The Effect of Coded Signals on the Precision of Autonomous Underwater Vehicle Acoustic Navigation

by

Benjamin Kerbin Evans

B.A., Physics, Williams College (1996)

Submitted in partial fullfillment of the
requirements for the degree of

OCEAN ENGINEER

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

September 1999

Author .....
Joint Program in Oceanography and Applied Ocean Science and Engineering
Massachusetts Institute of Technology/Woods Hole Oceanographic Institution
August 6, 1999

Certified by .....
Prof. Henrik Schmidt
Professor of Ocean Engineering, MIT
Thesis Supervisor

Certified by .....
Mr. Thomas Austin
Senior Engineer, WHOI
WHOI Advisor

Accepted by .....
Prof. Michael Triantafyllou
Chairman, Joint Committee for Applied Ocean Science and Engineering
Massachusetts Institute of Technology/Woods Hole Oceanographic Institution

The Effect of Coded Signals on the Precision of Autonomous Underwater Vehicle Acoustic Navigation

by

Benjamin Kerbin Evans

Submitted on August 6, 1999, in partial fullfillment
of the requirements for the degree of
OCEAN ENGINEER
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
and the
WOODS HOLE OCEANOGRAPHIC INSTITUTION
September 1999

Abstract

Acoustic coded signaling offers potentially significant improvements over traditional "toneburst" methods in many underwater applications where error due to noise and multipath interference is a problem. In this thesis, the use of these spread spectrum techniques is analyzed for navigation of the REMUS autonomous underwater vehicle. The accuracy of the current system using Turyn and Barker sequences, as well as toneburst, is quantified, and the sources of the remaining error are examined.

Thesis Supervisor: Prof. Henrik Schmidt
Professor of Ocean Engineering, MIT

WHOI Advisor: Mr. Thomas Austin
Senior Engineer, WHOI

Acknowledgments

I would like to thank Professor Henrik Schmidt for serving as my advisor and for supporting my graduate work in the Joint Program. I would also like to thank Mr. Thomas Austin for acting as my advisor in Woods Hole, and the rest of the OSL engineers, in particular Roger Stokey and Christopher von Alt, for their advice and assistance over the three years I have worked with their group. In addition, I would like to acknowledge Professor Jerome Milgram, who served as my interim advisor during Professor Schmidt's sabbatical and provided much of the equipment used in the experiments reported here.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

In the past decade, rising awareness of the importance of the coastal environment and man's effect on it, as well as the ongoing shift in naval strategy toward littoral warfare, has created demand for better understanding of the coastal zone. This need for more accurate and extensive near shore data has precipitated a transition of priorities in ocean science and engineering research away from deep water, open ocean projects to an increasing emphasis on near shore research.

During this same period, the capabilities of Autonomous Underwater Vehicles (AUVs) have been greatly improved and their costs decreased. Numerous advances have been made in the areas of power consumption, battery capacity, underwater docking, interchangeable instrument payloads, on board data storage and processing power, physical size, and general "user friendliness" which have led to the development of a generation of vehicles ideally suited for use in the coastal environment. Vehicles such as REMUS (Remote Environmental Monitoring UnitS), developed at the Woods Hole Oceanographic Institution Ocean Systems Laboratory, are small enough to be deployed by a crew of two or three from a small boat and easily programmed from a laptop computer, yet can carry a wide range of instruments, conduct missions as long as 18 hours, and dock at an underwater station to upload data, recharge batteries, and download new mission instructions [18]. These capabilities and ease of use, combined with comparatively low cost, make small AUVs a useful tool for conducting nearly real time coastal environmental surveys [19].

However, despite the recent improvements in AUV technology, navigation of these ve-

hicles has remained a major problem. Inertial navigation systems such as those used on manned submarines and larger AUVs are generally too large, heavy, and expensive for installation on small vehicles such as REMUS, and while the advent of the Global Positioning System (GPS) has revolutionized position finding on the surface and in the air, the basic physics of electro-magnetic wave propagation prevents the signals on which this system depends from penetrating the ocean to useful depths. While there have been several novel attempts to use GPS for underwater navigation by towing a surface antenna or surfacing frequently to obtain a fix, in general acoustics have proven to be the most accurate and practical means of underwater position finding for AUVs. Sound propagates efficiently in water due to its density, but is subject to sources of interference which can introduce severe error into acoustic navigation systems, particularly in shallow water. This lack of positioning accuracy presents the most serious current problem in AUV operations.

