Naval Research Laboratory
Stennis Space Center, MS 39529-5004

NRL/MR/7174--96-8006

Coherence Estimation of Shallow Water Acoustic Narrowband CW Pulsed Signals

CDR Samuel M. Nagle

Office of Naval Research
Science and Technology Reserve Det. 822
Seattle, WA

Roger W. Meredith

Ocean Acoustics Branch
Acoustics Division

August 2, 1996

REPORT DOCUMENTATION PAGE

Form Approved
OBM No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 2, 1996	Final

4. TITLE AND SUBTITLE

Coherence Estimation of Shallow Water Acoustic Narrowband CW Pulsed Signals

5. FUNDING NUMBERS

Job Order No. 571-6861-00
Program Element No. 0602435N
Project No.
Task No. UW-35-2-05
Accession No.

6. AUTHOR(S)

CDR Samuel M. Nagle* and Roger W. Meredith

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Acoustics Division
Stennis Space Center, MS 39529-5004

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/MR/7174--96-8006

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Acoustics Division
Washington, DC 20375-5320

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Office of Naval Research, Science and Technology Reserve Det. 822, Seattle, WA

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Different methods for estimating signal coherence from underwater acoustic data taken in shallow water are compared. Specifically this study details methods to estimate coherence from narrow-band, high-frequency, short pulse length, CW signals that typify the signals used by minehunting sonars in shallow water environments. Due to the short record length of the received pulses, the spectral frequency resolution was poor using classical Fourier techniques. This motivated alternate approaches for the spectral estimation and coherence estimates. The other methods compared are an autoregressive parametric based approach, harmonic wavelet approach, and concatenation of successive pulses followed by the classical method.

14. SUBJECT TERMS

MCM sonar systems designs, high frequency acoustics, spatial and temporal coherence

15. NUMBER OF PAGES

36

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

CONTENTS

COHERENCE ESTIMATION OF SHALLOW WATER ACOUSTIC NARROWBAND CW PULSED SIGNALS

1.0 INTRODUCTION AND BACKGROUND

A fundamental underlying assumption in the processing of underwater acoustic signals by sonar arrays, is that the signal is coherent across the spatial aperture of the array, and over the time duration of the processing. By summing the responses from an array of hydrophones when receiving and processing a signal, a sonar system can provide higher signal-to-noise levels than do individual hydrophones. Coherent signals add while incoherent signals cancel, conversely, array gain degrades when unwanted signals are coherent. In general, fluctuations in the ocean medium can act to decorrelate desired signals and increase correlation in unwanted signals, thus reducing array gain. Accordingly, the evaluation of the coherence of acoustic signals is important to predicting sonar performance.

The variability of signal coherence is of particular importance in minehunting and mine classification sonars that are required to detect and classify small objects. These sonars are often characterized by their small angular beamwidths (<3 degrees) and fine range resolution (<1.5 m). In order to achieve such capabilities, high frequencies (>20 kHz) are normally used over relatively short ranges (<1500 m). Minehunting typically occurs in a complex and dynamic shallow water environment. This environment can change significantly over short distances or over small time periods. Storms, tides, currents, and seasonal variations create significant environmental changes in time frames ranging from hourly to seasonally and in spatial frames ranging from a few tens of meters to several miles. An acoustic signal becomes "distorted" as it propagates in the medium due to multipath interference and scattering from the boundaries and from inhomogeneities in the medium. Most often the sound speed varies with depth resulting in sound speed gradients that cause the acoustic ray to refract or bend in the direction of decreasing sound speed. All of the aforementioned factors affect signal characteristics including amplitude, frequency, phase, duration, directivity, etc. The coherence of a signal provides a good overall measure of the net effect all of these pertubating factors, and how it will affect sonar performance.

In this paper we evaluate different methods for deriving a quantitative measure of signal coherence from underwater acoustic data taken in shallow water. Specifically

this study details methods to estimate coherence from narrow-band, high-frequency, short pulse length, CW signals that typify the signals used by minehunting sonars in shallow water environments.

Coherence is a statistic that is classically used as a measure of the linearity of a system, and relates two random processes, one an input process, the other an output. Using the notation from Bendat and Piersol [1], coherence is defined as follows:

$$\gamma(f)^2_{xy} = \frac{|S_{xy}(f)|^2}{S_{xx}(f)\ S_{yy}(f)} \tag{1}$$

where

$$S_{xx} = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-j2\pi f\tau} d\tau \tag{2}$$

$$S_{yy} = \int_{-\infty}^{\infty} R_{yy}(\tau) e^{-j2\pi f\tau} d\tau \tag{3}$$

are the respective autospectral density functions, and

$$R_{xx} = \frac{1}{T}\int_0^T x(t)x(t+\tau)dt \tag{4}$$

$$R_{yy} = \frac{1}{T}\int_0^T y(t)y(t+\tau)dt \tag{5}$$

are the autocorrelation functions. Similarly, the cross spectral density function is given by:

$$S_{xy} = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j2\pi f\tau} d\tau \tag{6}$$

where

$$R_{xy} = \frac{1}{T}\int_0^T x(t)y(t+\tau)dt \tag{7}$$

is the cross correlation of the two processes being evaluated.

In an ideal linear system, the coherence will be equal to unity. Coherence will be less than unity when any one of the following four conditions exist:

1. Extraneous noise is present in the measurements.
2. Resolution bias errors are present in the spectral measurements.
3. The system relating input to output is not linear.
4. The output is dependent on more that just a single input.

In the case of underwater acoustic signals, factors 1, 3, and 4 above can be affected by the medium, as previously discussed. As the coherence falls below unity, sonar performance will be degraded.

We examine four methods to estimate coherence, each method differs in the technique used to evaluate the spectral terms of Eq. 1 that are given by expressions 2, 3 and 6. In the first method we used classical spectral estimation based on Fourier transform methodology. We found that due to the limited observation time of the received pulses, the spectral frequency resolution was poor. This limitation motivated us to seek alternate schemes for the spectral estimation, and led us to try using the other techniques. In the second method, we implemented a parametric model of the received signals (Auto Regressive model). The third method used a recently developed tool for signal analysis known as the wavelet transform to extract time-frequency information from each pulse, and develop a statistic akin to coherence from the derived wavelet transform coefficients. The last method involves concatenation of successive pulses to increase the record length in order to increase spectral resolution.

2.0 DATA DESCRIPTION

The data set used in this study was collected as part of high frequency acoustic experiments conducted in shallow water. In the experiments, short duration pulsed sinusoids are transmitted from a projector to a receiving array, where the data is band

shifted to 5 kHz, low pass filtered, quadrature sampled, and stored for post processing. For acoustic measurements done in shallow water, short duration pulses are required so that the direct path contribution may be identified and separated in time of arrival from other components of the signal that are due to surface and bottom reflections, and reverberation.

A train of pulses were transmitted at a specific frequency, pulse length, and repetition rate for a set number of 150 pulses. Although the frequency of the actual transmitted tones varied from run to run, in all cases the received signals were modulated so that the center frequency was 5 kHz prior to being sampled at a 20 kHz rate. The data selected for our study consisted of 150 pulses; each pulse was a gated sinusoid at a frequency of 150 kHz, with a pulse width of 1 millisecond.

3.0 EVALUATION OF COHERENCE

To evaluate the variations in coherence, the sequence of pulses provides a train of input data that can be compared over time and space, with each pulsed sinusoid being evaluated to create the terms of Eq. 1. If we let:

S_{xx} = The autospectral density of P_{ref}

where P_{ref} is a reference pulse,

S_{yy} = The autospectral density of P_i

where P_i is a pulse indexed by i, and

S_{xy} = The cross spectral density of P_{ref} with P_i,

we have a method that lets us measure the linear relationship between different received pulses. In the case of evaluating temporal coherence, the reference pulse, P_{ref}, is typically selected as the first pulse in a given run. The coherence as a function of time is then determined by considering each subsequent pulse in the train (indexed by i) as an output of a system driven by the reference pulse. Although not done in our study, spatial coherence may also be evaluated using a similar method, where a reference hydrophone may be defined in the receiving array, and the received outputs from other hydrophones within the array can be treated as system outputs.

3.1 Evaluation of Coherence Based on Fourier Spectral Estimates

The key to determining coherence comes in the evaluation of the spectral densities, expressions 2, 3 and 6, through periodogram estimation. The best known classical method is implemented by taking the Fourier transform of the correlations, or of the original time history records. The correlations were used to facilitate estimating cross-spectrums using the auto-regressive model discussed in section 3.2. They are used with the other methods to be consistent. Typically the Fourier transform is executed through the use of the Fast Fourier Transform (FFT) algorithm, the squared magnitudes of the resulting complex numbers are used to generate the power spectral density function estimate. By taking the Fourier transform of the correlation function, which is always an even function, the resulting power spectrum is always real. Note that when dealing with a sampled data set such as this, the power spectrum estimates appear as discrete frequency components separated in frequency by

$$B_e = \frac{1}{T} \tag{8}$$

where T equals the record length. In section 3.4, concatenation of successive pulses is used to increase T. Once the auto and cross spectral densities are estimated, coherence is calculated by substituting the terms into Eq. 1.

As a demonstration of the utility of this method for determining coherence, this technique was used on a run of data acquired in a shallow water acoustic experiment conducted in Kiel, Germany. As previously stated, the particular run used in our analysis consisted of 150 pulses transmitted at a 1 Hz repetition rate. Each pulse was a 1 millisecond sinusoid CW signal at a frequency of 150 kHz (Kiel data, channel 30 of run #333). Recall that even though the actual transmission frequency is 150 kHz, the signal was band shifted to a center frequency of 5 kHz. Figure 1 illustrates a typical received pulse taken from this run, with only the direct path portion of the signal shown.

To obtain estimates of the auto and cross spectral densities, the Welch procedure of windowing and averaging was used [2]. In this procedure, the original data is divided into a integral number of segments, possibly overlapping. A window is applied

to these segments and the resulting modified periodograms are averaged to obtain an estimate of the power spectral density.

In our analysis, some choices were required in order to obtain the spectral estimates. To form a coherence estimate, we have two prime considerations, spectral averaging (needed for coherence estimate) and spectral bin width. To get a large number of averages and good resolution we need lots of sample points, but high resolution sonars require short pulse lengths to get good temporal resolution so we don't have many samples to begin with, plus we have to reduce the ones we have to do the spectral averaging needed for a statistical coherence estimates. If the ocean medium is stationary and ergodic then we can ensemble average spectrums over successive pings. This is not applicable in this data because we want to investigate how coherence changes temporally (from ping to ping as described above) over the time it might take to form a synthetic aperture. Our choices for parameters used in the spectral estimation are given as follows:

Window length: This factor determines the Fourier transform length and correspondingly, the resulting bandwidth of each spectral bin per Eq. (8). Since the direct path portion of our pulse consists of only about 32 points, we tried window lengths of 8 and 16. Such short time windows result in extremely broad spectral bins, but give us a sufficient number of windows so that an estimate can be achieved by averaging spectrums from a single ping.

Window type: The selection of the window type was not critical in our case, and was a bit arbitrary. A Hanning window was used.

Amount of overlap between windows: A 50% overlap was used (i.e. 4 points of overlap when an 8-point window was used, and 8 points of overlap when a 16-point window was used). Overlap processing allows ensemble averaging of spectral estimates, which reduces the variance on each spectral level estimate at the cost of increased computational load due to computing multiple FFTs.

