STATISTICAL ANALYSIS OF NOISE IN TOWED STREAMER ARRAYS

Charlotte Sanchis and Alfred Hanssen

Fugro Geoteam AS
Hoffsveien 1c, NO-0213 Oslo, Norway
phone: + (47) 22132321, email: c.sanchis@fugro.no

Dept. of Physics and Technology
Univ. of Tromsø, NO-9037 Tromsø, Norway
phone: + (47) 77645182, email: alfred@phys.uit.no

ABSTRACT

Towed arrays of hydrophones are commonly used in marine seismic to infer properties about the sea bottom. Normally, air guns are used as impulsive sources of pressure waves, and the hydrophone recordings are employed to estimate three dimensional maps of the geological layer structure of the subsurface sea bottom. In this paper, we present a statistical analysis of pure noise recordings, i.e., passive hydrophone recordings in the absence of impulsive sources. In particular, we show that a multitaper power spectral estimate of the noise processes yields a composite power law spectrum. Furthermore, on the basis of Parzen kernel estimates and higher-order moment analysis of the noise amplitudes, we deduce that the noise is slightly platykurtic (sub-Gaussian) and almost symmetric. These findings will have consequences for future noise reduction and signal detection algorithms.

1. INTRODUCTION

The goal of marine seismic exploration is to obtain an image of the sea bottom. In that purpose, an air gun is used as an impulsive source to generate pressure waves that travels down into the subsurface, where it is reflected. Some of the reflected energy is recorded by a long seismic streamer towed behind a vessel, together with unwanted noise and among others, weather and flow related noise. More precisely, streamers are subjected to forces from the vessel, from steering birds and from the motion of the sea. Minimizing the effects of unwanted noise is one of the main challenges faced by seismic data processors. Characterization of seismic noise has been performed in the past, with focus either on low frequencies or on hydrophone streamer noise [5]. Statistical properties of ocean noise have been studied as well, [1] and [4]. More recent marine streamers are now used in seismic acquisition. It is therefore pertinent to extend the previous results with noise characterization completed from modern equipments. Results should be used for future work on noise reduction.

Data consists in passive hydrophones recordings, i.e. in the absence of impulsive sources. It has been acquired during a Fugro campaign in the Mediterranean sea. The vessel's speed was 5 knots and weather conditions were pretty good with 1 to 1.5 meters waves. Nominal streamer depth is 8 meters. The noise data set was obtained using a solid seismic streamer which consists in a large number of hydrophones, grouped in subsections of 12.5 meters long. Within each of 12.5 meters subsections, inputs from hydrophones are summed and a trace representing recorded signal is produced and referred to as a channel. As illustrated in Figure 1, the seismic streamer is divided into 54 so-called active sections. Each active section is 150 meters long and contains 12 groups of hydrophones, providing 12 channels with noise records. Altogether, the streamer has 648 channels. There are two possible locations where extra units with depth keeping and ranging functionalities can be attached, at ~ 25 meters from the head and ~ 25 meter from the tail of each active section. Whether they cause extra noise, has to be determined.

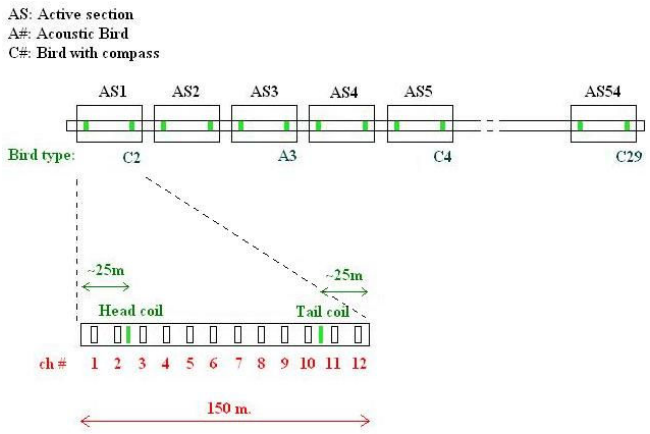

Figure 1: Streamer layout.

