Signal Detection in Underwater Sound Using Wavelets

Trevor C. BAILEY, Theofanis SAPATINAS, Kenneth J. POWELL, and Wojtek J. KRZANOWSKI

This article considers the use of wavelet methods in relation to a common signal processing problem, that of detecting transient features in sound recordings that contain interference or distortion. In this particular case, the data are various types of underwater sounds, and the objective is to detect intermittent departures (potential "signals") from the background sound environment in the data ("noise"), where the latter may itself be evolving and changing over time. We develop an adaptive model of the background interference, using recursive density estimation of the joint distribution of certain summary features of its wavelet decomposition. Observations considered to be outliers from this density estimate at any time are then flagged as potential "signals." The performance of our method is illustrated on artificial data, where a known "signal" is contaminated with simulated underwater "noise" using a range of different signal-to-noise ratios, and a "baseline" comparison is made with results obtained from a relatively unsophisticated, but commonly used, time-frequency approach. A similar comparison is then reported in relation to the more significant problem of detecting various types of dolphin sound in real conditions.

KEY WORDS: Multivariate density estimation; Segmentation; Short-time Fourier transform; Signal detection; Signal processing; Thresholding; Underwater sounds; Wavelet decomposition.

1. INTRODUCTION

For our purposes, underwater sounds can be considered to comprise acoustic events of interest superimposed on a background underwater sound environment. Throughout this article, we refer to the former as "signals," and the latter as "noise," although it should be appreciated that in doing so, we use the terms rather loosely.

In the search for signals, we have some a priori knowledge that allows us to focus attention on certain frequency bands, but otherwise the nature of signals is ill-defined, and they are perhaps best thought of as "transient features of potential interest." In particular, we do not have access to obvious models for, nor any reference library of, the signals that we wish to detect.

At the same time, the "noise" in our data will rarely conform to that implied by conventional statistical uses of that term. In general, the background continuum will contain features that depend on particular underwater conditions and on the recording apparatus used; furthermore, these features may change and evolve during any recording. Thus no obvious distributional assumptions can be made to model such noise, and its nature must be inferred from initial periods in any particular recording, when we can assume a priori that no signals of interest are present. From these periods, we must develop an empirical model of the noise and then allow this model to adapt with time and recording conditions as appropriate, so that we can use it as a baseline at any time for the identification of signals; that is, significant departures from it.

Our primary interest here is simply in detection (i.e., in the determination of where particular signals begin and end, sometimes referred to as "segmentation"), and we restrict ourselves solely to that objective. In passing, however, it should be noted that as well as having importance in its own right, segmentation also is often a necessary prerequisite to discrimination between different types of signals. The work reported here thus may have relevance to studies concerned with classification of sonar and other high-dimensional signals, either where neural-networks are used (see, e.g., Gorman and Sejnowski 1988; Smith, Bailey, and Munford 1993); or where wavelets are used as a feature extraction method for discrimination or classification (e.g., Coifman and Saito 1994; Learned and Wilsky 1995; Telfer et al., 1994).

2. UNDERWATER SOUND DATA

Our available data relate to underwater sounds of different types, including those produced by dolphins, shrimps, seals, whales, ice breaking, and others. They consist of selected extracts from lengthy recordings taken in real life conditions in the ocean, using a hydrophone placed several meters beneath sea level. The recording apparatus sampled the sound at 40.96 kHz with 16-bit resolution (near-CD quality), and thus the volume of raw data is considerable. The sampling rate is over twice the highest frequency of the signals that we potentially wish to identify, so there are few aliasing problems.

Figure 1 represents a subset of the raw data consisting of approximately 33,000 data points with values (amplitude) between $-32{,}768$ and $+32{,}767$ (16 bit). The section of data portrayed corresponds to dolphin sound, and principally relates to a "whistle," although there is also a secondary series of short term "clicks" toward the end of the depicted

Wojtek Krzanowski is Professor, Trevor Bailey is Senior Lecturer, and Kenneth Powell is a research student, Department of Mathematical Statistics & Operational Research, University of Exeter, EX4 4QE, U.K. Theofanis Sapatinas is Research Associate, Department of Mathematics, University of Bristol, BS8 1TW, U.K. This research was supported through a grant from the Defence Research Agency (DRA), Winfrith Technology Centre, Dorset, U.K. (Agreement No. 2041/045) and also through a Co-operative Award in Science and Engineering (CASE) studentship funded jointly by the Engineering and Physical Sciences Research Council (EPSRC) and DRA, Winfrith (No. 9350006X). Partial support for one of the authors is also acknowledged through EPSRC grant. GR/K/70236. The underwater sound data were made available by David J. Smith of DRA, Winfrith, who provided useful advice throughout the project. The authors also particularly thank the associate editor and referees for numerous helpful suggestions for improvements to the article.

Journal of the American Statistical Association
March 1998, Vol. 93, No. 441, Applications and Case Studies

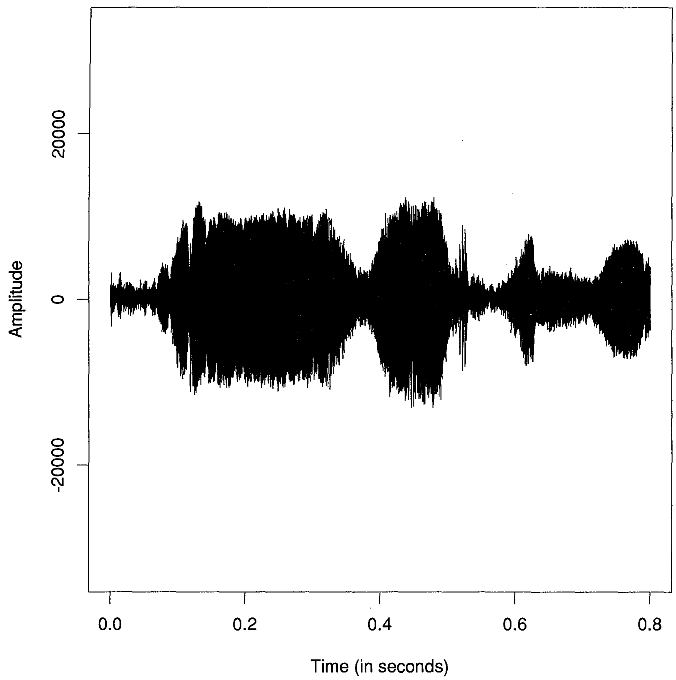

Figure 1. An Example of Dolphin Sound.

.8-second time period. In this example relatively little noise is present; however, many of our samples contain levels of background interference that make even the most dominant features, such as loud whistles, difficult to distinguish by the human ear.

Analysis of such underwater sounds has some similarities with work directed towards human speech recognition. However, as mentioned elsewhere (see, e.g., Creasey, Smith, and Gazey 1989; Powell, Sapatinas, Bailey, and Krzanowski 1995), the special recording circumstances and the nature of the sounds themselves give rise to unique features not found with the more commonly analyzed speech data. For example, underwater data have a larger amount of background interference than is generally the case in speech recognition work, and the structure of this noise is potentially more complex. Also in speech data, single frequencies tend to predominate, biased toward the lower end of the frequency range up to 20 kHz, while some of the underwater signals of interest to us are spread across a wider range of frequencies toward the upper end of this band. Finally, valuable a priori knowledge that can be brought to bear in the analysis of human speech is not available for underwater sounds; for example, we do not have a reference library of phenotypes to assist in the analysis.

3. GENERAL APPROACH

Historically, several approaches have been used for signal detection in underwater and other "noisy" sound data, either by working directly with the raw sound or by using some transformation of it.