Using the above choices for spectral estimation parameters, the auto spectral density estimates for the pulse in Figure 1 are illustrated in Figure 2 where the results using an 8-point FFT window and a 16-point FFT window are shown. This results in spectral frequency bin widths of 2500 Hz and 1250 Hz respectively. The bandwidth of the original CW pulse is on the order of 10 Hz. For CW signals with a high

signal-to-noise ratio, the spectral level is 3-6 dB higher for the signal frequency bin than for the neighboring bins and can be easily resolved. For lower SNR signals this will not be the case, resulting in larger fluctuations in spectral levels and thus more severe fluctuations in coherence. In addition, active sonar systems are often reverberation limited rather than ambient noise limited and reverberant returns tend to be more coherent than ambient noise. Accordingly, spectral estimates for short narrowband CW pulses are more difficult to estimate. This results in spectral frequency bins much larger than the original signal bandwidth, thus reducing signal-to-noise of the coherence estimate.

After obtaining the auto and cross spectral densities, the coherence as a function of frequency was derived from Eq. 1. Figure 3 displays the coherence of pulse number 150 taken from our dataset, using pulse number 1 as the reference pulse. Because of the spectral estimation problems, the resulting coherence estimate varies by less than 10% over a large frequency span. SNR is sufficient to resolve the coherence peak, but the relative difference in coherence between the peak and adjacent coherence bins is less than 3%. By repeatedly calculating the coherence at the signal frequency for each pulse in the experiment run, using pulse number 1 as P_{ref}, the temporal coherence over the duration of the run can be evaluated. Figure 4 illustrates the variability of coherence over the 150 second run period. Plots are shown for the center frequency bin (at the signal center frequency) and the two adjacent frequency bins. The signal bandwidth is 200 Hz; the spectral binwidth is 1250 Hz for the 16-point FFT and 2500 Hz for the 8-point FFT. Notice the high coherence in both adjacent frequency bins where no signal exists. This indicates a coherent background and the long term trends are significantly different for the two adjacent bins.

How much variability in the temporal coherence is due to spectral estimations problems and how much is due to the ocean medium? Some component of the high frequency oscillations are related to spectral estimation problems of a narrow band signal in wideband coherent noise. Some component may be associated with wind and wave motion with periods less than a few seconds, and perhaps, turbulence. Longer term trends in Fig. 4 are more closely associated with ocean tidal and swell motions with periods of a few seconds. Very long term trends on the order of minutes may be related to non-dispersive internal waves. Note that the coherence of the adjacent frequency bins show only slightly lower coherence values but larger high frequency oscillations over the 150 seconds, thus lending some credence to the hypothesis that the

high frequency variability is related to a narrow band signal in wideband coherent noise.

3.2 Evaluation of Coherence Using an Auto Regressive Parametric Model

Parametric model based power spectral estimators overcome the limitation on spectral resolution bandwidth by extrapolating the signal outside of the analysis window, effectively providing a longer record length for analysis (T, in Eq. 8). Several good references are available that provide a detailed description of the underlying mathematics that surround these methods [2, 3, 4], only a brief overview is given here.

Parametric methods fall into three subgroups: autoregressive (AR), moving average (MA), and autoregressive moving average (ARMA). All three methods model the original signal by estimating its poles and zeros, whose numbers also define the transfer-function model's order. For instance, an AR model of order N means its transfer-function is modeled by N poles; similarly, an MA model of order N is modeled by N zeros. Using the standard z-transform notation common for sampled data systems, the general system transfer function equation for an ARMA model is given by:

$$H(z) = \frac{b(0) + b(1)z^{-1} + b(2)z^{-2} \ldots b(q)z^{-k}}{a(0) + a(1)z^{-1} + a(2)z^{-2} \ldots a(p)z^{-k}} \tag{9}$$

In our study, we restricted ourselves to using the Maximum Entropy Spectral modeling method developed by Burg [5] to derive the polynomial coefficients $a(0)$ through $a(p)$ in the denominator of Eq. (9). Since we used an all pole AR model, we set $b(0)=1$, and $b(1) = b(2) = \ldots = b(q) = 0$.

To obtain the autospectral density and the cross spectral density estimates necessary for estimating coherence, we followed the method of first determining the autocorrelation and the cross correlation sequences from the raw data and then using these correlation values as inputs into the Burg modeling algorithm to derive the AR coefficients. The frequency response spectrums for the all pole models were then used as estimates for the auto and the cross spectrums respectively.

The power spectrum derived by the AR model is shown in Figure 5. In this case, a model order of 3 was used, and a data set size of 128 was specified, resulting

in spectral bins that are 156.25 Hz wide. Note that the source peak at 5000 Hz is much narrower with respect to the earlier FFT based spectral estimate (Fig. 2).

Problems occurred when coherence was calculated based on the AR spectral estimates. As was done for the Fourier methods, a plot of coherence vs. frequency for pulse number 150 referenced to pulse number 1, is shown in Figure 6. In a paper by Makhoul [6], it is shown that as the model order is increased, a signal's spectrum can be approximated arbitrarily close by an all pole model. In our case, we found that using model orders greater than 3 resulted in ill-conditioned matrices. The ill-condition occurrence at such a low model order is most likely due to the large dynamic range of the signal being modeled, and the limited samples available for generating the correlation matrix necessary to derive the AR coefficients. This limitation on our model order, results in large errors in the absolute magnitudes of the spectral density estimates used in our calculation of coherence. Thus the coherence formula (Eq. 1), results in values for coherence that are greater than 1. This makes interpretation difficult and accordingly, may be of no use as a "true" measure of coherence. However, if we consider this AR model based statistic as a new measure of "likeness," similar to coherence, it may be possible to characterize how it changes as a function of phase, amplitude, and frequency, and find some utility for it's use in predicting sonar performance. The signal coherence at 5000 Hz is more clearly resolved than the FFT estimate of coherence and the level of the adjacent bins drastically reduced. This has the potential for estimating our signal coherence more accurately. If a method could be found to raise the model order without resulting in an ill-conditioned correlation matrix, then perhaps this method would give better results.

As it is, Figure 7 shows the AR estimated temporal coherence for the same 150 pulses computed by classical methods shown in Figure 4. Note that both the structure and trends of the AR based coherence are different than from the FFT based computation shown in Figure 4.

3.3 Evaluation of Coherence Using the Discrete Wavelet Transform

Here we explore the use of wavelet theory as a method for determining coherence. Briefly, the wavelet transform is similar to the Fourier transform methodology in that it expands a function using a family of basis functions. A signal is decomposed into a set of component parts, that when summed together reconstruct

the signal. It differs, however, in two aspects from the Fourier transform: 1) the basis functions may be either finite or infinite in length, 2) all of the basis functions are time shifted as well as time compressed or time dilated versions of a single analyzing base wavelet function. The theory behind wavelets is complicated, and it is not the objective of this paper to describe wavelets in depth. Many references are available [7, 8, 9] that provide a good introduction to wavelet theory and their applications in engineering and science.

Using the notation from Newland [7], the goal of the wavelet transform is to decompose an arbitrary signal f(x) into an infinite summation of wavelets at different scales according to the expansion

$$f(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{j,k} W(2^j x - k). \tag{10}$$

The analyzing base wavelet is W(x), and each of the terms within the double summation of Eq. 10 represent time scaled (either compressed or stretched) and time translated versions of W(x). The index j is an integer commonly referred to as the wavelet "level" and can roughly be thought of as being analogous to frequency. With each ascending value of j, the wavelet becomes contracted in time by a factor of 2. The contraction in the time domain results is a doubling of the frequency bandwidth, and thus a constant time-bandwidth product for all wavelet levels is achieved, a universal characteristic of wavelet functions. The index k in Eq. 10 represents a time translation of a wavelet within the analysis interval for a particular level.

It can be shown that for $j<0$, $W(2^j-k)$ can always be expressed as a sum of terms like $\phi(x-k)$, which is termed a scaling function. Eq. (10) can then be expressed in the alternative form as:

$$f(x) = \sum_{k=-\infty}^{\infty} C_{\phi,k} \phi(x-k) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} C_{j,k} W(2^j x - k). \tag{11}$$

At a given wavelet level j, 2^j wavelets are required to span the record length under analysis. At wavelet level 0, a single analyzing wavelet spans the entire analysis range, at level 1, two wavelets are required, at level 2, four, and so on. By convention a wavelet level of -1 is used to account for the contribution from the scaling function,

and for our application this contribution reduces to the DC component (mean value) of a signal.

Wavelets have proven extremely useful in time-frequency analysis of non-stationary signals such as speech. The advantage of wavelets over Fourier methods is that it provides a means to localize frequency characteristics to specific time frames. An algorithm for computing the wavelet transform when the signal f(x) is sampled at equally spaced intervals over $0 \leq x \leq 1$, is the Discrete Wavelet Transform (DWT). This algorithm developed by Mallat [10], is similar to the Discrete Fourier Transform in that it assumes that the signal $f(x)$ is one period of a periodic signal. The algorithm computes the coefficients $c_{j,k}$ from Eq. 10. For a sampled dataset of length N, the number of wavelet levels provided by the DWT will be $\log_2 N$. As an example, a data set of length 64 will generate 6 wavelet levels, 0 through 5, not including the -1 level needed for the DC component. Assuming that the data are generated from sampling a continuous time process using a sample rate of 20 kHz, Table 1 summarizes the number of wavelets needed to span the record length at each level, and also specifies the corresponding spectral center frequency for the wavelets at a given level.

Table 1: Wavelet Levels

Level j	Number of wavelets needed to span analysis interval	Wavelet Center Frequency (Hz)
-1	0	DC
0	1	312.5
1	2	625
2	4	1,250
3	8	2,500
4	16	5,000
5	32	10,000

A common means of displaying information from the wavelet transform is to plot the squared coefficient values on a two dimensional grid base, where one dimension is wavelet level (index j, which is akin to frequency), and the other dimension being wavelet position (index k), a measure of time, in this manner a time-frequency map is

formed. Details about how to form the map are explained in [7], but the positions of the coefficients are adjusted to allow for the differing lengths of the wavelets they define within each level. An important feature of this mapping is that the mean-square value of a signal is given by the volume under the two-dimensional wavelet surface of the square values of the wavelet coefficients over the time-frequency plane. The resulting mapping is termed the "mean-square map" of the signal.

As an illustration of how the mean-square map is used for displaying time-frequency information, two examples are given. Figure 8 presents a frequency modulated chirp in the time domain and the corresponding contour plot of the mean-square map on the two dimensional time-frequency grid. The signal shown here represents a linear frequency shift that is changing in frequency at the rate of 200 Hz/ second. The interval under analysis is 1 second of data sampled at 1024 Hz. Note that the wavelet levels run from -1 through 9, consistent with that expected for 1024 data points. The signal energy is clearly defined by map contours and shifts upward in level as the frequency rises. Similarly, a mean-square map for the received acoustic pulse from Figure 1 is shown in Figure 9. This map was derived by windowing 64 points around the direct path portion of the signal, resulting in 5 wavelet levels. Note that the signal energy is concentrated at wavelet level 4, this level corresponds to the octave of frequencies centered about 5 kHz as would be expected (see Table 1), and near the center of the wavelet position. Most of the energy is contained in the single wavelet level. Only the wavelet levels adjacent to this level exhibit any other energy, and the lower levels exhibit no energy. At both ends of the wavelet position axis, the energy is reduced due to the samples kept prior to and after the actual pulse. The steepness of the contours indicate the high signal-to-noise ratio of the pulse. This is what a very narrowband, noiseless pulse looks like in the wavelet domain.