1.2 Previous Work

Acoustic navigation systems have been in use for many years in deep water, but shallow water has presented challenges that complicate the application of "off the shelf" open water systems in the coastal zone. In deep water, both the acoustic source and receiver are generally far from the surface and the bottom, and the sound propagates in the relatively homogeneous water in between, free from the interference caused by reflections from these surfaces. However, in shallow water, the proximity of the water's boundaries can result in strong reflections which make distinguishing the "direct path" arrival at the receiver from the reflected "multipath" arrivals difficult. In addition, rapidly changing bathymetry along the coast can lead to regions where the direct path from source to receiver is blocked by bottom or shoreline features, and the dynamic coastal environment often generates vertical stratification in the water column which further reflects and refracts sound. Also, the level of ambient noise, both natural and man made, is generally much higher in the coastal zone than the open ocean, and obviously shortens the maximum range at which a signal can be detected. Since the direct path arrival is generally most useful for determining the range from sources to receivers and generating a fix, any phenomenon which degrades the system's ability to identify the direct path has an adverse impact on navigational accuracy and precision.

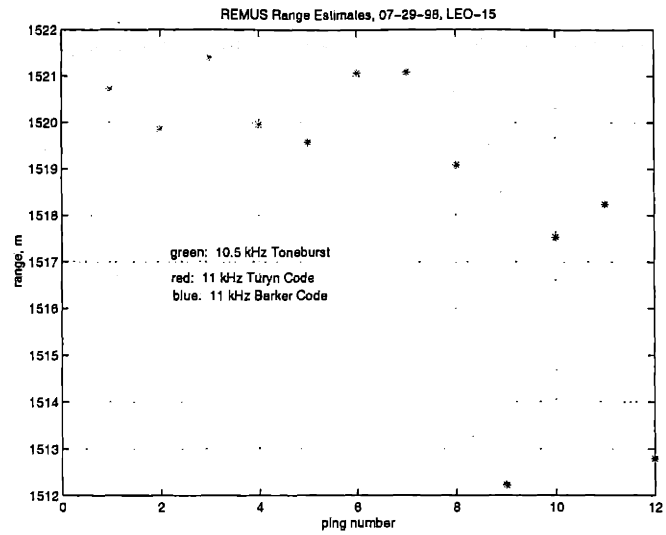

Figure 1-1: REMUS Range Estimates at LEO-15

As will be discussed further in Chapter 2, a signal with increased bandwidth is more easily localized in time, and therefore the direct path and multipath arrivals are more easily distinguished. One means of accomplishing this is to shorten the signal, but this has the unwanted additional affect of limiting the range of propagation since a shorter signal allows for less energy to be transfered to the water. Alternately, a "spread spectrum" coded signal of the same duration as the original increases bandwidth and therefore time resolution while maintaining range.

Computer simulation suggests that a system based on these coded signals would be significantly more robust in the presence of multipath interference and ambient noise than an otherwise equivalent toneburst scheme. However, when implemented on REMUS during field trials at LEO-15[1] in July 1998, coded signaling navigation appeared to be less precise than this modeling predicted. As shown in Figure 1.1, range estimates from a stationary vehicle to a moored transponder using coded signals were no more precise than those generated from a standard toneburst.

While four samples of each waveform hardly constitutes a statistically significant result, clearly investigation of this surprising discrepancy between prediction and practice was warranted. The research reported in this thesis was undertaken to bridge this gap with the

[1]**The Long term Environmental Observatory at 15 meters, located off the Rutgers University Marine Field Station in Tuckerton, NJ**

first quantitative field testing of underwater acoustic coded signaling in order to determine sources of the error observed at LEO and compare the precision of toneburst and coded acoustic navigation systems.

Incidentally, while the range estimation error at LEO-15 provided the impetus for this research, the source of this error was eventually found to be the probable result of a problem in the preamplifier hardware, rather than the properties of the acoustic waveguide. While not directly related to the primary focus of this thesis, for completeness a summary of this investigation is included in Appendix B.

1.3 Contributions of this Thesis

This primary contributions of this thesis are as follows:

1. The first extensive testing of the application of spread spectrum coded signaling techniques to underwater navigation under operational conditions was conducted. A series of tests generated thousands of samples with different code sequences and frequencies, the results of which are presented and analyzed.

2. The sources of the residual error are identified and their potential contribution to navigational inaccuracy assessed by both field testing and numerical modeling simulation.

3. The tradeoff between resolution and robustness of the signal waveforms is examined, in the context of vehicle hardware and variable properties of the acoustic waveguide.