Noise pressure values are recorded every 2 milliseconds and for a duration of 6 seconds, simultaneously by the 648 channels. They are gathered as one noise recording. Such a recording process is repeated 10 times at approximatively 6 seconds interval, so that a total of 10 noise recordings are available for the analysis.

2. MULTITAPER SPECTRAL ANALYSIS

It is well known that the standard periodogram estimator of the power spectral density is statistically inconsistent [3]. The shortcomings of the periodogram may be mitigated by tapering in time to reduce spectral leakage, and by smoothing in frequency or time to reduce the estimator variance. The state of the art tapering technique is David J. Thomson's multitaper estimator. This estimator reduces spectral leakage subject to certain optimality criteria, while simultaneously reducing variance through an averaging procedure over an orthonormal set of tapers. A time-frequency resolution bandwidth parameter must be chosen by the user. In practice, the analysis of seismic records is not very sensi-

tive to the actual choice of the bandwidth parameter, as long as an adequate compromise between smoothing and frequency resolution is identified. We regard the multitaper estimator to be the natural choice for our kind of data, and several examples of multitaper estimates will be shown in this paper.

2.1 Brief Outline of Multitaper Method

As the multitaper method may not be well known among seismic data analysts, we will now briefly outline the estimator before actually applying it to streamer data.

The multitaper (MT) technique proposed by Thomson [8], follows ideas from Slepian [7]. This method combines the use of optimal data tapers, with averaging over a set of power spectral estimates.

2.1.1 Discrete Prolate Spheroidal Sequences

Thomson [8] proposed to apply some stringent optimality criteria when selecting data tapers. He suggested to consider tapers that maximizes the "spectral concentration", or the energy contained in the mainlobe relative to the total energy of the taper. One therefore seeks the taper $v[n]$ with a discrete Fourier transform $V(f)$, that maximizes the window energy ratio

$$\lambda = \frac{\int_{-f_B}^{f_B} |V(f)|^2 \, df}{\int_{-1/2}^{1/2} |V(f)|^2 \, df} \tag{1}$$

where f_B is the wanted resolution half-bandwidth (a design parameter) of the taper. An ideal taper would therefore have $\lambda \simeq 1$ and f_B as small as possible (but note that $f_B > 1/N$).

Expressing $V(f)$ by its discrete Fourier transform, $V(f) = \sum_{n=0}^{N-1} v[n] \exp(-j2\pi fn)$ and maximizing the above functional with respect to $v[n]$, [7] showed that the optimal taper $\mathbf{v} = [v[0], v[1], \ldots, v[N-1]]^T$ obeys the eigenvalue equation

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{v} \tag{2}$$

where the matrix $\mathbf{A}$ has elements $[\mathbf{A}]_{nm} = \sin[2\pi f_B(n-m)]/[\pi(n-m)]$, for $n, m = 0, 1, \ldots, N-1$. Note that (2) is an N-dimensional eigenvector/eigenvalue problem, thus giving N eigenvector/eigenvalue pairs, $(\mathbf{v}_k, \lambda_k)$, where $k = 0, 1, \ldots, N-1$. The interpretation is thus that we obtain a *sequence* of orthogonal tapers (eigenvectors), $\mathbf{v}_k$, each with a corresponding spectral concentration measure λ_k. The first taper $\mathbf{v}_0$ has a spectral concentration λ_0. Then, $\mathbf{v}_1$ maximizes the ratio in (1) subject to being orthogonal to $\mathbf{v}_0$, and with $\lambda_1 < \lambda_0$. Continuing, we can thus form up to N orthogonal tapers $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{N-1}$, with $0 < \lambda_{N-1} < \lambda_{N-2} < \cdots < \lambda_0 < 1$. Only tapers with $\lambda_k \simeq 1$ can be applied, since $\lambda_k \ll 1$ implies a large undesirable leakage.