The wavelet transform used to generate the mean-square maps of Figs. 8 and 9, used a wavelet basis from a family of infinite length wavelet functions that were developed by Newland [7], and have the characteristic that each wavelet level represents a component from an octave band of frequencies. The wavelets are known as harmonic wavelets, are defined in the frequency domain by

$$W(\omega) = (1/2\pi)2^{-j}e^{-i\omega k/2^j} \qquad \text{for} \qquad 2p2^j \le w < 4p2^j \qquad (12)$$
$$= 0 \text{ elsewhere}$$

where j is the wavelet level. The Fourier transform of the wavelet has the characteristic of being of constant magnitude within an octave band of frequencies, and being 0 outside of this band. The magnitude is set to normalize the enclosed area to unity. By calculating the inverse Fourier transform of $W(\omega)$, we find the corresponding complex wavelet is of the form

$$w(2^j x - k) = \frac{e^{i4\pi(2^j x-k)} - e^{i2\pi(2^j x-k)}}{i2\pi(2^j x-k)} \tag{13}$$

An efficient DWT algorithm using this wavelet basis was developed by Newland [7], and when it is applied to a set of real sampled data, the coefficients generated are complex conjugate pairs. The squared magnitudes used for the coefficients are determined by multiplying each coefficient with its complex conjugate.

Figure 10 illustrates the magnitude of $w(x)$ defined by Eq. 12 for the case where j=1. Note from the shape of the function that the magnitude of the wavelet coefficients ($c_{j,k}$) derived by taking the inner product of the real and imaginary parts of this wavelet with the signal under analysis, will get their largest contribution from that portion of the signal localized around t=0 with respect to the wavelet function. In the case of these harmonic wavelets, the rate of decay in the time domain is inversely proportional to x, thus providing a time localization depending on where the particular wavelet is centered.

We have used the harmonic wavelet family and the resulting information provided in the mean-square maps, to calculate a numerical analog of coherence using the calculated wavelet coefficients. Whereas a signal's mean-square (power) can be determined by the area under its power spectral density curve derived from the Fourier transform, in wavelet analysis, the mean-square value of a signal is given by the volume under the two-dimensional wavelet surface of the square values of the wavelet coefficients over the time-frequency plane. Using this analog relationship of the wavelet transform derived mean-square maps, with the Fourier transform derived power spectrums, we form a statistic equivalent to that given in Eq. 1, except we substitute wavelet transform derived power terms for the auto power spectrums (denominator factors) and the cross power spectrum (numerator).

To get our coherence estimate from the wavelet coefficients, we first take autocorrelation functions (Eqs. 4 & 5) and the crosscorrelation function (Eq. 7) of two

pulses, a reference pulse, and a pulse indexed by i as previously described. The DWT is then applied to each of the correlations, and the mean-square maps for each of the correlation functions are derived from the coefficients. We select the wavelet level where the signal power is predominant, level 4 in this case, where the 5 kHz frequency is centered. The array of mean-square map coefficients from level 4 of the cross correlation function map is then divided, element by element, by the element by element product of square valued level 4 wavelet coefficients derived from the two autocorrelation functions. To be more explicit in how the wavelet coherence is developed, we make the following definitions:

i) Let $C_{xx,j}(k)$ be the array of squared wavelet coefficients at level j derived from applying the DWT to the autocorrelation of discrete time process x(n).

ii) Let $C_{yy,j}(k)$ be the array of squared wavelet coefficients at level j derived from applying the DWT to the autocorrelation of discrete time process y(n).

iii) Finally, let $C_{xy,j}(k)$ be the array of squared wavelet coefficients at level j derived from applying the DWT to the cross correlation of discrete time process x(n) with y(n). Now perform the following element by element operations on the arrays to create a new array of wavelet coherence values that are indexed by k, wavelet position.

$$\gamma^2_{Wav_{xy,j}}(k) = \frac{\left|C_{xy,j}(k)\right|^2}{\left|C_{xx,j}(k)\right|\left|C_{yy,j}(k)\right|} \qquad \text{for all k at level j} \tag{14}$$

The number of elements contained in the array generated by Eq. 14 depends on the original data set size under analysis, and the level j for which the calculation is being made. As previously shown in the Table 1, when a 64-point dataset is used, we get wavelet levels running from -1 to 5. Our wavelet level of interest is level 4, which has 16 wavelets equally spaced across the analysis interval, thus each wavelet coherence array will have 16 elements.

Now turning our attention back to the acoustic data, we can generate a wavelet coherence array for each pulse of the 150 ping run from which the data was acquired. For each ping an array of 16 elements is created. Each array element is a calculated wavelet coherence centered at a particular lag or time position within the correlation functions. The indices of the array are represented by the variable k, and run from

1 to 16. The lowest lag value is indexed as k=1, and the highest lag value is indexed at k=16. If we form a 2 dimensional plane with one axis being lag value (position k), and the other axis being time (ping number), we can sequentially plot the wavelet coherence arrays for each ping of the run to form a surface over the plane that represents the temporal wavelet coherence over the duration of the run. Figure 11 illustrates such a surface plot for the same run of pings where we had previously shown the temporal coherence for in Fig. 4.

What does the plot of Figure 11 tell us about coherence? Additional information on coherence as a function of time spanning the record length (not the 150 pulse data set) is gained by using wavelets. Coherence is now a function of wavelet position, k, corresponding to the record length as well as a function of pulse number. Here we see that the coherence has relative maxima near wavelet positions 4,5 and at 15 as well as relative minima near wavelet positions 10,11. They appear consistent throughout the data set. How much of this curvature is related to environmental changes in the medium, or processing parameters, or equipment fidelity is not known at this time, but it is new information. By examining the temporal wavelet coherence at a position near the center of the pulse (at elements k=8 and k=9), we get a plot that looks very similar to the temporal coherence derived using the Fourier methods. Figure 12 illustrates the temporal wavelet coherence with a position value of k=8. Note that the basic shape and structure of the wavelet coherence is much like that seen in Fig. 4, but the higher frequency oscillations are more severe. This indicates that similar results can be obtained with this wavelet transform as are derived with the Fourier methods. The constant time-bandwidth product characteristic of the wavelet transform, however, again prevents a narrow band analysis at the signal frequency, and thus suffers from the same SNR limitations seen when we used Fourier methods.

Other wavelet bases can be used besides the harmonic wavelet chosen for this work. The most efficient recursive algorithms for wavelet transforms require orthogonal wavelets, although wavelets do not necessarily have to be orthogonal. Orthogonality is not easily accomplished while maintaining compact support in both temporal and spectral domains and it is for this reason the harmonic wavelet was initially selected. The design of wavelets with specific properties is beyond the scope of this paper, however, future work may investigate coherence results using other wavelet basis and the feasibility of designing a wavelet specifically for coherence estimation.

3.4 Evaluation of Coherence Using Concatenation of Successive Pulses

As a final method for estimating coherence we examined pulse concatenation. Concatenation of successive pulses can be used to achieve higher spectral resolutions at the expense of reducing the ensemble of pulses. Concatenation presents unique difficulties, but since our interest is in a single frequency, and since the signal is very narrow band, the effect of concatenation on only one frequency need be considered.

The sample length for each pulse in the concatenation group can be pre-determined by requiring an integer number of cycles. Since this data was digitized at 20 kHz and the frequency of interest is 5 kHz, after modulation, the sample length should be 20/5 = 4 cycles, which corresponds to 1 millisecond, or 20 samples. Fifteen successive pulses were concatenated for each group, yielding ten groups spanning the 150 seconds of the entire data set. Each group of concatenated pulses was processed for coherence using the Fourier spectral estimates identical to that in Section 3.1. For comparison, the concatenated pulses were processed using 128, 64, and 16 pt FFTs.

Often the resulting power spectrum from each group of concatenated pulses was poor, the spectrum were noisy with several high sidelobe levels. The spectrum levels and structure are very dependent on the starting sample and the total number of samples chosen for concatenation. By extensive trial and error the samples that best maximized coherence level were determined. For the 128-point FFT coherence estimation, 19 samples per group maximized the coherence level, for the 64-point FFT estimate 16 samples per group were best. The 16-point FFT estimate was not so sensitive to the group size or starting sample and the same values as for the 64-point FFT were used.

The more pulses used in the concatenation, the better frequency resolution in both the spectral levels and coherence estimates, however that leaves fewer groups and less temporal resolution over the total 150 second time span of the data set. Also, the more concatenation, the higher potential for errors due to possible relative phase differences in the individual pulse alignments.

The coherence estimate was obtained based on Fourier spectral estimates as previously described using the Welch procedure of windowing and averaging FFTs. Figure 13 shows a typical spectra for a single concatenation of 15 pulses using three different size FFTs. Concatenation allows a larger size FFT which enhances spectral

resolution. However, with this data, other peaks occur at frequencies that do not exist in the data and must be an artifact of the concatenation. They are probably due to relative phase errors in aligning two pulses. The spectral resolution is much improved, the binsize for the 128-point FFT is now on the order of the signal's bandwidth.

Figure 14 shows the coherence over the 150 second long data set using ten groups of 15 concatenated pulses. The coherence estimate was highly sensitive on the samples used. Sensitivity was most pronounced for the highest FFT size and varied by as much as 0.4. The long term trend in these coherence estimates compares favorably with that shown in Fig 4 using the standard FFT estimate, however, the coherence levels are much reduced and the structure is totally missing. This is the expense of reducing the ensemble of pulses in order to concatenate. Concatenation can be useful for short monochromatic pulses at the expense of reducing the ensemble of pulses. The effect of possible mismatched phases between concatenated pulses is the principal drawback. Manual alignment of each pair of pulses may be needed and can be tedious. By using a slightly different number of samples for each individual pulse, the phase errors may be reduced.

4.0 CONCLUSION AND SUMMARY

Linear coherence has many applications: as a measure of an input/output system's linearity, as a prediction for system output, and as a measure of common spectral density between two signals. We have looked at four different methods for determining coherence. As previously discussed, frequency resolution is constrained to the inverse of the limited observation time, and for the short duration pulsed sinusoids examined in this study, this results in poor frequency resolution. Application of classical Fourier based spectral estimation techniques for generating a periodogram result in very broad frequency bins deemed unacceptable for such short duration pulses. Fourier Transform methods still provide the best method for computing linear coherence for an input/output system. Averaging overlapped and windowed FFTs always improves the coherence estimate, however, the sampling requirements for adequate frequency resolution cannot always be optimally satisfied.

The frequency resolution is much improved using AR methods, however, the spectral magnitude errors imposed by the low model order results in large errors in the

spectral magnitudes, and yields coherence results that are difficult to interpret. AR coherence is not useful for array performance prediction or for predicting system output based on the input and the coherence function. It is also not a measure of nonlinearity of the system. Comparing the AR coherence from one system with another system is difficult if not futile. However, the AR coherence can be used in a limited sense as a relative measure of the commonality or 'sameness' between two signals or as a measure of commonality between outputs of a single system. The AR method also has potential for yielding better results if higher model orders could be achieved without incurring an ill-conditioning of the correlation matrix. For the data described herein, the AR coherence offers a large advantage in frequency resolution that more accurately reflects the relative coherence between the signal frequency and other non-signal frequencies than does the FFT based coherence as demonstrated in Figures 3 and 6.

