark data and the true spectrum that generates the data.

10000 2-peak benchmark data, 5 - times over sampled

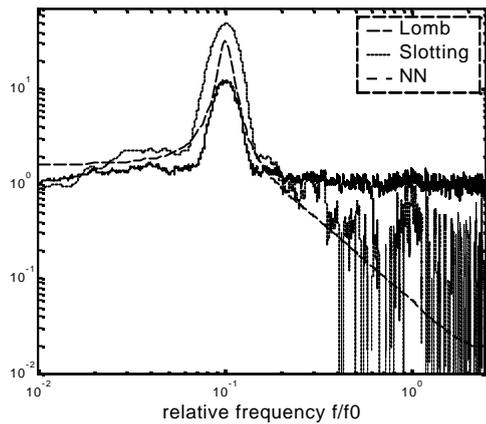


Fig.5 Spectra of two-peak benchmark data, with $\alpha=5$, 10000 observations.

100000 two-peak benchmark data, 5 - times over sampled

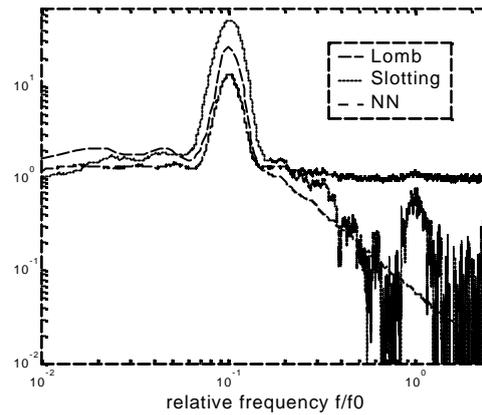


Fig.6 Spectra of two-peak benchmark data, with $\alpha=5$, 100000 observations.

relative peak at the frequency f_0 , but also one at $0.008f_0$ and $1.8f_0$. The same data of Fig.4 have been used for Fig.5. Now, the slotting spectrum has been smoothed with a Daniell window of width $N/100$. A comparison of the slotting spectra shows the enormous influence of the width of the window. Transforming the slotted covariance of 10000 benchmark data produced 40% negative spectral estimates; this is before the windowing in Fig.4 or Fig.5. In applying windowing, it is clear that it has a decisive influence whether the impossible negative spectral estimates are kept as they are, replaced by zeros or replaced by their absolute values. After variable windowing, negative spectra have been found in Fig.4 at low frequencies because of the narrow spectral window; they disappeared in Fig.5 because the spectral window was wider at low frequencies. On the high frequency end of the spectrum, the situation is reversed and the fixed Daniell bandwidth of $N/100$ is much smaller than the spectral width of the variable window in Fig.4. Hence, Fig.4 has only a couple of ripples and Fig.5 has many. It is disturbing that the choice of the width of the window has such a great influence on the appearance of the slotting spectrum, as has the treatment of “impossible” negative values. By varying the width, many different spectra can be obtained. However, it is difficult to determine whether some peak is really significant or present in the data.

In Fig.6 with 100000 observations, Lomb shows a weak ripple in the spectrum at f_0 . Slotting also gives peaks; the most significant peak is at f_0 , but many more peaks can be distinguished. The peaks at $0.45f_0$ and $0.65f_0$ are artifacts of the slotting method. It may be concluded that something can be retrieved in the slotting spectrum if a priori knowledge predicts a peak but that always many insignificant peaks will be found. Due to the fact that the raw slotting spectrum is not positive everywhere, no theoretical or statistical arguments about significance can be given. The very high quasi white noise background level of the Lomb-Scargle spectrum is remarkable. The figures 4, 5 and 6 indicate that none of the methods detects anything about the true slope at high frequencies. Furthermore, resampling with nearest neighbor (NN) followed by ARMAse1 does not detect anything at f_0 . This would not be better if periodogram based spectral estimates would be used instead of ARMAse1. The fact that the true irregular spectrum for frequencies above f_0 has no influence on the spectrum after resampling indicates that the true spectrum cannot be recovered for those higher frequencies with equidistant resampling. Hence, taking the resampling rate α greater than 2 will not give more information about the true irregular observations and $\alpha=2$ is chosen for further simulations.