It is usually safe to apply tapers up to order $k = 2Nf_B - 1$ [3].It is customary to standardize the tapers such that they are orthonormal, $\mathbf{v}_k^T \mathbf{v}_{k'} = \delta_{k,k'}$, where $\delta_{k,k'}$ is the Kronecker delta. The solutions $\mathbf{v}_k$ are referred to as "Discrete Prolate Spheroidal Sequences" (DPSS) [7]. These optimal tapers are not expressible in closed form. The eigenvalue equation (2) must therefore be regarded as the defining equation for these tapers. Recent versions of the Matlab Signal Processing Toolbox includes efficient solvers for the DPSS problem.

2.1.2 Basic Multitaper Spectral Estimators

The simplest definition of an MT estimate is simply the arithmetic average of K tapered "eigenspectra"

$$\widehat{S}_{MT}(f) = \frac{1}{K} \sum_{k=0}^{K-1} \widehat{S}_{MT}^{(k)}(f) \tag{3}$$

where the "eigenspectrum" of order k is defined by

$$\widehat{S}_{MT}^{(k)}(f) = \left| \sum_{n=0}^{N-1} v_k[n] x[n] \exp(-j2\pi fn) \right|^2 \tag{4}$$

where $v_k[n]$ denotes the elements of DPSS-taper of order k. Also data adaptive averaging schemes exist, see [8] and [3]. The adaptive averaging is necessary in several applications, and will be employed also in this paper.

The averaging of tapered spectral estimates, Eq. (3), leads to a decrease of the variance relative to any individual spectral estimates. Asymptotically,[8]

$$\mathrm{var}\{\widehat{S}_{MT}(f)\} \simeq (1/K) S^2(f), \tag{5}$$

where $S(f)$ is the true power spectrum.

Note that [2] compared the leakage, variance, and frequency resolution for the DPSS MT method with that of a standard weighted overlapped segment averaging (WOSA). He found that the MT method always performed better than the WOSA for each of the measures, when the other two measures were required to be equal for both estimation methods.

3. PARZEN AMPLITUDE ESTIMATES

The statistical distribution of the noise amplitude fluctuations i.e. the probability density function (pdf) of the measured signal amplitudes, is an important tool for characterization of the data.

The histogram is the simplest and most widely used estimator of the pdf. Despite its widespread use, the histogram has several fundamental drawbacks. Basically, the histogram amounts to counting the number of amplitude levels that falls within a specified bin $x_0 + mh \le x < x_0 + (m+1)h$, where x_0 is a chosen origin, h is a chosen bin width, and m is an integer. The histogram estimator is therefore a discontinuous and quantized piecewise constant function. Moreover, the histogram is sensitive to the choice of origin. The histogram estimator is particularly problematic when analyzing short data segments.

To overcome the many non-desirable properties of the histogram, we have chosen to estimate the probability density function by means of the so-called Parzen kernel estimator. The kernel estimator of density has good statistical properties, and it is well suited for the analysis of short data segments. Basically, the kernel

estimator is constructed by placing a smooth and symmetric normalized function (a "kernel") with its origin at each data point. By summing this collection of normalized functions, we obtain a smooth and statistically consistent estimate of the probability density.

In general, the kernel estimator of the pdf at amplitude x, given N data samples, $x_0, x_1, \ldots, x_{N-1}$ can be written as [6, 9]

$$\widehat{p}(x) = \frac{1}{N} \sum_{n=0}^{N-1} K_h (x - x_n) . \qquad (6)$$

Here, $K_h(\xi) \equiv (1/h)K(\xi/h)$, where $K(\xi)$ is the so-called smoothing kernel, and h is a scaling parameter that controls the degree of smoothing.

For Eq. (6) to be a valid estimator, one must require that $K(\xi) \geq 0$, $\forall \xi$, and $\int_{-\infty}^{\infty} K(\xi) d\xi = 1$. In addition, it is reasonable to restrict $K(\xi)$ to the class of symmetric kernels, $K(-\xi) = K(\xi)$. In practice, the resulting estimator is not very sensitive to the detailed shape of the kernel $K(\xi)$. The estimate however depends strongly on the value of the parameter h.