Using orthogonal wavelets, the discrete wavelet transform provides a new method to estimate coherence, and a temporal coherence map can be easily generated that displays the value for the wavelet estimate of coherence in a two dimensional form. This maps the coherence of both the signal frequency band(s) and the non-signal frequency bands as a function of time. The wavelet coherence estimate matches well with both the overall trend and structure of the Fourier transform based coherence estimate as shown in Figure 4, although the actual levels are slightly different. The wavelet position gives additional insight into coherence over time scales on the order of the record length. The constant time-bandwidth characteristic of wavelet decomposition prevents a narrow band analysis at our signal frequency. The big advantage of the wavelet coherence estimate will be for broad band and chirp signals where the signal frequency spans several wavelet levels over time spans less than the sampling period.

Both the AR coherence estimate and the wavelet coherence estimate over the 150 seconds show more severe oscillations than the Fourier transform based coherence estimate. For the AR estimate this is most likely due to the errors imposed by using such a low model order for modeling the signal. Why this is so for the wavelet estimate may be in part due not to the wavelet estimate but to the broadness of the FFT spectral estimate (due to the small number of samples) resulting in a smoothed coherence estimate.

Concatenation improved spectral resolution and coherence as a function of frequency estimates at the expense of coherence as a function of pulse number. Only

the general, overall trend was preserved. Because this is real data, not computer synthesized, it is difficult to know which method gives the most correct value for coherence, but all four methods have something unique to offer. Separating out numerical effects, averaging effects, and signal propagation effects is difficult. Although the estimated values of coherence may be different, the tends and structure can be similar.

5.0 ACKNOWLEDGMENTS

I would like to thank the United States Naval Reserve for their support in conducting this study. The bulk of this work was done while serving my 2 weeks of annual active duty for training while attached to the Office Naval Research, Science and Technology Reserve detachment 822. I also want to express my gratitude to the Olin Aerospace Company for their full support in allowing me to schedule and fulfill my annual training requirement in support of this research task.

6.0 REFERENCES

1) Bendat, J. S. & Piersol, A. G., *Engineering Applications of Correlation and Spectral Analysis*, 1980, John Wiley & Sons, Inc.

2) Therrien, C. W., *Discrete Random Signals and Statistical Signal Processing*, 1992, Prentice-Hall (Englewood Cliffs, NJ)

3) Marple, S. L. Jr, *Digital Spectral Analysis with Applications*, pp 595–596, 1987, Prentice-Hall (Englewood Cliffs, NJ)

4) Kay, S. M., *Modern Spectral Estimation, Theory and Applications*, 1988, Prentice-Hall (Englewood Cliffs, NJ)

5) J. P. Burg. "Maximum Entropy Spectral Analysis." *Proceedings of the 37th Meeting of the Society of Exploration Geophysicists*, 1967.

6) Makhoul, J., "Linear Prediction: A Tutorial Review," *Proceedings of the IEEE*, 63(4):561-580, April 1975.

7) Newland, D. E., *Random Vibrations, Spectral & Wavelet Analysis*, 1993, Longman, Harlow and John Wiley, New York

8) Daubechies, I., *Ten Lectures on Wavelets*, 1992, Society for Industrial and Applied Mathematics (SIAM)

9) Coifman, R. R., Meyer, Y., Wickerhauser, V., *Wavelet Analysis and Signal Processing*, Yale University, 1991

10) Mallat, S., "A Theory for Multiresolution Signal Decomposition: the Wavelet Representation," *IEEE Trans. on Pattern Analysis and Machine Intell*, 11(7):674-693, July 1989

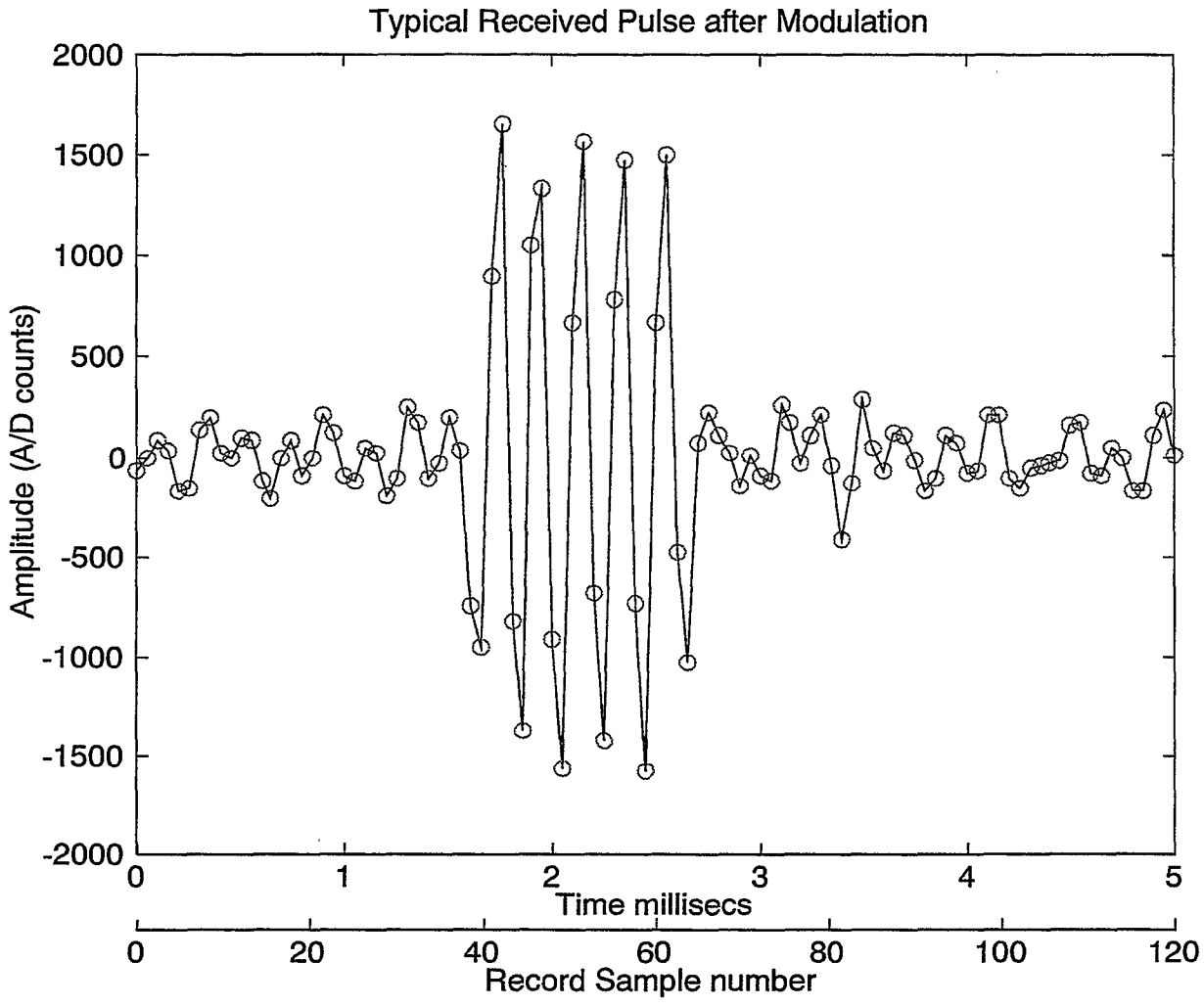

Figure 1: Typical narrowband, CW pulse signal received at 150 kHz, modulated to 5 kHz, and sampled at 20 kHz.

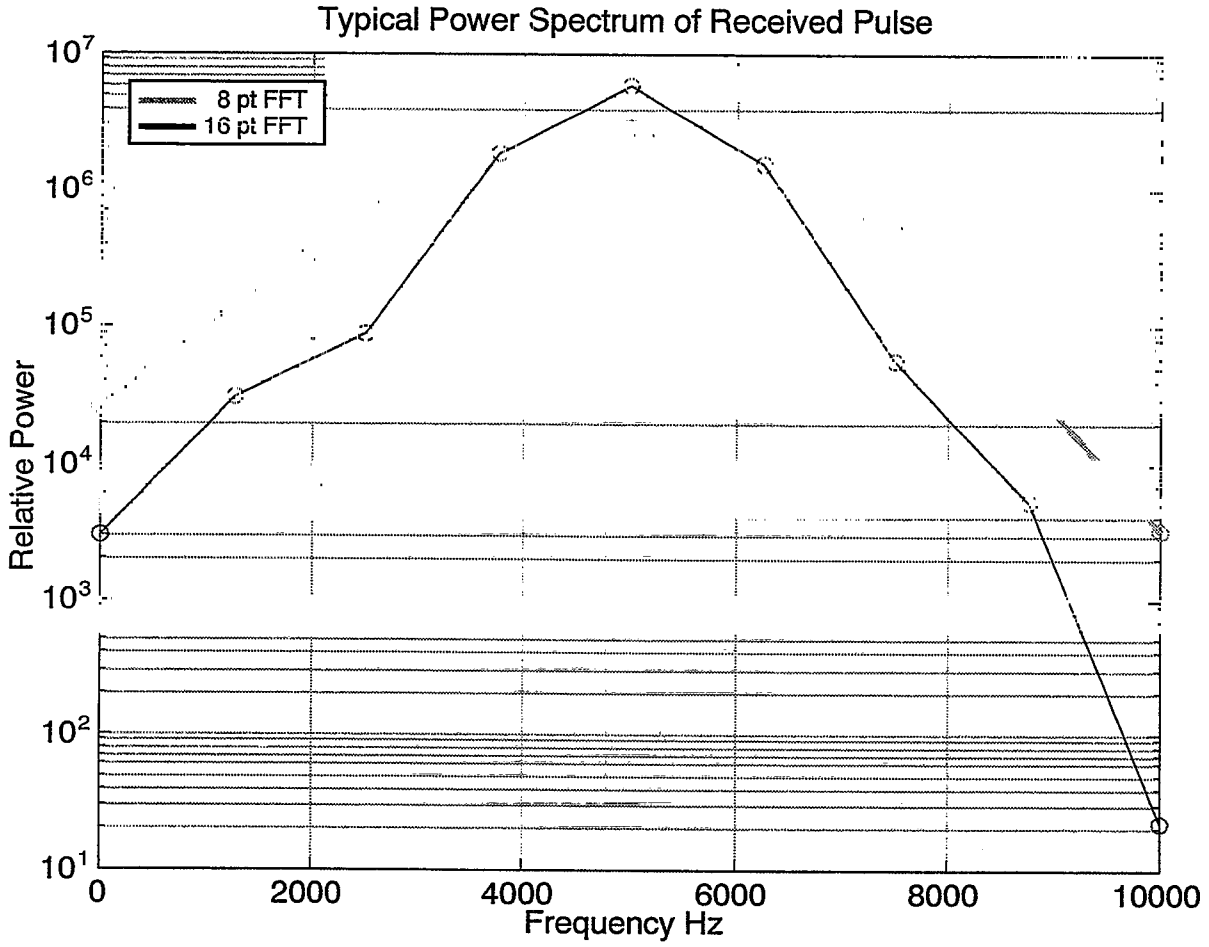

Figure 2: Power spectral density of received pulsed CW sinusoid using both 8-point and 16-point FFT sizes, 50% overlap, and a Hanning window.