5. SIMULATIONS

Simulations are indispensable for the calibration or evaluation of new spectral estimation methods. The first thing to do is to retrieve spectral properties or details in simulations, where the true process is exactly known. Only if one can retrieve these details with some processing method, it can be hoped that the method can be applied successfully to practical data. If not, it is useless to try. It is certainly *not sufficient* to show the successful operation on some practical data sets to show the applicability of a method. Nobach et al. (1998) warn for the danger of misinterpretation, because slopes in spectra are influenced by filtering characteristics that are specific for a method. It is a good practice to use simulations to calibrate methods of signal processing. Only methods that pass severe simulation tests are possible candidates to be evaluated on practical data. All real data are much more complicated than any simulation can deal with. In turbulence, the distribution of data may not be normal, the distribution of time instants will not be Poisson, the velocity bias may be present, additive measurement noise may not be white, the possibility of simultaneous particles in the measurement volume, multiple validation and many more practical deviations from theoretical models. In short, if a

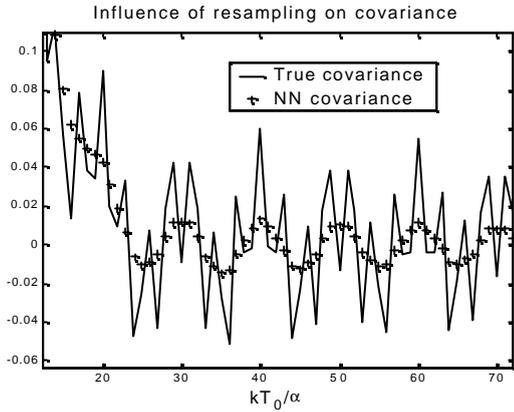


Fig.7 The influence of nearest neighbor resampling on a part of the covariance function of a simulated turbulence signal with two peaks.

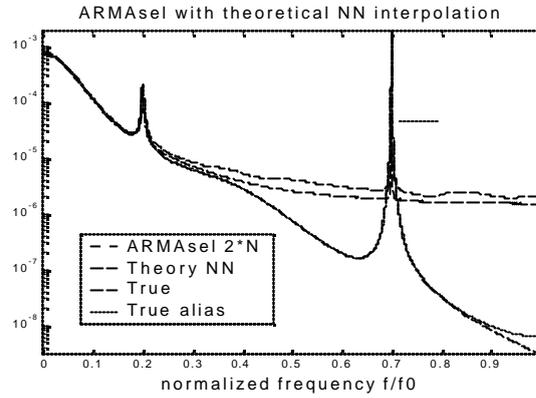


Fig.8. The true equidistant spectrum for frequencies below f_0 , the aliased spectrum obtained by sampling with $2f_0$, the theoretical NN spectrum and the ARMAsel spectrum estimated from 10000 irregular observations with $a=2$.

method passes simulation tests, it may work on practical data; if not, the method is at least unreliable. A good sequence will be: simulations, benchmark tests and real practice. Benchmark data are a subset of simulations, with predefined data. This is necessary to exclude the possibility that the simulations cover only partially the application range for which a method is proposed. The poor performance of all methods on the benchmark data in section 4 is a good reason to go back to simulations.

For simulations, non-equidistant data have been generated with a spectrum for turbulence:

$$\mathbf{h}(\omega) = \frac{\mathbf{c}}{\left(1 + \frac{\omega}{\omega_1}\right)^{5/3} \left(1 + \frac{\omega}{\omega_2}\right)^{16/3}} \quad (21)$$

with $\omega_1=0.08\pi/T_0$ and $\omega_2=0.8\pi/T_0$, so with changes of the slope at relative frequencies $0.04f_0$ and $0.4f_0$ respectively. It has two regions with different slopes in the spectrum: $h \sim \omega^{-5/3}$ and $h \sim \omega^{-7}$. The actual computer program generated 20N equidistant data with as sampling distance 0.05 times the average irregular sampling interval T_0 . An irregularly sampled set of observations has been taken by randomly selecting N numbers from a uniform distribution between 1 and 20N and keeping only the observations at the N selected time instants. With true white noise as a process, it has been verified that this irregular data generation with a hidden grid at 1/20 does not differ significantly from truly irregular data for our simulation purposes. It is possible to add separate peaks to the spectrum, with a variable frequency, bandwidth and power. In this way the capacity of methods to detect slopes and/or peaks could be studied systematically. The main questions are: how large should a peak be before it can be detected, which method is most sensitive and what is the highest frequency at which resampled spectra may contain significant information.