A good standard choice of the smoothing kernel is the Gaussian $K(\xi) = (1/\sqrt{2\pi}) \exp(-\xi^2/2)$. The choice of the smoothing parameter h is non-trivial, and several techniques exist for estimating a value of h that obeys some optimality criterion.

It is straightforward to show that the expected value of the basic kernel estimator in Eq. (6) can be written as [6, 9]

$$E\{\widehat{p}(x)\} = K_h(x) * p(x), \qquad (7)$$

where $*$ denotes the convolution operator, and $p(x)$ is the true amplitude probability density function (pdf). Thus, we may interpret the pdf estimator in Eq. (6) as a smoothed version of the true (but unknown) underlying pdf. As a consequence, the kernel estimate of the pdf does not exhibit the unphysical discontinuities present in the naive histogram based pdf estimator.

The bias of the estimator can readily be approximated by

$$b\{\widehat{p}(x)\} \equiv E\{\widehat{p}(x)\} - p(x) \approx \frac{\sigma_K^2}{2} h^2 p''(x), \qquad (8)$$

where $\sigma_K^2 \equiv \int_{-\infty}^{\infty} \xi^2 K(\xi) d\xi$. Assuming that the data are statistically independent, a useful approximation for the estimator variance is given by [6],

$$\mathrm{var}\{\widehat{p}(x)\} \approx \frac{\mathcal{E}_K}{Nh} p(x), \qquad (9)$$

where $\mathcal{E}_K \equiv \int_{-\infty}^{\infty} K^2(\xi) d\xi$ is the energy of the kernel.

3.1 Moments of the Amplitude Probability

Since the full pdf contains statistical information at a detailed level, it is often convenient to consider higher order moments of the data. The coefficient of skewness (or simply the skewness) is the dimensionless third order central moment defined by $\gamma_3 = \mu_3/\mu_2^{3/2}$. The kurtosis (or coefficient of excess) is the dimensionless fourth order central moment defined by $\gamma_4 = \mu_4/\mu_2^2 - 3$. Here, the central moment of order ν for a stochastic variable X is defined by $\mu_\nu = E\{[X - E\{X\}]^\nu\}$. Symmetric densities have $\gamma_3 = 0$, while densities skew to the right (left) have $\gamma_3 > 0$ ($\gamma_3 < 0$). The kurtosis is a flatness measure relative to the Gaussian. If $\gamma_4 < 0$, the density is termed platykurtic, and the density is flatter than the Gaussian close to the maximum.

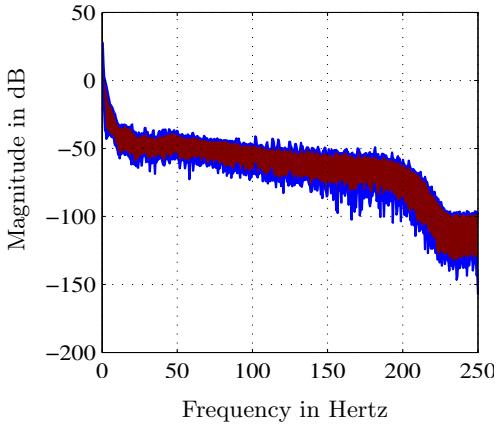

Figure 2: Envelope of the 648 power spectrum estimates.

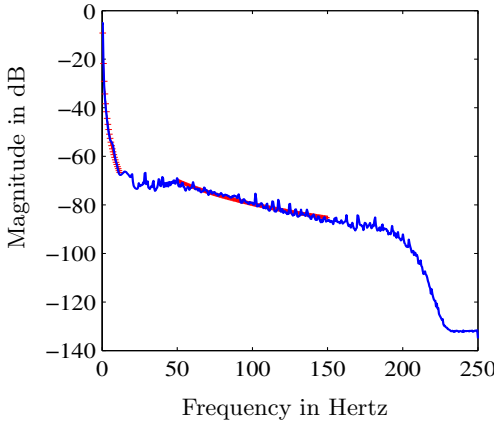

Figure 3: Averaged power spectrum estimate with linear fit of the $(0, 12]$ and $[50, 150]$ Hz intervals (red lines).