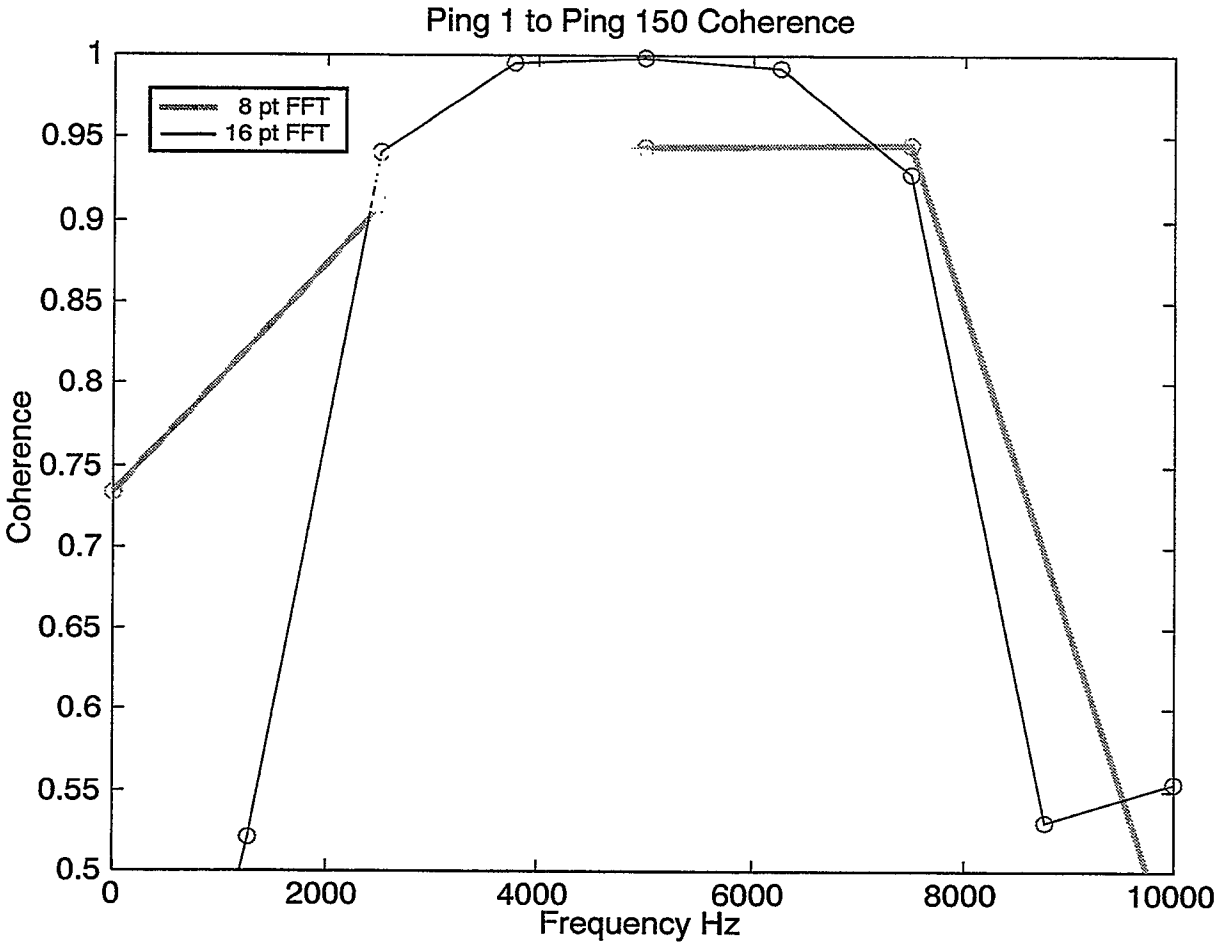

Figure 3: Coherence of pulse 150 with the reference pulse 1, both 8-point and 16-point FFT sizes for spectral estimation.

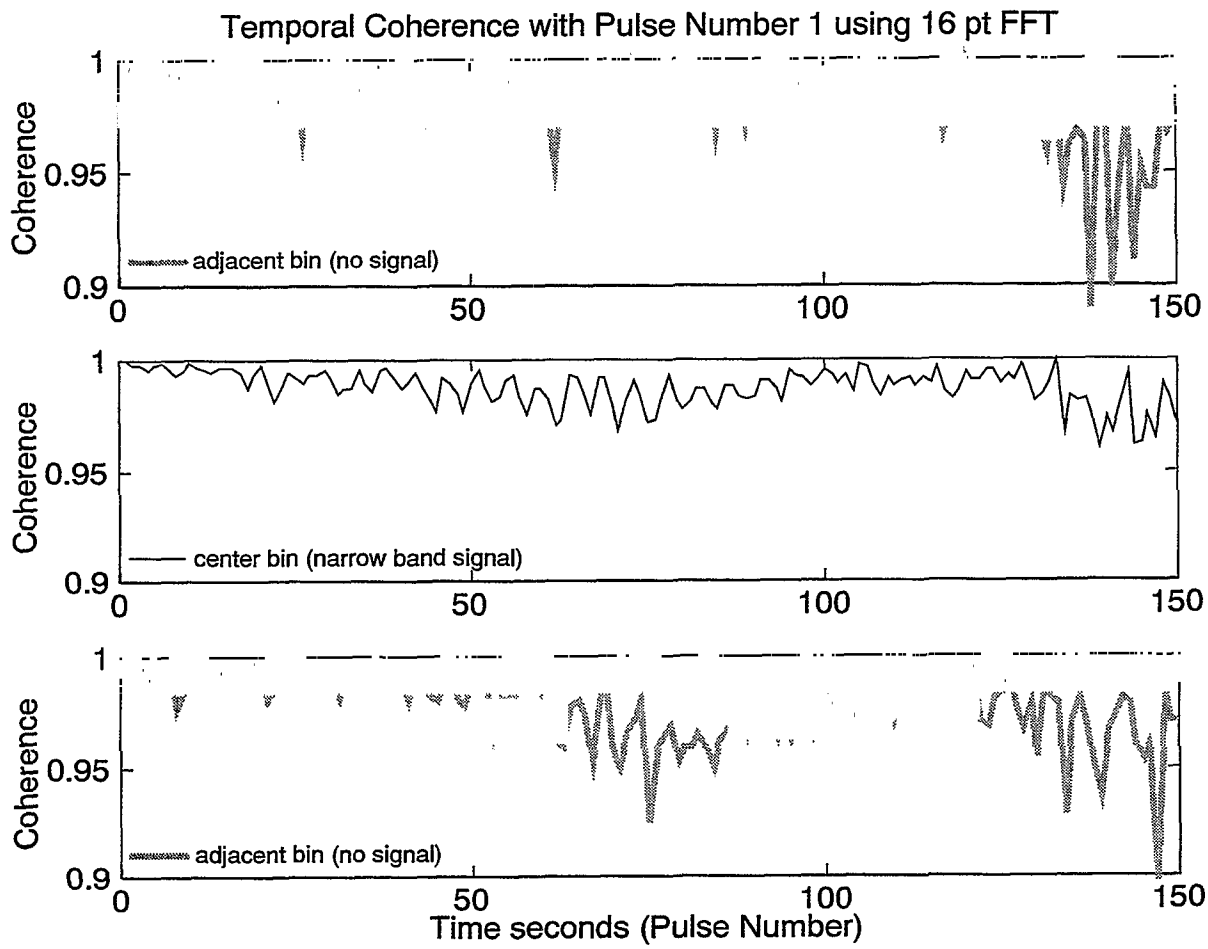

Figure 4: Coherence as a function of pulse number using pulse 1 as reference. Since the pulse repetition rate is 1 pulse per second, there is a 1-to-1 correspondence of pulse number with time in seconds.

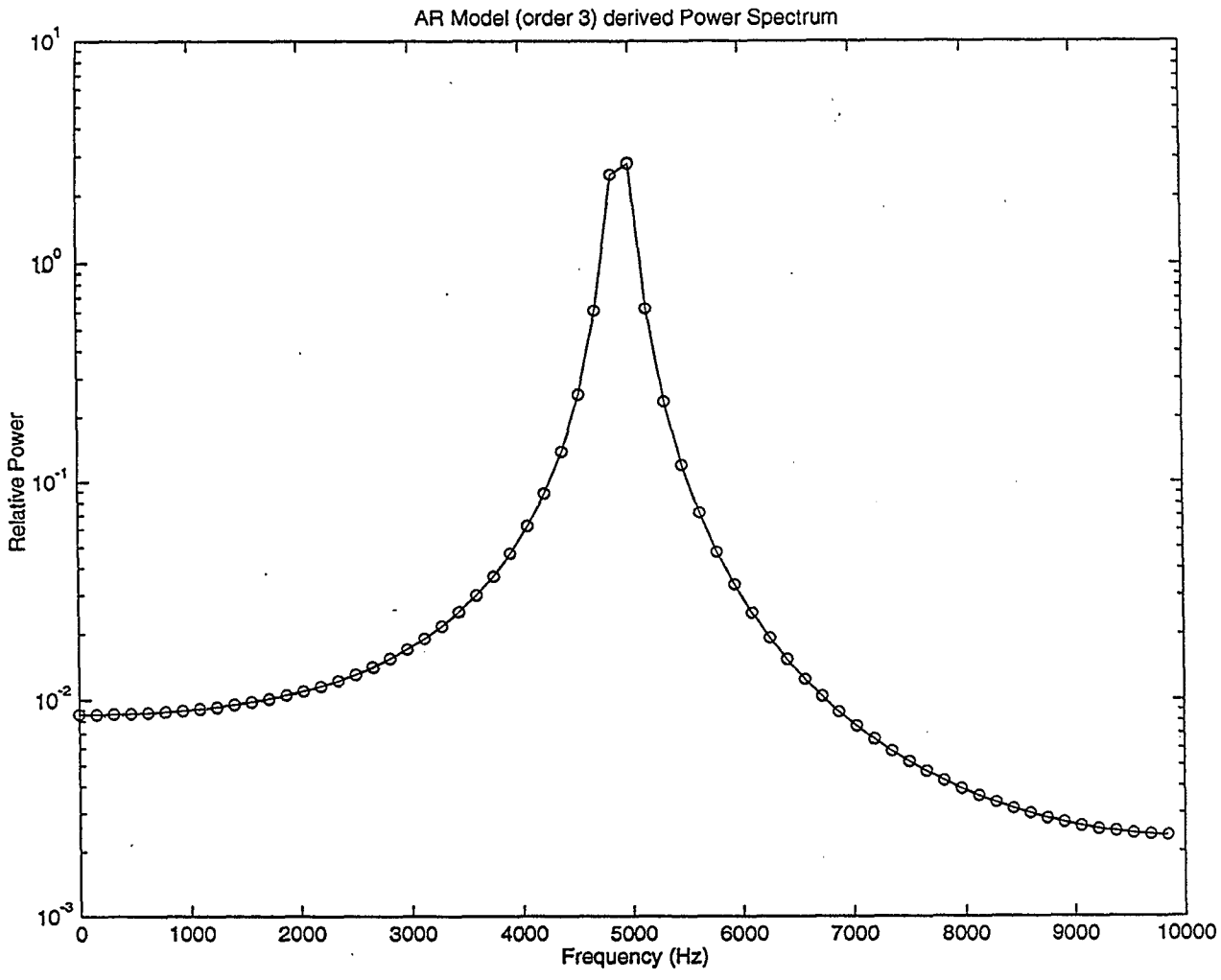

Figure 5: Power spectral density of received pulsed CW sinusoid using AR model for the autocorrelation of the received signal. Notice the increase in spectral resolution over that of Figure 2.