Clearly, many mainstream developments in digital signal processing and filtering techniques are relevant to signal detection, and numerous texts cover that general field (e.g., Ludeman 1986; Smith and Mersereau 1992). Additional methods that have been used include some based on techniques used in multivariate statistical process control (e.g., Wierda 1994); others that use techniques for detecting outliers in multivariate time series (e.g., Khattree and Naik 1987); and others that use methods arising from seismic signal analysis and image processing applications (e.g., Chen 1985; Pal and Pal 1993).

Many of the various methods for the analysis and detection of sound signals make use of the Fourier transform. Although this transform is extremely useful and well established, it does have drawbacks—principally difficulties in analyzing short-term transient sound behavior. Various short-time Fourier transforms (STFT), using a variety of "windows" with different relative advantages, have been developed to address this problem. In addition, alternatives to the STFT with better time-frequency localization have been suggested; for example, the Wigner distribution and its variants.

These various time-frequency approaches have been extensively reviewed (e.g., Boashash 1990; Cohen 1989; Jones and Parks 1992), and we do not discuss them further here. Instead, we concentrate on exploring uses of the wavelet transformation in signal detection. Our justification for wishing to experiment with this alternative kind of transformation is simply that signals with very short time duration are frequently those of most interest in our particular underwater sound data. The resolution of the wavelet transformation has local adaptivity, and this potentially enables it to "zoom in on" irregularities and characterize them more specifically than is possible with alternative transformations. A good account of the relative advantages of wavelet versus Fourier transforms has been given by Strang (1993).

Some of the potential uses of wavelets for statistical problems have been recently discussed and developed by Donoho (1993), Donoho and Johnstone (1994, 1995), and Donoho, Johnstone, Kerkyacharian, and Picard (1995). They have demonstrated how, by developing appropriate thresholds on the coefficients resulting from a wavelet decomposition, a function of unknown smoothness can be recovered from sampled data contaminated with white noise. Subsequently, their ideas have been used by various researchers to uncover useful structural information from complex noisy datasets. For instance, Wang (1995) has applied these techniques in the detection of data "jumps" and "sharp cusps," and Ramsey, Usikov, and Zaslavsky (1995), have applied wavelet decompositions to look for structure in the U.S. stock market price indices. Recently, a variety of further thresholding rules has been suggested based on, for example, cross-validation methods, multiple-hypothesis testing procedures, and Bayesian approaches (e.g., Abramovich and Benjamini 1996; Abramovich, Sapatinas, and Silverman 1996; Nason 1995, 1996; Neumann and Spokoiny 1995; Ogden and Parzen 1996; Wang 1996; Weyrich and Warhola 1995). In addition, Johnstone and Silverman (1997) have developed a "level-dependent" threshold approach for data with correlated noise, and some of the previous methods mentioned can also be extended to this case.

Although these various thresholding methods are interesting, they may not necessarily be the most relevant wavelet approach in our signal detection problem. Such thresholding techniques are essentially concerned with the recovery of "smooth" features against a background which is often assumed to be "white" or "colored" noise. In our application the signals of interest are far from smooth and, as mentioned earlier, the real-life background underwater environment may well evolve over time and include features arising from the recording apparatus—it cannot necessarily be equated with random error, whether this be correlated or otherwise. Such arguments are supported by Powell et al. (1995), who applied a number of such thresholding methods to underwater sound data and showed them to have deficiencies in relation to feature detection; in particular, some higher-frequency signals of interest were often "thresholded-out."

Recent work that is more relevant to the application here involves the use of wavelet thresholding in postprocessing to find adaptive segmentations of a possibly nonstationary stochastic signal, based on local trigonometric or time-scale decompositions. For example, von Sachs, Nason, and Kroisandt (1996) have used nonlinear wavelet shrinkage for smoothing to estimate the evolutionary wavelet spectrum consistently. Their work parallels in the wavelet domain that of Donoho, Mallat and von Sachs (1996) based on local cosine packets to estimate the covariance of a locally stationary process, and of Neumann and von Sachs (1997) based on tensor wavelet thresholding to smooth pseudo–Wigner–Ville time-frequency distributions.

Such segmentation approaches have considerable potential, but all involve predefined models that are essentially designed for specific situations, although in some cases they allow for a fairly broad class of signals. In this article we prefer to approach the problem from an alternative direction. Instead of imposing an a priori model, we attempt to build an adaptive model of the background continuum using density estimation of the joint distribution of certain summary features of unthresholded wavelet decompositions. Observations considered to be outliers from this density estimate are then flagged as signals.

This approach has some similarities to methods that have been applied to data from an entirely different arena, and for a different purpose, by Nason and Silverman (1995). These authors used the "stationary" wavelet transform and then applied kernel smoothing to the coefficient behavior in each resolution level to produce useful analyses of datasets from astronomy and veterinary anatomy. Their approach was applied to individual coefficient values within separate levels of the stationary wavelet decomposition, while we concentrate here on summary features of the coefficients and on exploiting their multivariate behavior both across and within levels of the discrete wavelet transform.

The article is organized as follows. In Section 4 we very briefly summarize some relevant aspects of wavelet theory. In Section 5 we suggest summary features of wavelet decompositions of our underwater sounds that appear to exhibit distinctive multivariate behavior when signals occur. In Section 6 we discuss density estimation of the joint distribution of these summary measures during periods of noise and propose a method for detecting multivariate outliers from this estimated distribution. We then apply these ideas in Section 7 to detect signals in artificial and real sounds of different types and give a "baseline" comparison with results obtained from the use of a straightforward time-frequency approach. Finally, in Section 8 we summarize and discuss some implications of these results.

4. THE WAVELET DECOMPOSITION

One way to view wavelets is as orthonormal basis functions for various function spaces (although strictly they need not be orthonormal). With a wavelet series expansion, the basis functions are all dilations and translations of a single function referred to as the "mother wavelet" and denoted by ψ. The dilation and translations of the mother wavelet are given by:

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), \qquad j,k \in \mathbb{Z},$$

where j is the dilation factor and k is the translation factor. For suitable choices of ψ, the set of functions $\{\psi_{j,k}\}$ forms an orthonormal basis for $L^2(\mathbb{R})$. Wavelets can also form unconditional bases for other function spaces, such as Besov or Triebel spaces (see, e.g., Meyer 1992).

For a function $f \in L^2(\mathbb{R})$, the wavelet series representation is

$$f(x) = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} w_{j,k}\psi_{j,k}(x),$$

where the wavelet coefficients $w_{j,k}$ are given by

$$w_{j,k} = \int_{\mathbb{R}} f(x)\psi_{j,k}(x)\, dx.$$

Intuitively, $\psi_{j,k}$ represent "smooth wiggly functions." In contrast to standard Fourier sine and cosine series, wavelets are well localized in time (via translations) and in frequency/scale (via dilations).

One commonly used series of mother wavelets was constructed by Daubechies (1988), with each ψ in the series indexed by N. These wavelets are compactly supported in the time domain (0 outside a finite interval), and have high regularity (ψ and all its derivatives up to some order are continuous); the regularity is proportional to the index N. Such wavelets have another interesting property, that of vanishing moments. We say that a mother wavelet ψ has m vanishing moments if

$$\int_{\mathbb{R}} x^k\psi(x)\, dx = 0, \qquad k = 0,\ldots,m-1.$$

This property allows the use of wavelets for compression techniques, because they provide sparse representations of functions. Daubechies (1992, pp. 242–244) has showed that this property ensures that the fine-scale wavelet coefficients will be large only where a function or its derivatives have singularities.