The first example adds two peaks, a broad one at $0.02f_0$ and a narrow peak at $0.07f_0$, both with a power of 0.04 times the power in the turbulence spectrum (21). The true covariance function of the turbulence signal given by (21) with the additional two peaks is approximated by an AR(500) process, that is used to generate the 20N equidistant observations. The number of irregular observations N is 10000. Fig.7 shows how the high frequency peak with the quickly oscillating covariance is strongly smoothed by the convolution in (10); the low frequency oscillations are almost not damped.

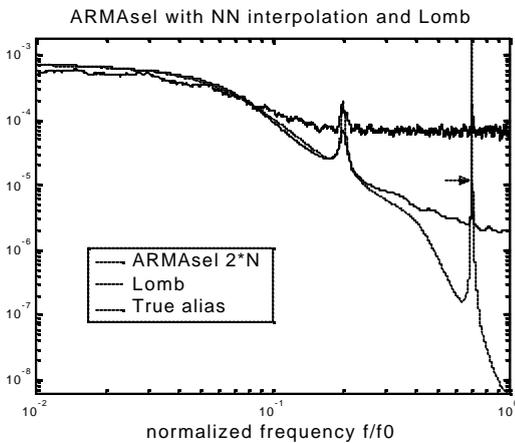


Fig.9. The same data from Fig.8 with the Lomb spectrum, on a logarithmic frequency scale, estimated from 10000 irregular observations with $a=2$.

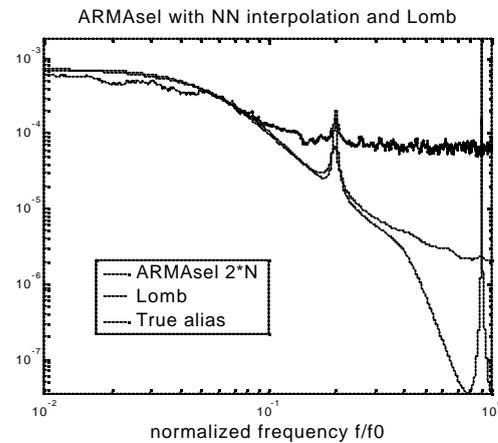


Fig.10. Simulation with second peak at $0.9f_0$, estimated from 10000 irregular observations with $a=2$.

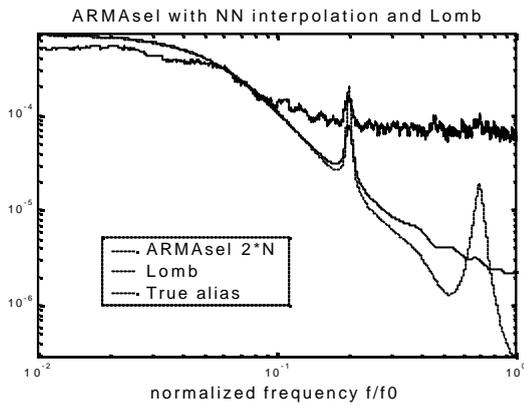


Fig.11. Simulation with second wider peak at $0.7f_0$, estimated from 10000 irregular observations with $\alpha=2$.

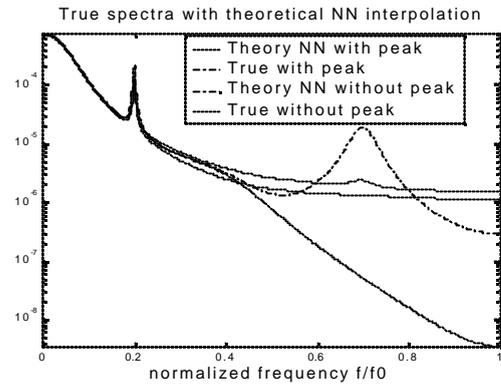


Fig.12. Theoretical difference between true and NN resampled spectra. Comparison of process with wide second peak and no second peak.