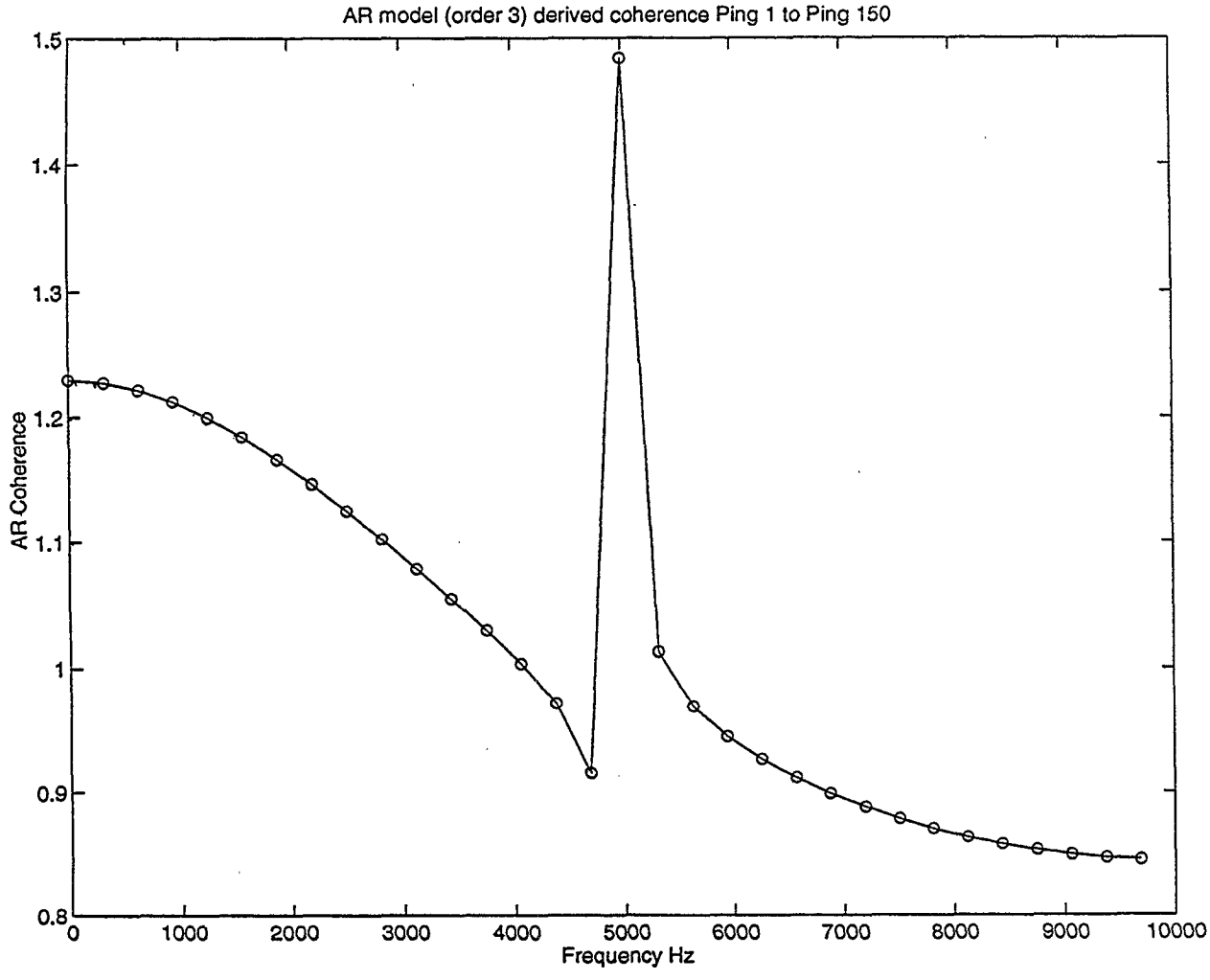

Figure 6: Coherence of pulse 150 with the reference pulse 1 using AR spectrum estimation. Notice the increase in spectral resolution over that of Figure 3.

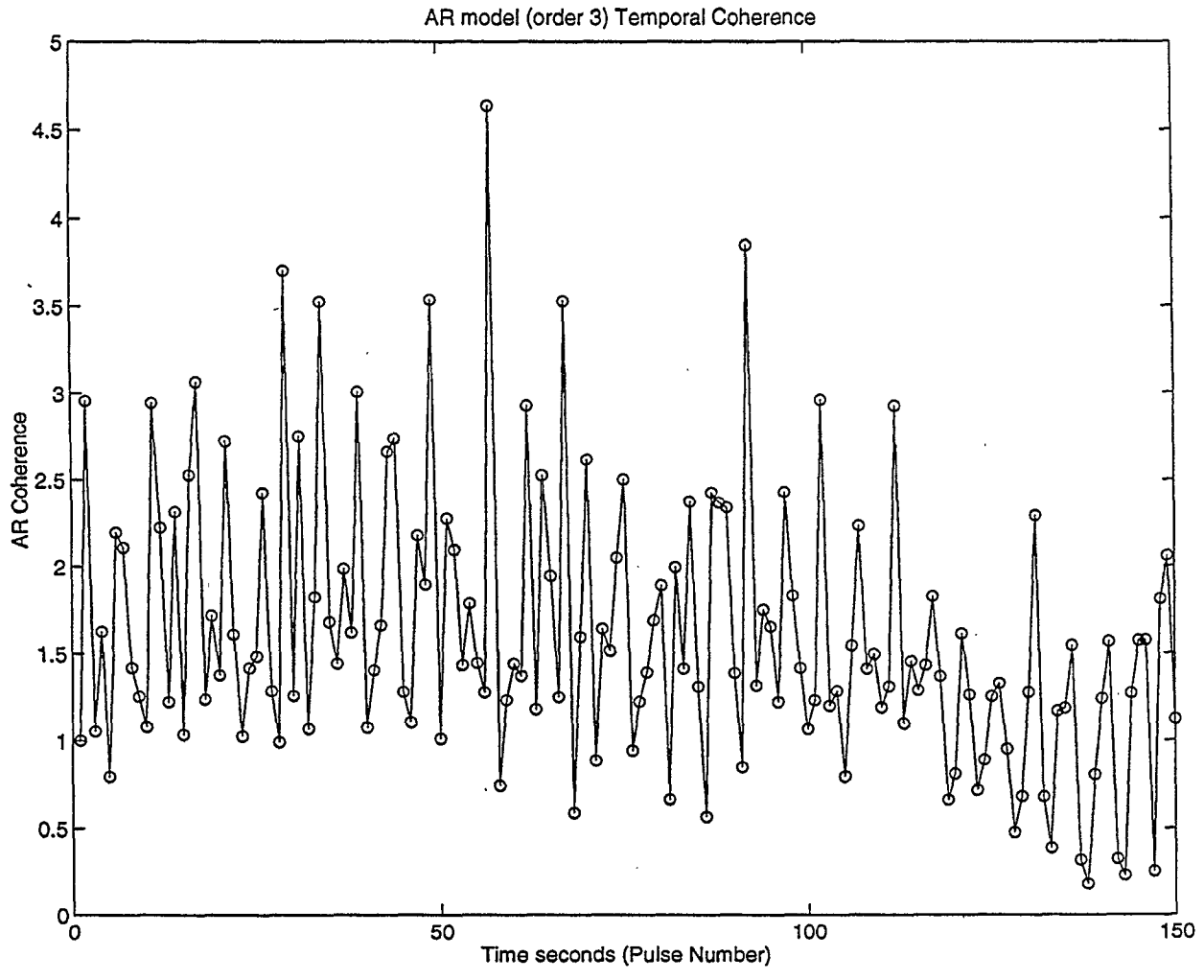

Figure 7: Coherence as a function of pulse number using pulse 1 as reference and AR spectrum estimation. Pulse repetition rate is 1 pulse per second. Compare to Figure 4.

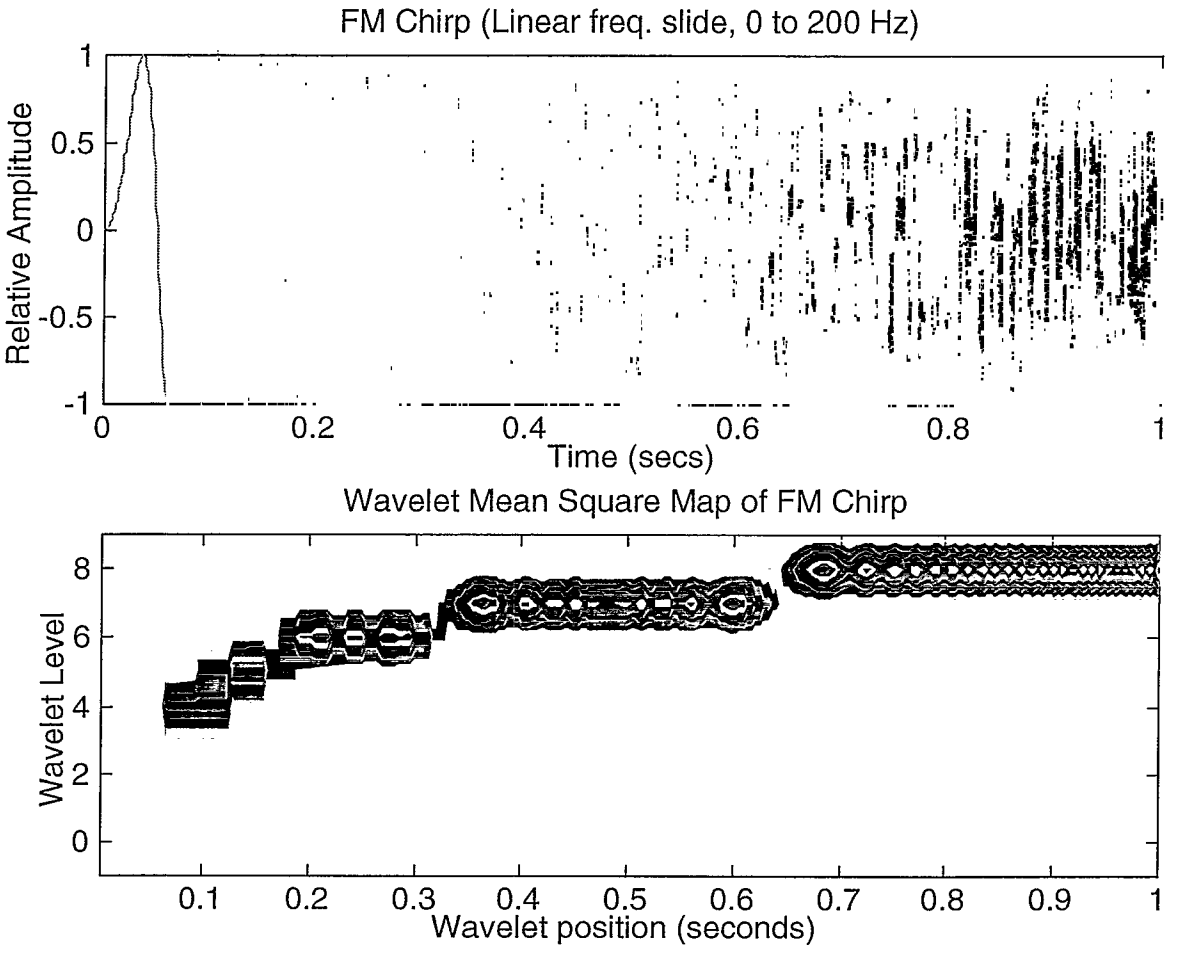

Figure 8: Synthesized frequency modulated chirp time series and contour plot of the resulting wavelet mean-square- energy map displaying wavelet position (time) and wavelet level (frequency) information. This purely illustrative example demonstrates the usefulness of wavelet processing for broadband signals.