The expansion of a function $f \in L^2(\mathbb{R})$ given earlier has a discrete analog for a dataset $f_1, f_2, \ldots, f_n$, where $n = 2^J$

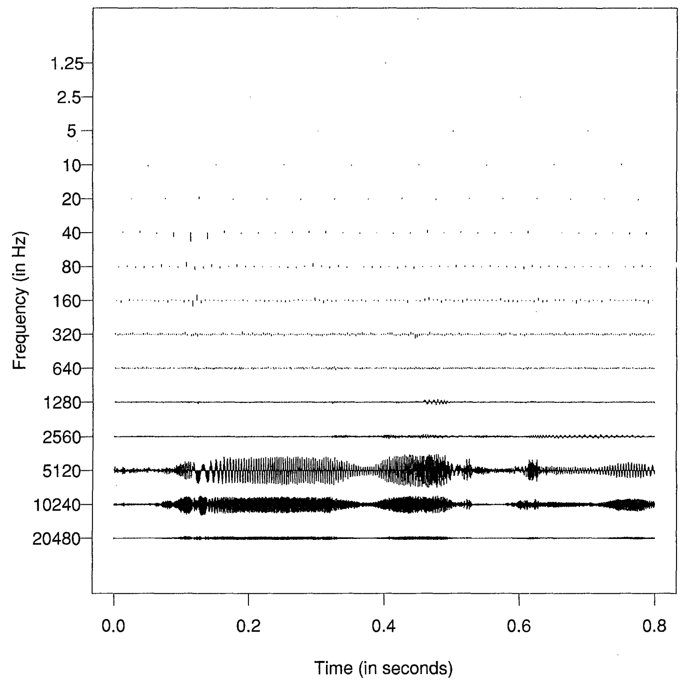

Figure 2. Discrete Wavelet Coefficients, $d_{j,k}$, for the Dolphin Sound Example Shown in Figure 1, Using Daubechies's Extremal Phase Wavelet With N = 8.

for some J. This discrete wavelet transform (DWT) is performed using an efficient algorithm that requires only $O(n)$ operations (see Mallat 1989) and results in $n-1$ discrete wavelet coefficients $d_{j,k}$, where $j = 0, 1, \ldots, J-1$ and $k = 0, 1, \ldots, 2^j - 1$, and one scaling coefficient labeled as $c_{0,0}$. Each $d_{j,k}$ describes the contribution around temporal location $2^{-j}k$ with frequency proportional to 2^j, whereas $c_{0,0}$ is a weighted total of all data.

All the subsequent wavelet analyses in this article use the `WaveThresh` package developed within S-PLUS by Nason (1993) and described by Nason and Silverman (1994); this package incorporates a DWT that handles periodic boundary conditions. Throughout the applications here, we used the Daubechies extremal phase wavelet with $N = 8$ (Daubechies 1992, table 6.1, p. 195); however, the methods we suggest could be used with alternative wavelet families, if required. Figure 2 shows the wavelet decomposition of the dolphin sound presented in Figure 1 (involving $2^{15} = 32{,}768$ data points) based on the Daubechies wavelet mentioned earlier. `WaveThresh` displays $d_{j,k}$ in the form of a "pyramid." There are 2^{15-1} coefficients at the highest frequency level (20.48 kHz), which corresponds to the bottom, or lowest level, of the pyramid ($j = 14$); there are then 2^{15-2} coefficients at the second highest frequency, and so on upward through the pyramid.

5. WAVELET DECOMPOSITIONS AND SIGNAL DETECTION

We now consider how we might summarize a wavelet decomposition, such as that pictured in Figure 2, in a way that lends itself to use in signal detection.

A preliminary consideration is the choice of an appropriate time scale, or time "window," over which to effect any summarization. Our recording apparatus sampled data at 40.96 kHz, and thus a window containing 128 coefficients at the lowest level of the decomposition (256 data points) corresponds to time period of .00625 seconds. The shortest signals of interest in our applications have a minimum duration of around .01 seconds, and so this window was considered to be easily sufficient to capture significant changes in the structure of the sound.

Several summary characteristics of the wavelet decomposition within such temporal windows may be relevant to signal detection. The summarization that we adopt here draws on the fact that the total sum of squares of the original data values is preserved in the sum of squares of the coefficients in the wavelet decomposition. If the former is loosely interpreted as an overall measure of the "energy" in the raw sound, then the sum of squares of the coefficients in a particular level of the decomposition can be thought of as relating to "energy at a particular frequency." Thus if we consider within each temporal window a set of variables, each of which is based on the sum of squares of wavelet coefficients in a different level, then we are effectively looking, within small time frames, at the energy levels in the original sound at different frequencies.

More specifically, we first subdivided any given underwater sound recording into periods of .8 s (with each period containing $2^{15} = 32{,}768$ data points). We chose these fairly long periods to provide a compromise between computational convenience and the need to minimize edge effects when summaries of the decompositions for each segment were subsequently reassembled. Each of these .8-second segments was subjected to a wavelet analysis, and each resulting decomposition was split into 128 time windows each with a length of 128 coefficients at the lowest level of the decomposition. The wavelet coefficients in levels covering the maximal frequency range that would ever be of interest to us (i.e., .1–20 kHz) were then summarized in terms of eight mean sums of squares, $(x_{t,1}, \ldots, x_{t,8})$, for each time window t, so that

$$x_{t,j-6} = \frac{\sum_{k=1}^{2^{j-7}} d_{j,k}^2}{2^{j-7}}, \qquad j = 7, \ldots, 14,$$

where $d_{j,k}^2$ are taken from time window t.

Observations on these variables for the 128 time windows in each of the separate, successive .8-s wavelet decompositions were then assembled sequentially to produce a continuous set of observations referring to time windows covering the entire duration of the original recording. Essentially this may be viewed as similar to what others have referred to as a "scalogram," or "time marginal of scalograms $d_{j,k}^2$," in the literature (see, e.g., Abry, Goncalves, and Flandrin 1995).

For our particular application, mean sums of squares corresponding to lower frequencies were actually of little interest; a priori the signals with which we are concerned are in the higher frequency ranges, and our focus in most of the later discussion will actually be on the behavior of $(x_{t,5}, x_{t,6}, x_{t,7})$, which cover the range 2–10 kHz. Subsequently, we use $\mathbf{x}_t = (x_{t,5}, x_{t,6}, x_{t,7})$ to denote a vector of multivariate observations on these three variables in time window t. In other applications, where a different frequency range is important, fewer or more of the mean sum of squares may need to be considered, and the definition of $\mathbf{x}_t$ can then be adjusted accordingly.

To illustrate the relevance of $\mathbf{x}_t = (x_{t,5}, x_{t,6}, x_{t,7})$ to signal detection in our particular case, we compare, in Figures 3 and 4, its behavior during typical dolphin signals (Fig. 3) and during a period of background interference when no 'signals' are present (Fig. 4). Each figure relates to a period of .8 s, and both are taken from the same original recording. The dolphin signals contained in the first case are a whistle (narrowband and relatively long time duration), followed by a short series of clicks (broadband and very short time duration). The two plots have been standardized so that the total energy across all frequencies is the same in each of the .8-s periods. Informally, then, we are comparing sounds of equal "volume."

It is clear from these plots that the selected mean sums of squares jointly behave in a significantly different way during the dolphin signals than they do during background interference with the same total energy when no signals are present. Moreover, although we do not reproduce further plots here, similar patterns in these variables tend to be repeated in signals of the same general type. Notice however, that it is the *multivariate* behavior of the variables that is important—in some cases, the behavior of any of $x_{t,5}, x_{t,6}$, or $x_{t,7}$ taken singly may remain quite similar during signal and noise phases, but their joint behavior is distinctive.