4. ANALYSIS OF SEISMIC NOISE

4.1 Spectral Analysis

The multitaper estimator is used to produce the spectrum estimate of a 6 seconds signal recorded simultaneously by 648 channels. Parameters for computation of the estimator are chosen to be: fast Fourier transform of length $N = 1024$ samples, time-bandwidth product $Nf_0 = 4$ and Thomson adaptive weighting as averaging scheme (sampling frequency F_s is 500 Hz). The power spectrum is estimated for each channel, which means that a total of 648 spectrum estimates are produced. For practical reasons, Figure 2 shows the envelope of the 648 estimated power spectra plotted on the $[0, F_s/2]$ interval. It appears that the difference in noise level is significant and can reach 50 dB re(W/Hz) at 1m. However, all power spectrum curves have the same basic shape. In order to obtain a better overview of the particular

shape of this curve, the averaged estimate is computed and shown in Figure 3. We are now able to identify 3 distinct regions in the averaged noise spectrum:

- Low frequency region: $[0, 12)$ Hz, with a magnitude attenuation of 125 dB, which contains most of the noise.
- Intermediate frequency region: $[12, 190)$ Hz, which is pretty flat and has a magnitude attenuation of 49 dB.
- High frequency region: $[190, 250]$ Hz, where we can observe the effects of the a hardware lowpass filter, with cut-off frequency at 200 Hz and attenuation of 370 dB/oct.

Where $[a, b) = \{x \in \mathbf{R} | a \leq x < b\}$.

Because seismic data of interest are normally within the $[0, 150]$ Hz interval, we choose to determine power-laws on the $(0, 12]$ Hz and the $[50, 150]$ Hz intervals. Power-laws are of the type $P(f) \sim f^p$. Generalizing slightly, we aim to perform a piecewise fit of our spectrum to the spectral shape $P(f) = 10^{p_2} f^{p_1}$, which is equivalent to

$$\log_{10}(P(f)) = p_1 \log_{10}(f) + p_2$$

where p_1 and p_2 are coefficients to be determined. Here, p_1 is the spectral decay coefficient of interest, while p_2 is a scaling parameter of little interest in the present paper. The parameter p_1 can be estimated by plotting the region of interest on log-log scale, and using basic linear fitting tools to measure the slope. With such a procedure, we get the following estimates, drawn in Figure 3 by the red curves:

- For the first region, $(0, 12]$ Hz:

$$p_1 = -4.15 \tag{10}$$

- For the second region, restricted to $[50, 150]$ Hz:

$$p_1 = -3.24 \tag{11}$$

Such spectral analysis have been repeated with other noise recordings, from the same seismic campaign and from other seismic campaigns. Similar characterization of the seismic noise has been repeatedly observed.

4.2 Amplitude Probability Density Estimates

4.2.1 Focus on a single data record

It is important to have an overview of the way the noise behaves. For convenience, we only plot a few time series from channels located far apart along the cable, see Figure 4. We choose to display channels number 1, 10, 100, 200, 300, 400, 500, 600 and 640, which spans over 8.1 kilometers. The x-axis represents the time in seconds, while the y-axis represents the noise pressure in millibars. One can observe that the very first ones (channels 1 and 10) and the last one (channel 640) are corrupted with low amplitude noise. However, all the time series present the same slow up and down fluctuations, whose amplitude can reach significant values.

To go further into the characterization of the noise amplitude, we use the Parzen kernel estimator to estimate its statistical distribution. The smoothing kernel is chosen to be the Gaussian function $K(\xi) = (1/\sqrt{2\pi}) \exp(-\xi^2/2)$ and the smoothing parameter is chosen as $h = 0.102$, as described in Section 3.