Aliasing is only seen in Fig.8 near f_0 because the generating process has a slope of $h-\omega^{-7}$ for frequencies until $10f_0$, so very little power is aliased. Fig.8 shows that NN with ARMAseI gives a good representation of the first peak at $0.2f_0$ and a reduced estimate of the second peak, according to the expectation with the theory for NN which is the spectrum calculated from the covariance obtained with (10) and the true covariance. The arrow in Fig.8 gives the almost coinciding peak values of the ARMAseI estimate and the NN theory. It is clear that the selected ARMAseI time series model closely follows the theoretical description of the influence of nearest neighbor resampling on the spectrum. This has been verified in a number of examples. In this example, the resampled spectrum gives a good estimate up to about $0.3f_0$, a little bit further than the 3 dB point at $0.22f_0$ in the white noise of Fig.2. None of the methods detects anything of the change of the true spectral slope at $0.4f_0$. However, NN resampling would detect such a change if it is at a frequency that is lower than $0.22f_0$. It has also been verified that using the ARMAseI algorithm on 2N equidistant data gives an estimate very close to the true spectrum with alias in Fig.8, so with detection of the change in slope. In other words, time series analysis with ARMAseI is very well capable to describe all spectral shapes and forms, but resampling with nearest neighbor is the cause of the difference between the resampled spectrum and the true spectrum. However, a small difference remains in Fig.8: the estimated ARMAseI spectrum at high frequencies is a fraction greater than expected with the theory; this is found in most examples. The theory is quite accurate but not yet perfect. Fig.9 additionally gives the Lomb spectrum for the same data. The ‘noise level’ is very high. Below that level, hardly anything is detected in the Lomb spectrum, but also the high peak at $0.7f_0$ is strongly attenuated, relatively seen still more than the ARMAseI peak in the resampled data. As all theoretical spectra are available as high order AR(500) processes, the model error ME of (6) can be used to compute the accuracy of estimated spectra. The ME of the ARMAseI estimate after resampling is 30000 with respect to the true process and 245 with respect to the NN theory.

Fig.10 shows that higher frequencies make an end to the possibility of NN ARMAseI to detect a peak. The peak at $0.95f_0$ is not detected by ARMAseI in this simulation run. It will sometimes be retrieved if many more observations are available, but f_0 can be considered as a soft upper boundary for the application of NN resampling with ARMAseI. The performance of Lomb-Scargle remains comparable to Fig.9, and this method can detect peaks at still much higher frequencies, *if the peak is strong and narrow*. Fig.11 has the second peak with the same power at the same location as in Fig.8 and 9, but it is made wider. Neither ARMAseI nor Lomb-Scargle can retrieve the wider peak. This shows that details in this frequency region must be extremely strong to be noticed after NN resampling. The same conclusion can be drawn from Fig.12. Here, the example with the wide peak of Fig.11 is compared with a situation where no second peak is present. The true spectra and the theoretical spectra after NN resampling are presented. It is obvious that large differences in true spectra give minor differences in the NN resampled spectra.

5. CONCLUDING REMARKS

An analysis of the covariance estimates shows that the distortion of nearest neighbor resampling becomes noticeable for frequencies above 0.44 times the Nyquist frequency, which is $\sqrt{2}$ times greater than with sample and hold. ARMAseI time series analysis can best be used to obtain the spectra after resampling. Peaks in the spectrum of irregular data can be found after resampling until a maximum frequency that is about two times the Nyquist frequency for equidistant sampling. However, peaks are attenuated much at those relatively high frequencies and slopes of spectra cannot be determined at all. So the performance is good until half the Nyquist frequency, deteriorates until twice that frequency and loses all information about the underlying data above twice the Nyquist frequency.

Lomb-Scargle can detect peaks until very high frequencies if the peaks are sharp and the noise is limited.