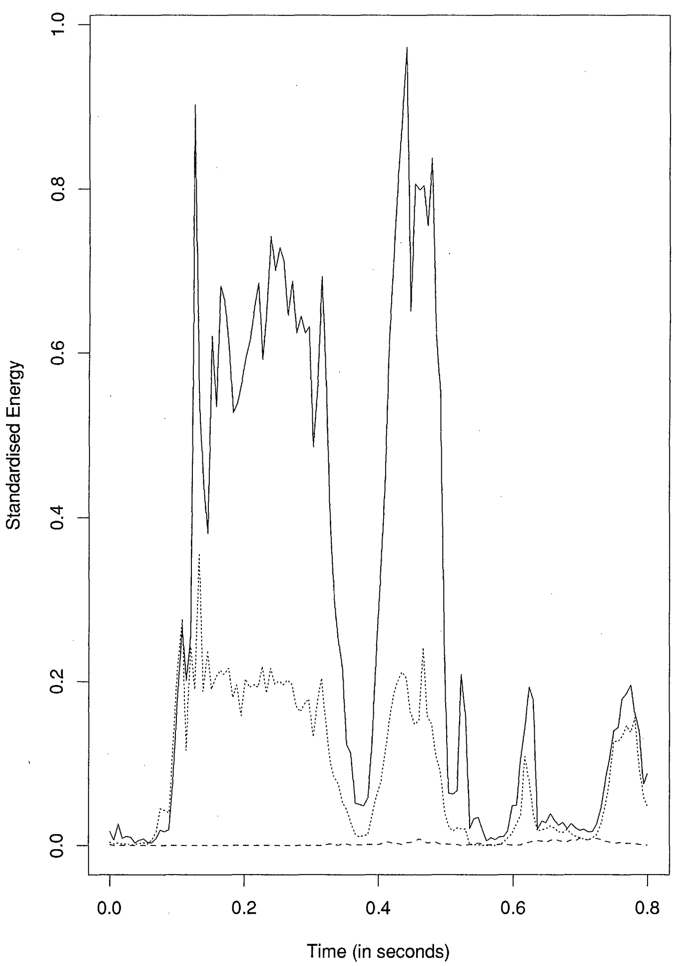

Figure 3. Behavior of $\mathbf{x}_t$ for Typical Dolphin Signals. – – –, $x_{t,5}$; ——, $x_{t,6}$; and · · ·, $x_{t,7}$.

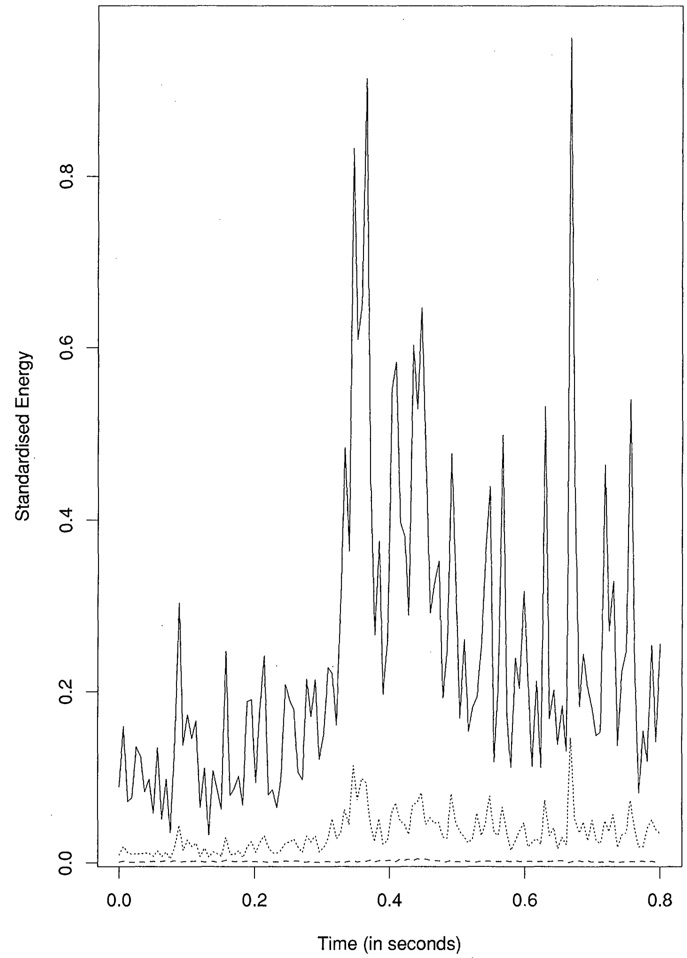

Figure 4. Behavior of $\mathbf{x}_t$ for Typical Background Interference. – – –, $x_{t,5}$; ——, $x_{t,6}$; and · · ·, $x_{t,7}$.

6. A SIGNAL DETECTION TEST

The results of the previous section suggest that signal detection might be achieved through formal detection of changes in the joint behavior of the mean sums of squares described there. This in turn requires a reference model of their joint distribution during periods of background interference and, in addition, a method of detecting multivariate outliers from this distribution.

If one were able to make some simple assumption about the joint distribution of the mean sums of squares in the case of such noise (e.g., multivariate normality), then one could simply estimate the parameters of this distribution from a period of known noise, and various techniques would be readily available to detect single multivariate outliers, or signals—for example, the use of Mahalanobis distance as recently described by Penny (1996). However, the possibly complex and changing nature of the background underwater continuum precludes any such simple assumptions and suggests taking a nonparametric and flexible data-dependent approach to the estimation of the joint distribution.

The idea proposed here is to develop an initial multivariate density estimate of the joint distribution of the mean sums of squares from a sufficiently long period of known background interference or noise. Then each successive new observation on the variables will be used to update the current density estimate of the joint distribution, unless it is detected as a potential outlier from this distribution, in which case it will be flagged as part of a signal and then ignored. The noise model thus adapts over time, thereby reflecting any changes in underwater background environment that are occurring.

To obtain density estimates, we used a multivariate kernel density estimator with a multivariate normal kernel of the form given by Silverman (1986, p. 78). Recall that our practical focus here is on a trivariate distribution for $\mathbf{x}_t = (x_{t,5}, x_{t,6}, x_{t,7})$, where t refers to the temporal window, so that, given data $\mathbf{x}_1, \ldots, \mathbf{x}_T$, the specific estimator used was

$$\hat{f}(\mathbf{x}) = \frac{(\det S)^{-1/2}}{(2\pi)^{3/2}Th^3} \sum_{t=1}^{T} \exp\left(-\frac{(\mathbf{x}-\mathbf{x}_t)' S^{-1} (\mathbf{x}-\mathbf{x}_t)}{2h^2} \right),$$

where S is a robust estimate of the covariance matrix (see, e.g., Tukey and Tukey 1981), and h is a suitable global bandwidth chosen on the scale of standardized data rather than original data. Optimal choice of h was determined by simple cross-validation, maximizing the approximate log-likelihood for the data, $\sum_{t=1}^{T} \log(\hat{f}_t(\mathbf{x}_t))$, where $\hat{f}_t$, $t = 1, \ldots, T$ refers to the kernel density estimate based on all the observations $\mathbf{x}_1, \ldots, \mathbf{x}_T$, except $\mathbf{x}_t$ itself. All these methods are well established (for more details and bibliographies, see, e.g., Silverman 1986 or Wand and Jones 1995). More sophisticated methods for bandwidth selection, such as generalized cross-validation, could be adopted if required, but the straightforward approach worked sufficiently well for our purposes. We also assume here that $\mathbf{x}_1, \ldots, \mathbf{x}_T$ are serially independent or at least have only restricted dependency. Without this assumption, the accuracy of the kernel estimate obtained is clearly questionable and this will have a corresponding adverse effect upon the signal detection method described in subsequent sections. Although wavelet coefficients at any particular scale are not necessarily serially uncorrelated, it has been demonstrated that the orthogonal wavelet transform does have a strong decorrelation property when applied to data with correlated noise (see, e.g., Johnstone and Silverman 1997). In addition, each $\mathbf{x}_t$ is a summary of several wavelet coefficients within a time window, so that with the values of T adopted here, the sample $\mathbf{x}_1, \ldots, \mathbf{x}_T$, actually refers to runs of several thousand individual wavelet coefficients. The possibility of serial dependence thus does not present a major practical difficulty in our case.