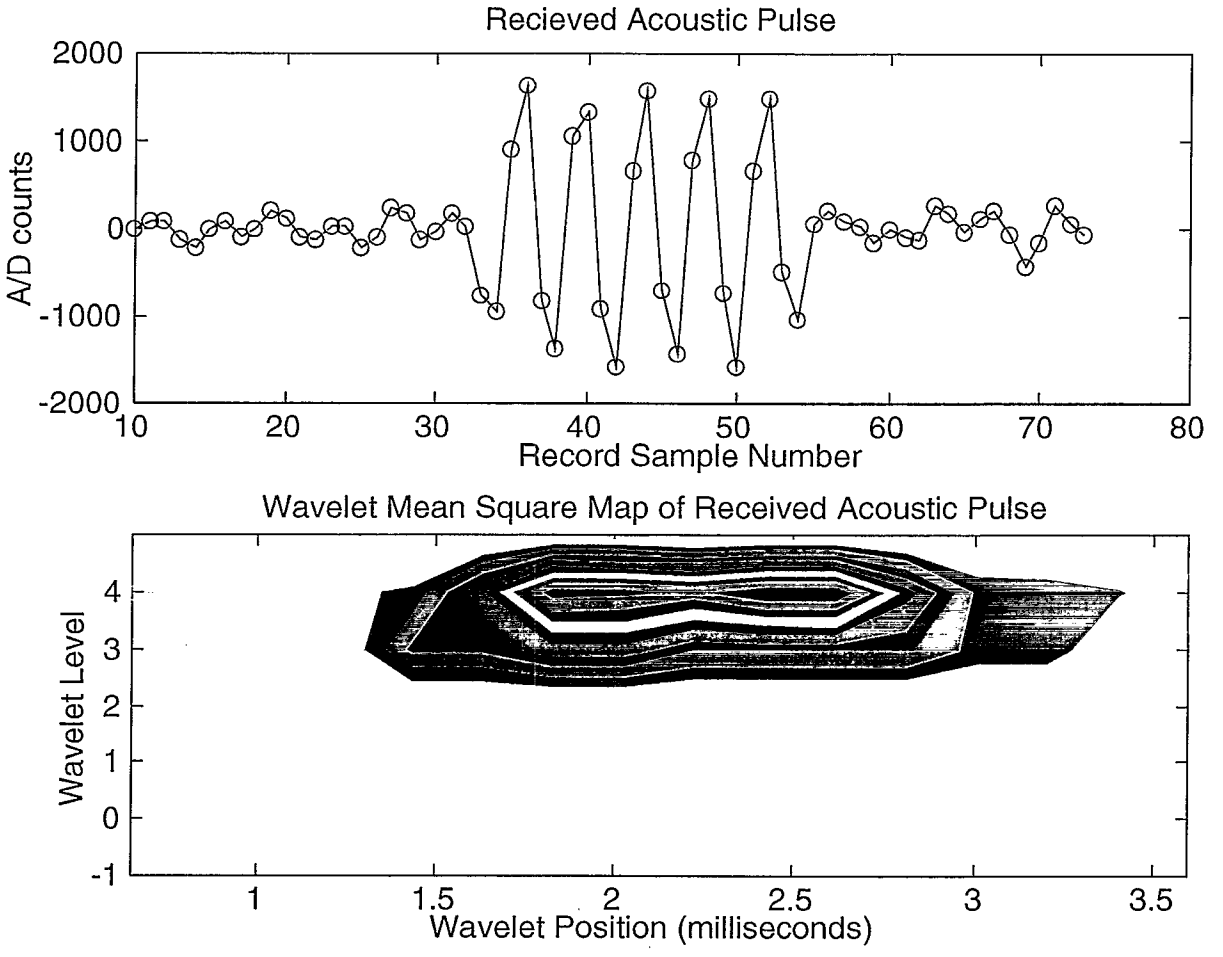

Figure 9: Time series and contour plot of the resulting wavelet mean-square-energy map displaying wavelet position (time) and wavelet level (frequency) information for the reference pulse, pulse number 1.

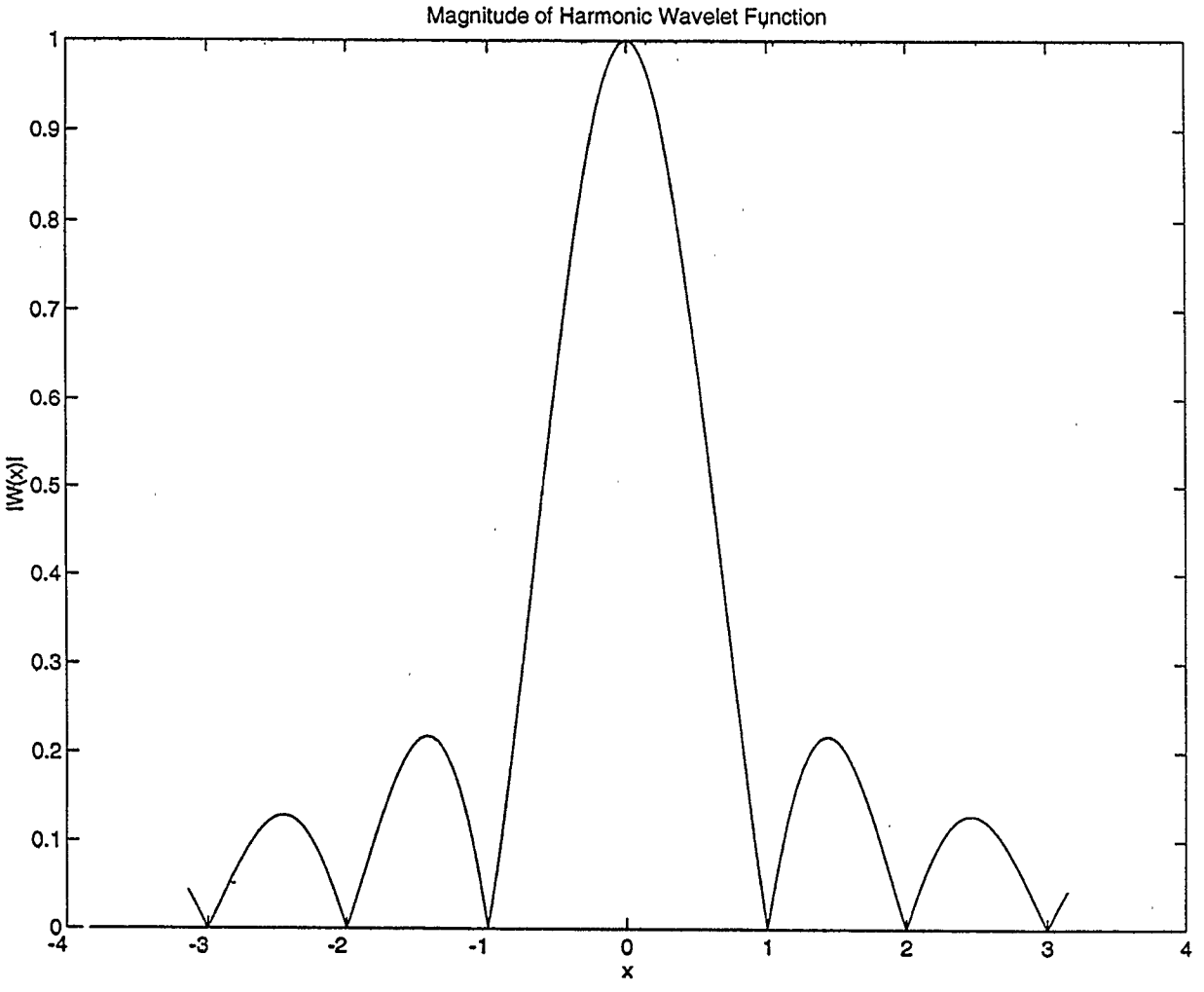

Figure 10: Magnitude of the complex harmonic wavelet developed by Newland and used to generate the wavelet coefficients used for computing the mean-square-energy maps.

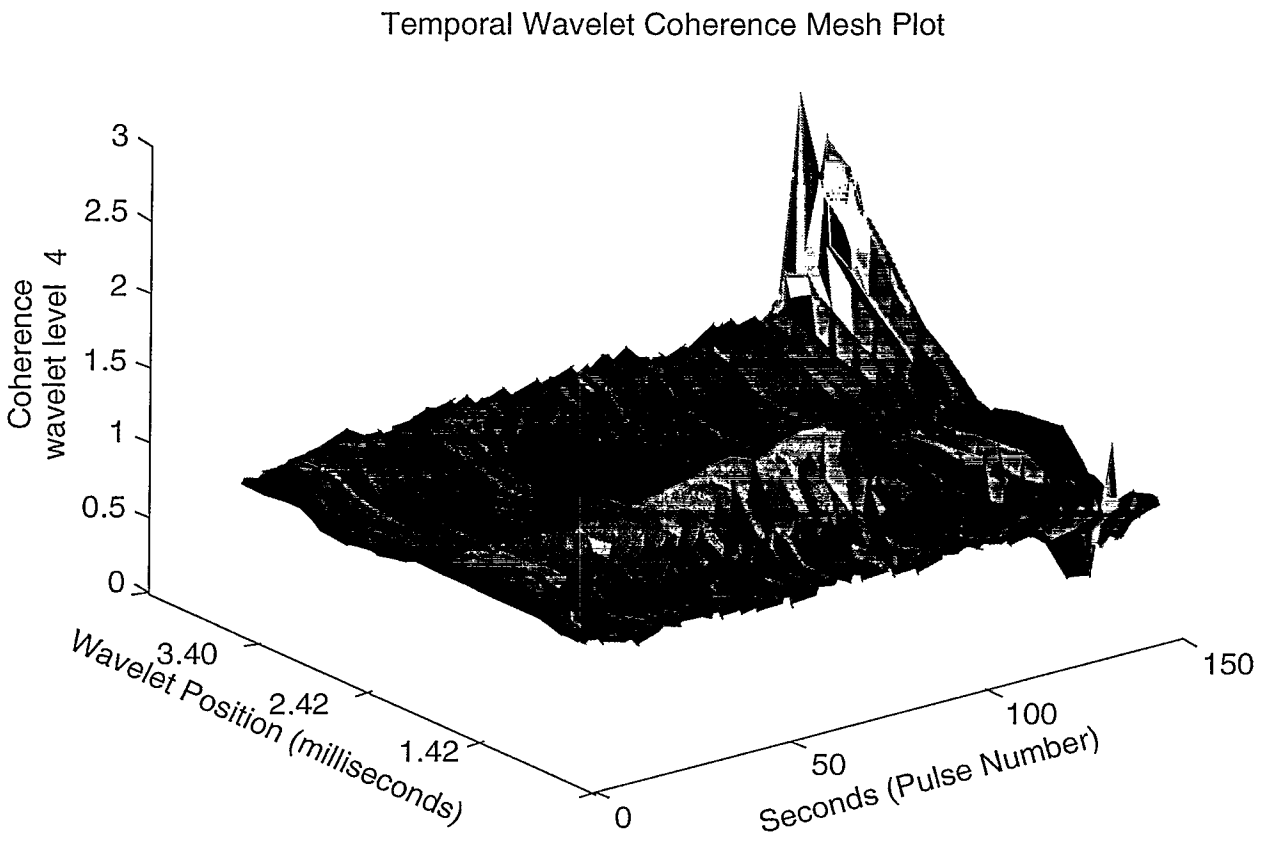

Figure 11: Three-dimension coherence surface from the wavelet mean-square-energy map for wavelet level four as a function of wavelet position and pulse number.