In addition, while this thesis concentrates exclusively on the problem of AUV navigation, the spread spectrum techniques employed and the acoustic propagation results described are also applicable to the area of underwater acoustic communications. This field is also currently the subject of considerable research [7].

1.4 Thesis Organization

This thesis is divided into five chapters, including this introduction. Chapter 2 discusses the theoretical background of acoustic navigation and the coded signaling techniques employed in this thesis. The actual experimental method and apparatus are described in Chapter

3. Chapter 4 presents the data collected, and describes the analysis conducted, including the propagation modeling employed, to assess the contribution of each potential source of error. Finally, Chapter 5 presents the results and conclusions drawn from this work, and discusses the unresolved issues raised. The appendices contain a note on modifications made to REMUS hardware and presentation of data not shown in Chapter 4.

Chapter 2

Background and Theory

2.1 Acoustic Navigation Primer

Although actual means of position finding is not the primary focus of this thesis, some discussion of the way in which a fix is generated from acoustic signals is appropriate. This section describes the location schemes most commonly employed in acoustic navigation. It should be noted that while these general methods are discussed in the context of acoustics, they can be implemented by other means, such as the radio signals used in LORAN and GPS. In all cases, however, accurate determination of the signals' time of arrival is crucial.

A discussion of navigation techniques would be incomplete without mention of dead reckoning, the most basic method of position finding. Dead reckoning allows the navigator to estimate current position from knowledge of an initial fix, time since that fix, and course and speed over that time. While the means of determining the course and speed can be as simple as a magnetic compass and counting propeller turns or as complex as gyro compasses, accelerometers, and bottom Doppler locked sonar, the fundamental principal has been in use for hundreds of years. Despite its simplicity, most vehicles, whether in the water, on the surface, or in the air, rely on dead reckoning for position finding between fixes by more accurate means.

2.1.1 Long Baseline Navigation

Long baseline navigation ("LBL") is conceptually the most simple method of generating a fix. The "baseline" referred to is a line or array of points whose positions are known. By

finding the location of the vehicle relative to these points, the position in global coordinates can be determined. In an underwater acoustic navigation system, these points are sound sources fixed on the bottom or attached to moorings. Depending on knowledge of the timing of these signals, LBL systems can usually be classified as either spherical or hyperbolic.

Spherical LBL

In spherical LBL, the range from the vehicle to several transponders is estimated, and position within the transponder array is then determined by triangulation. Range calculation requires accurate knowledge of the propagation speed and travel time of the signal. One means of determining this time is to have perfectly synchronized clocks at the source and receiver, and have the source broadcast at predetermined intervals. While this can be accomplished by using either extremely accurate clocks at the source and receiver, or using additional signals to correct a less accurate receiver clock, this is not a practical solution for AUV navigation where the expense and complexity of either accurate clocks or additional signals would be prohibitive. In underwater acoustic navigation, the sound speed can be readily determined, and travel time is usually found by interrogating the transponders with a ping from the vehicle, and measuring the time required to receive replies. Range to each transponder is then simply calculated according to,

$$r = \left(\frac{t-\tau}{2}\right) c \tag{2.1}$$

where r is the range to the transponder, t is the total round trip time from interrogation to receipt of the reply, τ is the sum of system delays such as the time taken by the transponder to reply to a received signal, and c is the speed of sound in water (typically 1500 m/s, but dependent on temperature and salinity as discussed later). Once this range is determined, the vehicle is known to lie somewhere on the surface of a sphere of radius r surrounding the transponder. In the general case, four transponder ranges are required for a three dimensional fix, as a point is uniquely determined by the intersection of four spheres.[1] However, by being slightly clever in assumptions about which side of a line array the vehicle will always operate on, positioning all the transponders at the same known depth, and using

[1]The intersection of two spheres forms a circle. Adding a third sphere reduces the possible position to two points on that circle, and the fourth determines which of the points is the fix.