The previously described kernel estimator was first applied to an initial sample, $\mathbf{x}_1, \ldots, \mathbf{x}_T$, taken from the beginning of the sound recording that was known a priori to consist of pure background interference with no signals of interest. The way in which our recordings were obtained ensured that lengthy periods of this nature were always available. The duration, T, of the initial sample must clearly be sufficiently large enough to form a reasonably stable kernel estimate of the joint distribution, and as a basic guide we used double the minimum theoretical sample size suggested by Silverman (1986, p. 94) to ensure acceptable mean squared error at the origin, when estimating a standard three-dimensional normal density using a normal kernel and a bandwidth that minimizes mean squared error at the origin. But other, separate considerations also governed our choice of T, and we return to this point later.

Once the initial density estimate was established, we then successively considered subsequent observed vectors of the selected mean sums of squares in the sound recording. Denote the new vector at any stage as $\mathbf{x}$, and let $\mathbf{x}_t$, $t = 1, \ldots, T$ be the set of most recent preceding points identified as noise; that is, those on which at any time the most recent kernel estimate of the noise distribution is based. We then identify $\mathbf{x}$ as an outlier from the current kernel density estimate if it produces a significantly small significance level, the latter being given by

$$\frac{\#\{\hat{f}_t(\mathbf{x}_t) \leq \hat{f}(\mathbf{x})\} + 1}{T+1},$$

where "$\#\{\cdot\}$" means "the number of" for $t = 1, \ldots, T$; $\hat{f}$ refers to the kernel density estimate based on $\mathbf{x}_1, \ldots, \mathbf{x}_T$; and $\hat{f}_t$ refers to the kernel density estimate based on $\mathbf{x}_1, \ldots, \mathbf{x}_T$ but excluding $\mathbf{x}_t$.

This detection criterion, or test, represents a simple comparison of $\hat{f}(\mathbf{x})$ with T ordered examples of effectively the same quantity for points known to be noise, with these examples obtained by cross-validation on $\mathbf{x}_1, \ldots, \mathbf{x}_T$. Its structure resembles tests which are common in resampling methodology and it has some similarities to those adopted in kernel discrimination methods (see, e.g., Remme, Habbema, and Hermans 1980). The minimum possible empirical significance level is clearly determined by the value of T; the power of the test would tend to increase with T. A judicious choice of T thus allows a degree of fine tuning to particular circumstances whereby a lower bound is established through the required significance level (given that this exceeds the level required for a stable density estimate; see earlier), and then a trade-off is established between a value sufficiently large to ensure adequate power and a value sufficiently small to maintain the time localization of the current density estimate. Note that the $\hat{f}_t$, $t = 1, \ldots, T$ are automatically available from the last iteration of the most recent cross-validation used to optimize kernel bandwidth, and so little extra computation is involved in applying the test.

If $\mathbf{x}$ is identified as an outlier by this detection test, then it is flagged as a part of a signal and subsequently ignored. Otherwise, $\mathbf{x}_t$ is replaced by $\mathbf{x}_{t+1}$, $t = 1, \ldots, T-1$, $\mathbf{x}$ replaces $\mathbf{x}_T$, and then the optimal bandwidth and the kernel density are updated before the entire process is repeated for the next $\mathbf{x}$. Thus the kernel density estimate and the critical region of the detection test develop recursively over time in response to any slow changes in underwater background environment that might be occurring. In this sense the approach addresses similar issues as methods that incorporate explicit models to allow for a slowly varying time change in an otherwise almost stationary noise component (see, e.g., Dahlhaus 1997; Donoho et al. 1996).

If the nature of the signals being sought requires it, further adjustment to this procedure is available through the introduction of an appropriate "lag" into the updating of the kernel estimate. Under such a scheme, new observations are successively tested but subsequently placed into a "buffer" of appropriate length, without updating the kernel. When the buffer is full, the set $\mathbf{x}_t$, $t = 1, \ldots, T$, and the kernel are updated using all vectors in the buffer identified as noise. The buffer is then emptied. In this way signal detection is always based on a slightly delayed background environment. Such a lag will clearly reduce the computation involved in updating the kernel after each new vector that is not classified as signal. However, the more significant motivation for a lag is to help to reduce the misclassification probability of the current vector being increased by a preceding signal in the immediate past being wrongly classified as noise when detecting certain signals (such as dolphin "whistles") characterized by a slow build-up.

7. APPLICATIONS

Here we examine the results when our signal detection method is applied to artificial data consisting of known signals contaminated with artificial underwater noise using a range of signal-to-noise ratios (SNRs) and compare these results to those obtained from an alternative and more conventional time-frequency approach. We do this simply to establish that our method is capable of reasonable recovery of known signals in well-understood circumstances, rather than to draw any general conclusions about performance. This established, we then report the results of applying the method to a substantial time period of "noisy" real underwater sound involving dolphin "clicks" and "whistles" and again compare the results to those obtained from the time-frequency approach.

In the comparisons here we use a straightforward time-frequency approach based on the STFT. As discussed earlier, there exist a range of more sophisticated competing model-based time-frequency or time-scale approaches for segmentation of a possibly nonstationary stochastic signal. Indeed, we could even experiment with a segmentation method derived from applying the sequential kernel density estimation ideas described in earlier sections to local STFT-based periodograms, rather than wavelet decompositions. However, our objective here is not to establish that our approach is better than other signal detection methods, but rather to provide an additional model-free methodology that performs relatively effectively when compared to a straightforward "baseline" technique.

7.1 Artificial Sound Data

In constructing artificial sound data, it was convenient to use easily generated human speech as the basic known signal. Accordingly, a single spoken word was recorded in laboratory conditions to ensure a negligible amount of noise in the recording; this word provided a series of 13 signals when split into temporal "windows" of the duration described in Section 5. To produce simulated underwater noise with which to "contaminate" these known signals, a section of typical interference was extracted from our real underwater recordings, an STFT was performed to obtain its spectral decomposition, and an inverse filter was then designed using the Yule–Walker algorithm of Friedlander and Porat (1984). The designed filter was then applied to simulated white noise, thereby generating artificial underwater noise with a similar spectral decomposition to the real underwater sound background. The series of known speech signals was then embedded additively toward the end of the period of this simulated underwater noise, leaving an initial section known to contain no signals.

We repeated this contamination procedure using different SNRs to produce a range of test datasets with different levels of interference present where for our purposes, SNR (in dB) is defined on a logarithmic scale, as the ratio of the energy (sum of squares) of the signal to the energy of the noise. For the different datasets, we used SNRs of 5, 0, -5, -8, and -10 dB, where the positive value implies that the signal is more powerful than the noise and vice versa for negative values, with 0 corresponding to signal and noise having equal energy. In practice, at an SNR exceeding -8 dB, speech is barely distinguishable by the human ear, whereas at the other extreme, an SNR of 5 dB represents clearly audible speech. Different realizations of the simulated underwater noise were used for contamination at each different SNR level.

Each resulting test dataset was designed to provide 128 successive multivariate observations, $\mathbf{x}_t$, for the mean sums of squares described in Section 5. Of these observations, 115 were known to relate only to artificial underwater noise, whereas, as mentioned previously, 13 related to the human speech signals with noise superimposed.