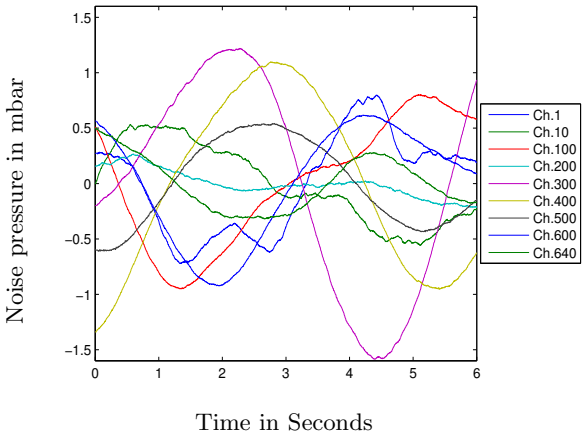

Figure 4: A few noise records with respect to time.

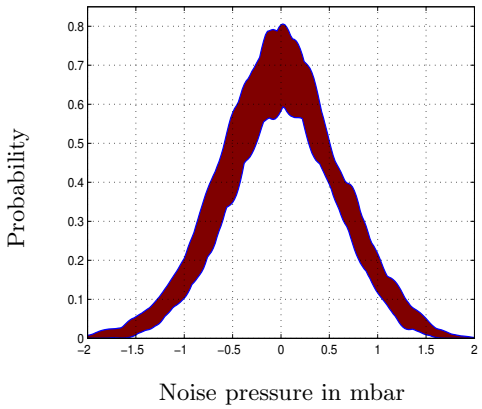

Figure 5: Envelope of the 3000 pdf estimates.

We can now estimate the distribution of the noise with respect to time. Since noise measurements are available for 3000 time samples, then 3000 probability density functions are estimated based on data recorded by 648 channels. The envelope of the 3000 probability density estimates is shown in Figure 5. If we look at distribution estimates one after the other, we observe that they change very progressively over time. So, even if differences among noise distributions are not negligible, they remain reasonable, and considering the averaged probability density function in order to characterize seismic noise distribution is acceptable. Figure 6 shows the averaged probability density function together with fitted Gaussian and t-distributions. The averaged pdf is slightly skewed to the right and none of the two standard distributions can fit with this skewness.

4.2.2 Extension to 10 data records

Estimation of the noise amplitude distribution is repeated with data recorded under the same conditions few seconds or minutes later. The smoothing parameters remain unchanged. Averaged pdfs resulting from estimations for each data record are plotted in Figure 7. All pdf curves have a similar shape, but they differ in the maximum amplitude and in their flatness.

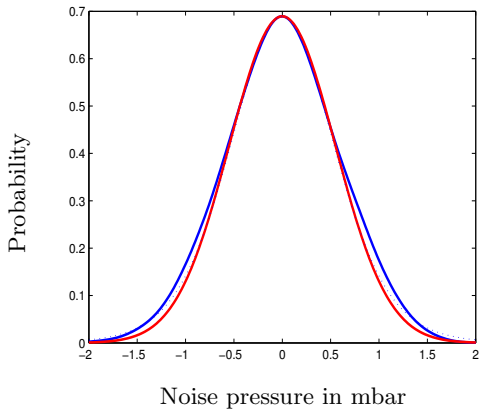

Figure 6: Averaged pdf estimate (solid thick) shown together with Gaussian fit (solid, red) and t-distribution fit (dashed line).

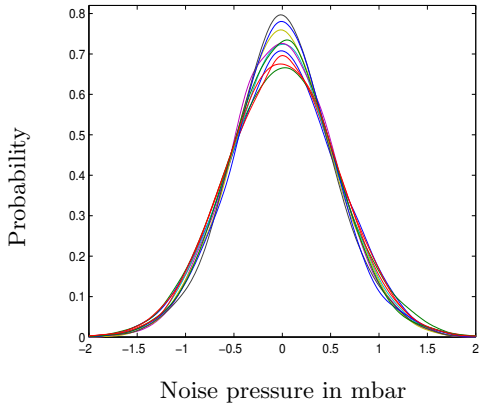

Figure 7: Overlay plot of the mean pdf of 10 noise recordings.