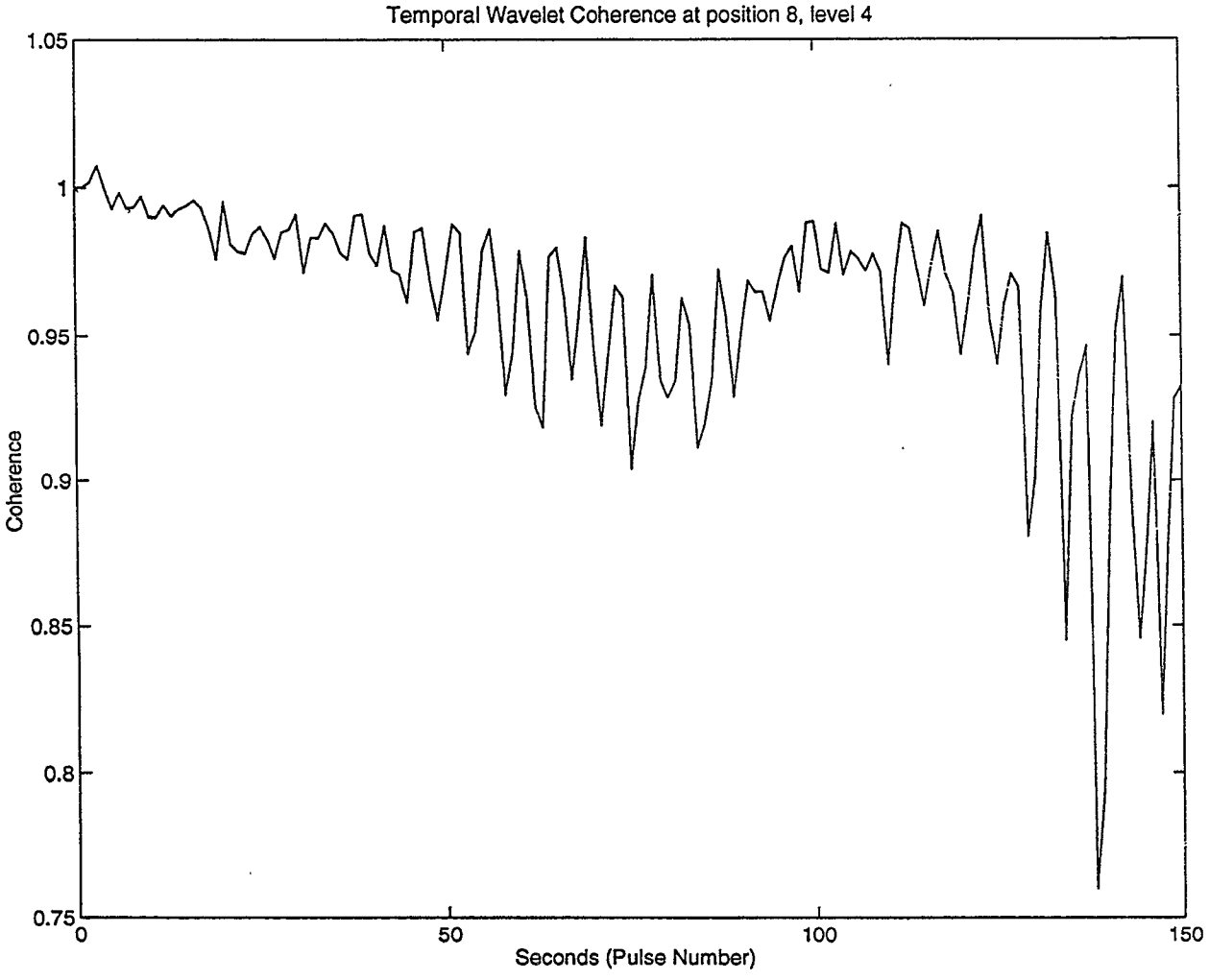

Figure 12: Coherence as a function of pulse number for wavelet level 4 and wavelet position 8. This is one slice of the surface in Figure 11. Compare to Figures 4 and 7.

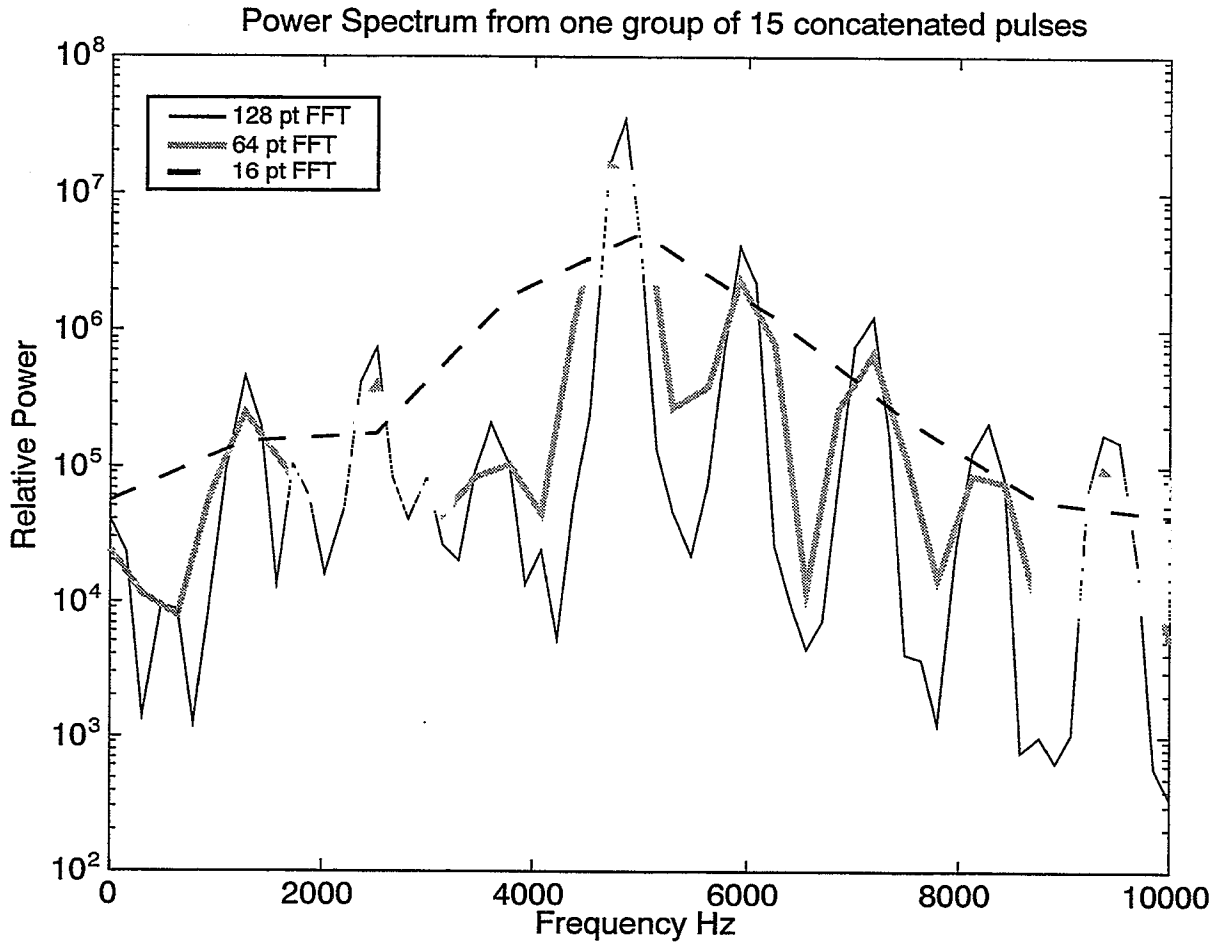

Figure 13: Power spectral density for the first group of 15 concatenated pulses using 128-point, 64-point, and 16-point FFT sizes, with 50% overlap, and a Hanning window. Compare to Figure 2.

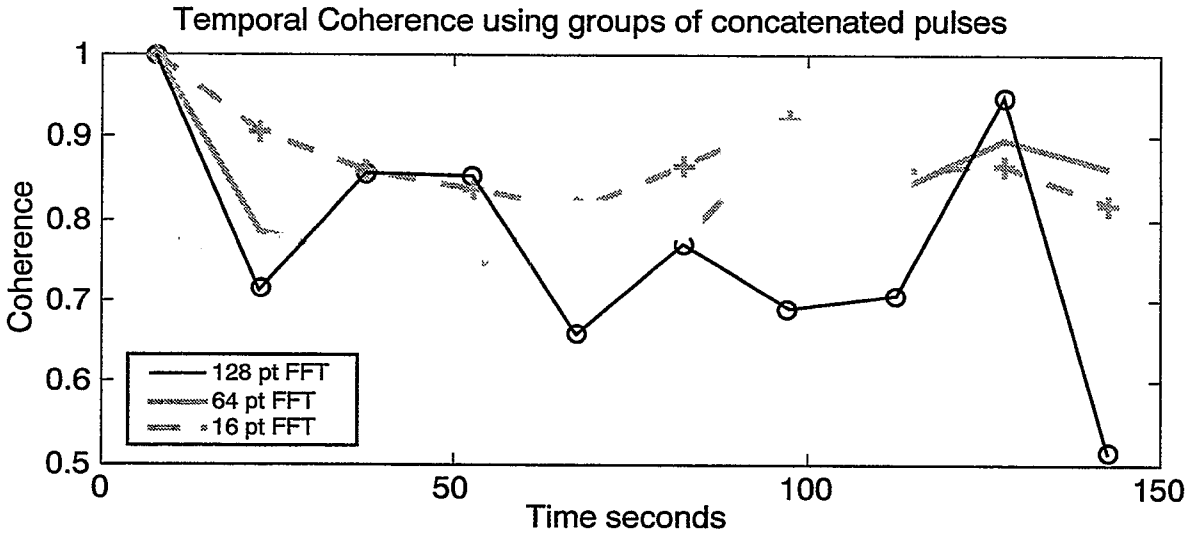

Figure 14: Coherence as a function of pulse number using 10 groups of 15 concatenated pulses. The pulse repetition rate is 1 pulse per second. Compare to Figures 4, 7, and 12.