In this application, a priori knowledge of the signals to be detected suggests use of only two mean sums of squares, those covering the 1–3 kHz range. Accordingly we take $\mathbf{x}_t = (x_{t,4}, x_{t,5})$ as simply a bivariate observation in this case and take the first $T = 30$ observations in each of the test sets as sufficient to provide an initial stable kernel estimate of the two-dimensional density of these variables. Recall that these observations are known to contain no signal from the way in which the test data were constructed. The successive vectors in the remainder of each test set were then considered, and the signal detection method described in the previous section was applied using a 5% significance level. As a result, each of the corresponding 98 observations were detected as signal or as background noise. Table 1 summarises the results obtained over the various test datasets, where "false positive" refers to the percentage of misclassifications when noise is classified as signal and "false negative" gives corresponding results when signal is classified as noise.

To obtain some measure of the performance of our method, we compared the results to those arising from an alternative and "baseline" approach to signal detection. As discussed briefly in Section 3, many methods for time-frequency analysis use the STFT, and so we adopted a

Table 1. Performance of Wavelet (DWT)-and Fourier (STFT)-Based Signal Detection Methods

	% false + VEs		% false − VEs	
SNR (dB)	DWT	STFT	DWT	STFT
+5	4.7	3.5	0.0	7.6
0	6.2	0.0	0.0	53.8
−5	4.7	1.2	15.4	100.0
−8	11.8	1.2	15.4	92.3
−10	10.6	3.5	38.5	100.0

NOTE: Comparison relates to identifying 13 known human speech signals contaminated with artificially generated underwater noise at different SNRs, so as to form a section of sound of a total length of 98 observations.

commonly used technique based on this transform for the comparison here. This method compares the energy in frequency "bins" within successive time "windows" with their means over the whole period of sound, and flags a time window as a signal if one or more of the standardized differences exceeds predefined thresholds. This method, along with others, has been described in detail by Boashash and O'Shea (1990). The results of applying it here with optimal thresholds chosen so as not to identify any signals in the first 30 observations of known noise are given alongside the results of our method in Table 1.

In signal detection work of this nature, it is generally the false-negative rate that is of most concern, because the cost associated with overlooking true signals is judged to be high relative to that incurred by "false alarms." Given the relatively small number of signals present in the test data and the size of T, the results of our method indicate acceptable numbers of such misclassifications across the whole range of SNRs. Some variability in the trend in the rates across different SNRs arises from the fact that different examples of simulated underwater noise are used in each of the test datasets, but as might be expected, the number of false negatives broadly increases as the degree of noise increases. However, even at SNR levels where the signal is virtually inaudible to the human ear, our method experiences only 15% false negatives, whereas the STFT method fails to detect any of the signals present. Within reasonable limits, the numbers of false positives is of less importance in practical applications, and although the Fourier method performs better than our method in this respect, the latter gives acceptable results, misclassifying less than 12% even in the worst case.

Clearly, these comparisons are not extensive enough to draw any general conclusions about performance, which would require repetitions within SNR levels. Unfortunately, numerous repetitions at each SNR level are not feasible here, because of the lack of a sufficient number of suitable sections of known underwater noise on which to base the contamination. However, the results do serve the purpose for which they were designed—namely, to indicate that our signal detection approach can reasonably recover known sound signals against background interference of the type likely to be encountered in real underwater applications, and that it compares favorably in this objective with a commonly used time-frequency method.

7.2 Dolphin Sounds

The comparisons on the artificial data are encouraging, but the more important question is how well our method performs in real situations. Thus now consider a lengthy and "noisy" section of underwater sound of 12 s duration (491,520 original data points, or 1,920 of the successive temporal windows referred to in earlier sections). The selected section of sound involved dolphin clicks and whistles. It was immediately preceded on the recording by a .8-s period that could be guaranteed a priori to contain no such signals. This preceding period provided $T = 128$ successive multivariate observations $\mathbf{x}_1, \ldots, \mathbf{x}_T$, on the three selected mean sums of squares described in Section 5; that is, $\mathbf{x}_t = (x_{t,5}, x_{t,6}, x_{t,7})$. This was taken as sufficient to provide an initial stable kernel estimate of the three-dimensional density of these variables. Successive vectors in the immediately following 12-s period of the recording were then considered, and the detection method was applied using a 1% significance level. As a result, each of the corresponding 1,920 observations was detected either as signal or as background noise, with output consisting of a white (noise) or black (signal) indicator for each observation. The results are depicted for the full 12.8-s period in Figure 5, below the corresponding wavelet coefficients for the sound in the three levels used in the detection; that is, between frequencies from 2–10 kHz.

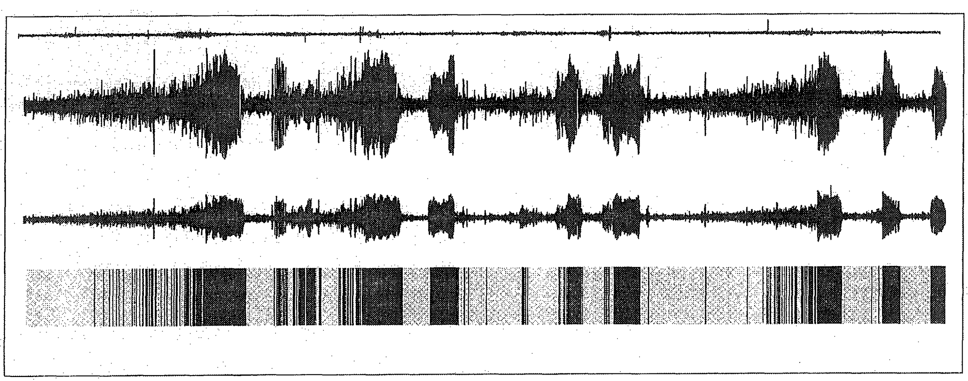

Figure 5. Discrete Wavelet Coefficients, $d_{j,k}$, Using Daubechies's Extremal Phase Wavelet With $N = 8$, in Frequency Bands 2.5 kHz, 5.1 kHz, and 10.2 kHz (with 2.5 kHz at the Top) and Wavelet Detection Results (Bottom, With Black Indicating Signal) for Selected 12.8 s (2,048 Time "Windows") of Dolphin Sound. The first .8 s (128 time windows) are known to contain no signals.

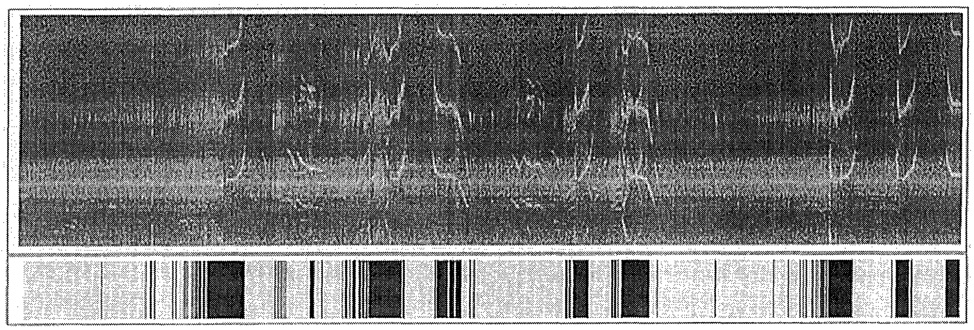

Figure 6. Spectrogram (Top) and STFT Detection Results (Bottom, With Black Indicating Signal) for Selected 12.8 s (2,048 Time "Windows") of Dolphin Sound. The first .8 s (128 time windows) are known to contain no signals.

In total, the wavelet method identifies 727 signals in the 1,920 unclassified observations relating to the final 12 s of the recording. Because the true signals are unknown in this case, we have no direct way to assess the accuracy of these detection results. However, in most cases there is a convincing visual match between salient features visible in the wavelet coefficients and the signal detection output.