Skewness and kurtosis coefficients, third and fourth order central moments respectively, describe the distortion of pdfs relatively to the normal distribution. They are calculated as described in Section 3.1 for each noise record and the results are shown in Table 1. Distribution resulting from the 1^{st} record presents the biggest right-skewed value, as observed in Figure 6. Distributions resulting from the 2^{nd}, 5^{th} and 10^{th} records present negative kurtosis numbers whereas the ones resulting from the 6^{th} and 8^{th} records present small, positive kurtosis numbers. In Figure 7, they are respectively the flattest and more peaked curves of the graph. All skewness numbers fall into the $[-0.05, 0.14]$ interval, which indicates that distributions are almost symmetric. Kurtosis numbers fall into the $[-0.44, 0.15]$ interval, which means that some distributions are slightly flatter than the normal distribution. The coefficient values confirm the impression given by the figures, and therefore, they allow us to conclude that distributions of these seismic noise records are almost symmetric and slightly platykurtic.

5. CONCLUSIONS

Some pure noise recordings using solid streamers have been analyzed. The multitaper power spectral estimate of the noise processes yields a composite power law spectrum. In addition, Parzen kernel estimation and higher-order moment analysis of the noise amplitudes, shows that the noise is slightly platykurtic and almost symmetric. Such a characterization of seismic noise can be used in the future for implementation of signal detection and noise reduction algorithms.

X	Rec 1	Rec 2	Rec 3	Rec 4	Rec 5
γ_3	0.14	-0.04	0.13	-0.05	0.11
γ_4	-0.02	-0.12	-0.03	0.03	-0.20
X	Rec 6	Rec 7	Rec 8	Rec 9	Rec 10
γ_3	-0.05	0.00	-0.02	0.08	0.11
γ_4	0.11	-0.01	0.15	-0.10	-0.44

Table 1: Estimated skewness and kurtosis coefficients for 10 noise records.

6. ACKNOWLEDGMENTS

The authors thank the Research Council of Norway for support under the PETROMAKS project no. 175921/S30. We thank Peder Berentzen and Thomas Elboth for their help, and Fugro Multi Client Services for the permission to use and publish the presented datasets. A. Hanssen thanks the Research Council of Norway for support under the YFF project no. 162831/V30.

REFERENCES

[1] P.L. Brockett, M. Hinich, and G.R. Wilson, "Nonlinear and non-Gaussian ocean noise," *J. Acoust. Soc. Am.*, vol. 82, pp. 1386–1394, Oct. 1987.

[2] T.P Bronez, "On the performance advantage of multitaper spectral analysis," *IEEE Trans. Signal Proc.*, Vol. 40, pp. 2941–482, 1992.

[3] D.B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications: Multivariate and Conventional Univariate Techniques*, Cambridge: Cambridge Univ. Press, 1993.

[4] L.A. Pflug, P. Jackson, G.E. Ioup, and J.W. Ioup, "Variability in Higher Order Statistics of Measured Shallow-Water Shipping Noise," in *IEEE signal processing workshop on Higher-Order Statistics 1997*, 1997, pp. 0400.

[5] M. Schoenberger and J.F. Mifsud, "Hydrophone streamer noise," *Geophysics*, vol. 39, pp. 781–793, Dec. 1974.

[6] B.W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, 1986.

[7] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty; V: The discrete case," *Bell Syst. Tech. J.*, Vol. 5, pp. 1371–1429, 1978.

[8] D.J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, Vol. 70, pp. 1055–1096, 1982.

[9] M.P. Wand and M.C: Jones, *Kernel Smoothing*, Chapman & Hall, 1995.