REFERENCES

- Adrian, R.J. and C.S. Yao, "Power Spectra of Fluid Velocities Measured by Laser Doppler Velocimetry", *Exp. Fluids*, 5, pp. 17-28, 1987.
- Benedikt, L.H., H. Nobach and C. Tropea, "Benchmark Tests for the Estimation of Power Spectra from LDA Signals", *Proc. 9th Int. Symp. On Applications of Laser Techniques to Fluid Mechanics*, Lisbon, paper 32.6, 1998.
- Britz, D. and R.A. Antonia, "A Comparison of Methods of Computing Power Spectra of LDA Signals", *Meas. Sci. Technol.*, 7, pp. 1042-1053, 1996.
- Broersen P.M.T., "Estimation of the Accuracy of Mean and Variance of Correlated Data", *IEEE Trans. On Instrumentation and Measurement*, vol. 47, pp. 1085-1091, 1998a.
- Broersen, P.M.T., "The Quality of Models for ARMA Processes", *IEEE Trans. on Signal Processing*, vol. 46, pp. 1749-1752, 1998b.
- Broersen, P.M.T., "Robust Algorithms for Time Series Models", *Proc ProRISC/IEEE CSSP Conf.*, pp. 75-82, 1998c.
- Broersen, P.M.T. and S. de Waele, "On the Accuracy of Time Series Models for Turbulence", *Proc. 8th EALA Conf. On Laser Anemometry, Advanced and Applications*, Rome, Italy, pp. 215-222, 1999.
- Broersen, P.M.T. "Facts and Fiction in Spectral Analysis of Stationary Stochastic Processes", *IEEE Trans. On Instrumentation and Measurement*, vol. 49, no. 4, August 2000.
- Gaster M. and J.B. Roberts, "The Spectral Analysis of Randomly Sampled Records by a Direct Transform", *Proc. R. Soc. London A*, vol. 354, pp. 27-58, 1977.
- Lomb, N.R., "Least- Squares Frequency Analysis of Unequally Spaced Data", *Astrophysics and Space Science*, vol. 39, pp. 447-462, 1976.
- Maanen, H.R.E. van and A. Oldenziel, "Estimation of Turbulence Power Spectra from Randomly Sampled Data by Curve-fit to the Autocorrelation Function Applied to Laser-Doppler Anemometry", *Measurement Science and Technology*, vol. 9, pp. 458-467, 1998.
- Mood, A.M., F.A. Graybill, D.C. Boes, "Introductuon to the theory of statistics", Tokyo, McGraw-Hill, 1974.
- Müller, E., H. Nobach and C. Tropea, "Model Parameter Estimation from Non-equidistant Sampled Data Sets at Low Data Rates," *Measurement Science and Technology*, vol. 9, pp. 435-441, 1998.
- Nobach, H., E. Müller and C. Tropea, "Efficient Estimation of Power Spectral Density from Laser-Doppler anemometer Data", *Exp. Fluids*, 24, pp. 499-509, 1998.
- Nobach, H., "LDA Benchmark Generator II", available on the internet, 1999.
- Priestley, M.B. "*Spectral Analysis and Time Series*", Academic Press, New York, 1981.
- Rousseeuw, P.J. and G. Molenberghs, "Transformation of non Positive Semi definite Correlation Matrices", *Commun. Statist. Theor. Meth.*, vol. 22, pp. 965-984, 1993.
- Scargle, J.D., "Studies in Astronomical Time Series Analysis. III. Fourier Transforms, Autocorrelation Functions, and Cross-Correlation Functions of Unevenly Spaced Data", *The Astrophysical Journal*, 343, pp. 874-887, 1989.
- Söderström, T. and M. Mossberg, "Performance Analysis of some Methods for Identifying Continuous-time Autoregressive Processes", *Signal Processing IX, Proc. Eusipco Conf. Rhodes, Greece*, pp. 2377-2380, 1998.
- Tropea, C., "Laser Doppler Anemometry: Recent Developments and Future Challenges", *Measurement Science and Technology*, vol. 6, pp. 605-619, 1995.
- Tummers, M.J. and D.M. Passchier, "Spectral Estimation Using a Variable Window and the Slotting Technique with Local Normalization", *Measurement Science and Technology*, vol. 7, pp. 1541-1546, 1996.
- Waele, S. de and P.M.T. Broersen, "A Time Domain Error Measure for Resampled Irregular Data", *Proc. IMTC Conf., Venice , Italy*, vol. 2, pp. 1172-1177, 1999a.
- Waele, S. de and P.M.T. Broersen, "Reliable LDA-Spectra by Resampling and ARMA-Modeling", *IEEE Trans. On Instrumentation and Measurement*, vol. 48, pp. 1117-1121, 1999b.

10th International Symposium on Application of Laser Techniques to Fluid Mechanics, Lisbon, July 10-13, 2000

Waele, S. de, W.S.J. Uijttewaai, P.M.T. Broersen and R. Booi, "Application of Time Series Analysis to Laser Doppler Anemometry Data", Proc.8th EALA Conf. On Laser Anemometry, Advanced and Applications, Rome, Italy, pp. 191-198, 1999c.

Waele, S. de and P.M.T. Broersen, "Error Measures for Resampled Irregular Data" IEEE Trans. on Instrumentation and Measurement, vol. 49, no. 2, April, 2000.