A more objective way of assessing the results is to compare them with those obtained using the detection method based on the STFT described earlier. One way of summarizing the spectral decomposition obtained from a STFT is through a spectrogram (see, e.g., Boashash 1990). In such a representation, the horizontal axis relates to time window and the vertical axis to frequency bin; the energy of the sound at each time/frequency combination (which is proportional to the square of the amplitude of the appropriate sinusoid) is then displayed as a grayscale, usually black for low values through white for high values. The spectrogram for the full 12.8-s section of sound is shown in Figure 6, with 128 frequency bins covering the range up to 16 kHz and with time windows analogous to those used in the earlier wavelet analysis. The results of signal detection based on this spectrogram using the energy detection method described earlier are also included in Figure 6, below the spectrogram. Each time window is flagged as white (noise) or black (signal). Optimal thresholds for the method were chosen so as to detect no signals in the first 128 observations of known noise.

In total, the method based on the STFT identifies 578 signals in the 1,920 unclassified observations relating to the final 12 seconds of the recording. Comparison of the spectrogram and the detection results in Figure 6 reveals a reasonable visual match between the dominant sound features visible in the spectrogram and the output of the signal detection method. However, the STFT method has identified 149 fewer signals than the number identified by our wavelet approach on the corresponding sound section.

When the results of our method are compared in more detail to those based on the STFT, the two methods agree on the identification of 519 signals and 1,134 noise observations. In the remaining 267 cases, the STFT method identifies 59 signals that are flagged as noise by our method, and our method identifies 208 signals that are flagged as noise by the STFT method. What is clearly unknown is whether the latter are true signals that have been missed by the STFT method; the results obtained on the artificial data considered earlier would tend to suggest this hypothesis, and it is further supported by reference to Figure 5, where it is difficult to accept that as many as 208 of the signals detected fail to correspond to significant visible features in the wavelet decomposition. Of course, significant visible features in the wavelet coefficient sequences could still a priori be true noise, because we allow for possibly varying background recording conditions. One way of examining this would be through a more detailed look at the multivariate density estimate for noise in the vicinity of these sound sections, rather than the wavelet coefficient patterns. An alternative is to identify sections of the recording detected by the wavelet but not by the STFT method, then subject these sections to an audio analysis. Although somewhat informal and subjective, the latter suggests that many of the features overlooked by the STFT approach can actually be identified by the human ear as short-term dolphin clicks, albeit with some difficulty. Thus it would appear that there are a substantial number of signals in this period of sound that do not appear in the spectrogram and are not detected by the STFT method, but are visible in the wavelet decomposition and are appropriately picked up by our method. This finding lends support to our earlier arguments for preferring wavelet over Fourier transforms in this particular application and to preferring the wavelet detection results over those generated by the baseline STFT method.

8. CONCLUSIONS AND DISCUSSION

Wavelet analysis has been applied to underwater sounds of differing types with the objective of detecting signals against a potentially changing background sound environment or noise in the data. Our motivation for seeking a signal detection method based on the use of wavelets over one using Fourier analysis is that wavelets provide better time-frequency localization for transients, and signals of very short-term duration are often of primary interest in the kind of application considered in this article.

The proposed method of signal detection constructs an adaptive model of the background underwater sound environment in the space of the wavelet representation, using recursive kernel estimation of the joint distribution of certain summary measures of the wavelet coefficients. Observations considered to be outliers from the kernel estimate at any time are then flagged as signals. We argue for this approach, as opposed to one based on straightforward thresholding of wavelet coefficients or the recent use of wavelet thresholding in postprocessing for adaptive segmentation, because the former are directed primarily toward the recovery of smooth features, which is not our primary concern, and the latter incoporate models that may be inappropriate in our application.

Overall, given a priori knowledge of the broad frequency band of interest, so that appropriate levels of the wavelet decomposition may be selected, and also given a time series that contains an adequate initial period that is known to contain no signals of interest, the techniques that we have described allow the rest of the series to be identified as signal or noise at intervals of very short time duration with moderate computational effort. Our method does not require any prior knowledge of suitable thresholds or features of the noise distribution, because this knowledge is built adaptively through time. Furthermore, the suggested approach is general and easily adaptable to different situations. It thus may be applicable to data from other sources, such as those obtaining from medical scanning, spectroscopy, or tomography of various descriptions. Moreover, the method used lends itself well to applications where the ultimately aim is subsequent discrimination between detected signals. The results clearly identify the length of signals, and their signatures are characterized by the multivariate behavior of the dominant frequency/energy measures, $\mathbf{x}_t$, throughout the signal. This provides a basis for later detailed classification of the type of signal involved.

Our detection of outliers is based on empirical considerations, and it could be argued that it is a contradiction to do this; instead, we should use basic knowledge about the general distributional structure of the data generation process. For example, it might be suggested that it is better to examine broad specifications of the noise distribution such as symmetry, or to look at location-slippage explanations of outliers (see, e.g., Barnett and Lewis 1994). But, such specifications simply are not available for our background sound environment, and we have no alternative but to accept the limited assessment of outliers yielded by an essentially nonparametric empirical approach. Empirical arguments also apply to justify the test procedure that we have used to decide whether any new observation is an outlier from the most recent kernel estimate of the noise distribution. Because our noise distribution is data dependent and contains features that change and evolve over time, it is not possible to theoretically develop a test statistic that would have any meaningful generality in the situations likely to be encountered in practice.

In summary, we have presented an approach that reasonably recovers known signals contaminated with varying degrees of simulated noise. More important, it also produces plausible results when applied to noisy real recordings of dolphin sound. We have demonstrated that these results compare favourably with corresponding results obtained using a more conventional, straightforward time-frequency signal detection method. Knowledge of the true signals present in our recordings is not available, and so clearly we cannot conclude on the basis of these comparisons that our approach is better than other signal detection methods. However, establishing this point is not our main interest. Signal detection is a difficult practical problem and one probably best approached using a multiplicity of methods, rather than trying to identify a single best technique. Our method has simply been presented as a possibly useful addition to the growing range of signal-processing tools.

[Received April 1996. Revised July 1997.]

REFERENCES

Abramovich, F., and Benjamini, Y. (1996), "Adaptive Thresholding of Wavelet Coefficients," *Computational Statistics and Data Analysis*, 22, 351–361.

Abramovich, F., Sapatinas, T., and Silverman, B. W. (1996), "Wavelet Thresholding via a Bayesian Approach," Technical Report, University of Bristol, U.K., Dept. of Mathematics.

Abry, P., Goncalves, P., and Flandrin, P. (1995), "Wavelets, Spectrum Analysis and $1/f$ Processes," in *Wavelets and Statistics* (Lecture Notes in Statistics 103), eds. A. Antoniadis, and G. Oppenheim, New York: Springer-Verlag, pp. 15-29.

Barnett, V., and Lewis, T. (1994), *Outliers in Statistical Data.* (3rd ed.), Chichester, U.K.: Wiley.

Boashash, B. (1990), "Time-Frequency Signal Analysis," in *Advances in Spectrum Analysis and Array Processing*, ed. S. Haykin, Englewood Cliffs, NJ: Prentice-Hall, pp. 418–517.

Boashash, B., and O'Shea, P. (1990), "A Methodology for Detection and Classification of Some Underwater Acoustic Signals Using Time-Frequency Analysis Techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38, 1829–1841.

Chen, C. H. (1985), "Recognition of Underwater Transient Patterns," *Pattern Recognition*, 18, 485–490.

Cohen, L. (1989), "Time-Frequency Distributions—A Review," *Proceedings of the IEEE*, 77, 941–981.

Coifman, R. R., and Saito, N. (1994), "Construction of Local Orthonormal Bases for Classification and Regression," *Comptes Rendus Academie Sciences, Paris*, Série A, 319, 191–196.

Creasey, D. J., Smith, B. V., and Gazey, B. K. (1989), "Underwater Acoustics and Sonar Systems—Signal Processing and Systems," Vol. 2, Lecture Notes, Birmingham University.

Dalhaus, R. (1997), "Fitting Time Series Models to Nonstationary Processes," *The Annals of Statistics*, 25, 1–37.

Daubechies, I. (1988), "Orthonormal Bases of Compactly Supported Wavelets," *Communications in Pure and Applied Mathematics*, 41, 909–996.

——— (1992), *Ten Lectures on Wavelets*, Philadelphia: SIAM.

Donoho, D. L. (1993), "Non-Linear Wavelet Methods for Recovery of Signals, Densities and Spectra From Indirect and Noisy Data," in *Proceedings of Symposia in Applied Mathematics: Different Perspectives on Wavelets*, 47, ed. I. Daubechies, San Antonio: American Mathematical Society, pp. 173–205.

Donoho, D. L., and Johnstone, I. M. (1994), "Ideal Spatial Adaption by Wavelet Shrinkage," *Biometrika*, 81, 425–455.

——— (1995), "Adapting to Unknown Smoothness via Wavelet Shrinkage," *Journal of the American Statistical Association*, 90, 1200–1224.

Donoho, D. L., and Johnstone, I. M., Kerkyacharian, G., and Picard, D. (1995), "Wavelet Shrinkage: Asymptopia?" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 57, 301-369.

Donoho, D. L., Mallat, S., and von Sachs, R. (1996), "Estimating Covariances of Locally Stationary Processes: Consistency of Best Basis Methods," in *IEEE Conference on Time-Frequency/Time-Scale Methods*, June 1996, Paris.

Friedlander, B., and Porat, B. (1984), "The Modified Yule–Walker Method of ARMA Spectral Estimation," *IEEE Transactions on Aerospace Electronic Systems*, 20, 158–173.

Gorman, R. P., and Sejnowski, T. J. (1988), "Analysis of Hidden Units in a

Layered Network Trained to Classify Sonar Targets," *Neural Networks*, 1, 75–89.

Johnstone, I. M., and Silverman, B. W. (1997), "Wavelet Threshold Estimators for Data With Correlated Noise," *Journal of the Royal Statistical Society*, Ser. B, 59, 319–351.

Jones, D. L., and Parks, T. W. (1992), "A Resolution Comparison of Several Time-Frequency Representations," *IEEE Transactions on Signal Processing*, 40, 413–420.

Khattree, R., and Naik, D. N. (1987), "Detection of Outliers in Bivariate Time Series Data," *Communications in Statistics, Part A—Theory and Methods*, 16, 3701–3714.

Learned, R. E., and Wilsky, A. S. (1995), "A Wavelet Packet Approach to Transient Signal Classification," *Applied and Computational Harmonic Analysis*, 2, 265–278.

Ludeman, L. C. (1986), *Fundamentals of Digital Signal Processing*, Cambridge: Harper & Row.

Mallat, S. G. (1989), "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 674–693.

Meyer, Y. (1992), *Wavelets and Operators*, Cambridge, U.K.: Cambridge University Press.

Nason, G.P. (1993), "The WaveThresh Package; Wavelet Transform and Thresholding for S," available from the Statlib archive.

——— (1995), "Choice of the Threshold Parameter in Wavelet Function Estimation," in *Wavelets and Statistics* (Lecture Notes in Statistics 103), eds. A. Antoniadis, and G. Oppenheim, New York: Springer-Verlag, pp. 261–280.

——— (1996), "Wavelet Shrinkage Using Cross-Validation," *Journal of the Royal Statistical Society*, Ser. B, 58, 463–479.

Nason, G. P., and Silverman, B. W. (1994), "The Discrete Wavelet Transform in S," *Journal of Computational and Graphical Statistics*, 3, 163–191.

——— (1995), "The Stationary Wavelet Transform and Some Statistical Applications," in *Wavelets and Statistics*, (Lecture Notes in Statistics 103), eds. A. Antoniadis, and G. Oppenheim, New York: Springer-Verlag, pp. 281–300.

Neumann, M. H., and Spokoiny, V. G. (1995), "On the Efficiency of Wavelet Estimators Under Arbitrary Error Distributions," *Mathematical Methods of Statistics*, 4, 137–166.

Neumann, M. H., and von Sachs, R. (1997), "Wavelet Thresholding in Anisotropic Function Classes and Application to Adaptive Estimation of Evolutionary Spectra," *The Annals of Statistics*, 25, 38–76.

Ogden, T., and Parzen, E. (1996), "Change-Point Approach to Data Analytic Wavelet Thresholding," *Statistics and Computing*, 6, 93–99.

Pal, N. R., and Pal, S. K. (1993), "A Review on Image Segmentation Techniques," *Pattern Recognition*, 26, 1277–1294.

Penny, K. I. (1996), "Appropriate Critical Values When Testing for a Single Multivariate Outlier by Using the Mahalanobis Distance," *Applied Statistics*, 45, 73–81.

Powell, K. J., Sapatinas, T., Bailey, T. C., and Krzanowski, W. J. (1995), "Application of Wavelets to the Pre-Processing of Underwater Sounds," *Statistics and Computing*, 5, 265–273.

Ramsey, J. B., Usikov, D., and Zaslavsky, G. M. (1995), "An Analysis of U.S. Stock-Price Behaviour Using Wavelets," *Fractals—An Interdisciplinary Journal on the Complex Geometry of Nature*, 3, 377–389.

Remme, J., Habbema, J. D. F., and Hermans, J. (1980), "A Simulative Comparison of Linear, Quadratic and Kernel Discrimination," *Journal of Statistical Computation and Simulation*, 11, 87–106.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.

Smith, D. J., Bailey, T. C., and Munford, A. G. (1993), "Robust Classification of High Dimensional Data Using Artificial Neural Networks," *Statistics and Computing*, 3, 71–81.

Smith, M. J., and Mersereau, R. M. (1992), *Introduction to Digital Signal Processing*, New York: Wiley.

Strang, G. (1993), "Wavelet Transforms Versus Fourier Transforms," *Bulletin (New Series) of the American Mathematical Society*, 28, 288–305.

Telfer, B. A., Szu, H. H., Dobeck, G. J., Garcia, J. P., Ko, H., Dubey A., and Witherspoon, N. (1994), "Adaptive Wavelet Classification of Acoustic and Backscatter and Imagery," *Optical Engineering*, 33, 2192–2203.

Tukey, P. A., and Tukey, J. W. (1981), "Graphical Display of Data Sets in 3 or More Dimensions," in *Interpreting Multivariate Data*, ed. V. Barnett, Chichester, U.K.: Wiley, pp. 189–275.

von Sachs, R., Nason, G., and Kroisandt, G. (1996), "Spectral Representation and Estimation for Locally Stationary Wavelet Processes," in *Proceedings of the Workshop on Spline Functions and Wavelets*, April 1996, Montreal, Canada.

Wand, M. P., and Jones, M. C. (1995), *Kernel Smoothing*, London: Chapman and Hall.

Wang, Y. (1995), "Jump and Sharp Cusp Detection by Wavelets," *Biometrika*, 82, 385–397.

——— (1996), "Function Estimation via Wavelet Shrinkage for Long Memory Data," *The Annals of Statistics*, 24, 466–484.

Weyrich, N., and Warhola, G. T. (1995), "De-Noising Using Wavelets and Cross-Validation," in *Approximation Theory, Wavelets and Applications*, NATO ASI Series C, 454, ed. S.P. Singh, Dordrecht: Kluwer, pp. 523–532.

Wierda, S. J. (1994), "Multivariate Statistical Process Control—Recent Results and Directions for Future Research," *Statistica Neerlandica*, 48, 147–168.