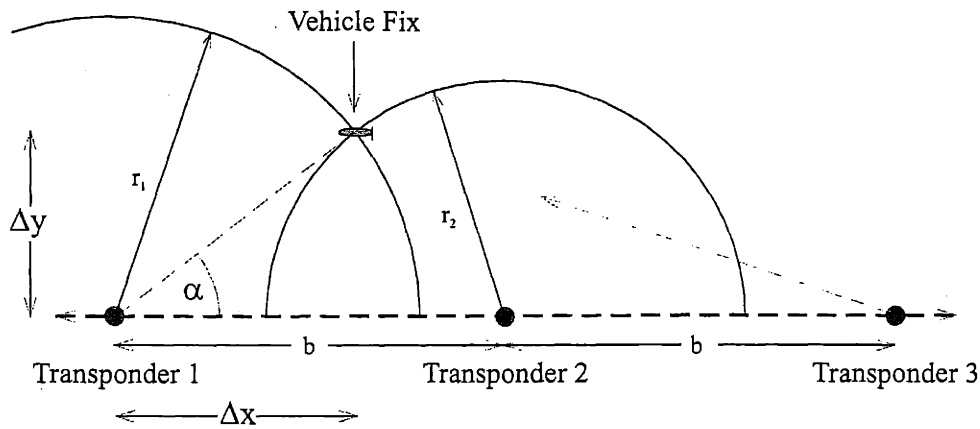

Figure 2-1: Long Baseline Navigation: In this arrangement, it is assumed that the vehicle will always operate above the transponder baseline, and will determine depth independently, so only two ranges are required for a fix. As shown the vehicle is at the intersection of the circular lines of position generated by ranges r_1 and r_2, as transponder three is out of range.

independent instrumentation to determine vehicle depth, in practice the problem can be reduced from the intersection of four spheres to the crossing point of two semi-circles, as shown in Figure 2-1. With two ranges and the baseline length between the transponders, the angles of the triangle formed by the transponders and vehicle can be found by the Law of Cosines, and the position of the receiver determined by trigonometry. For example, for the arrangement shown in Figure 2-1, position relative to transponder 1 can be calculated as follows:

$$\alpha = \arccos\left(\frac{r_1^2 + b^2 - r_2^2}{2r_1 b}\right) \tag{2.2}$$

$$\Delta x = r_1 \cos\alpha \tag{2.3}$$

$$\Delta y = r_1 \sin\alpha \tag{2.4}$$

Alternately, the position can be calculated by solving the system of equations generated by the two circular lines of position (LOPs). This results in two solutions, one of which is rejected because it falls on the wrong side of the baseline.

Despite its simplicity, spherical navigation has several disadvantages. First and most

serious is the need to know the exact time at which the transponder signal was transmitted. While systems such as GPS have sufficient resources to synchronize source and receiver times with atomic clocks at the transmitters and an extra signal to correct the inaccurate clock at the receiver, this is usually impractical for small, low cost AUVs because of the additional equipment and processing required. However, the system of interrogation and response described above as a substitute for source synchronization introduces several additional potential errors to the range estimate. This method requires two successful detections and arrival identifications instead of just one, both at the transponder and the vehicle, doubling the potential for error at this stage of the location process. In addition, the transponder delay must be accurately determined and removed from the round trip travel time.

Spherical LBL also presents practical problems for multiple receiver operations. If each transponder must be individually interrogated by each vehicle in order to determine ranges, a system to identify each vehicle's range response must be implemented to keep order. While it is easy to conceive of a table of different codes or frequencies to identify which replies are intended for which receivers, such a scheme quickly adds complexity to the system. For this reason, systems with multiple receivers generally rely on hyperbolic navigation systems.

Hyperbolic LBL

Unlike spherical LBL, hyperbolic navigation requires only that all the source signals be transmitted at the same time (or offset by a known delay), not that the exact time of transmission be known at the receiver. The receiver then measures the time delays between the received signals, and from this the vehicle's position can be determined. Since the receiver is passive and does not need to have its clock synchronized with those of the transmitters, hyperbolic navigation is useful for systems of multiple receivers all using the same beacons.

Again using the propagation speed of the signals, the time delay between arrivals from two transmitters can be converted to a range difference. In three dimensions, the locus of points for which the range to the two beacons is equal to this difference defines a hyperboloid, but as in the spherical case, the problem can be reduced to a plane by determining the depth of the transmitters and receiver independently. The receiver must then lie somewhere along the hyperbolic line of position generated by the time delay between the two transmitters,

and a third beacon is required to generate a second LOP crossing the first for a fix.[2]

While it does offer simpler timing, hyperbolic LBL is less often used for underwater acoustic navigation than spherical because it requires a larger array of transponders to generate fixes of equivalent accuracy. Spherical LBL works quite well with just a line of transponders, as only two signals are required to generate a fix. However, hyperbolic LBL requires a third transponder, preferably off the baseline created by the first two. Navigation errors are minimized when position is estimated from LOPs which are nearly perpendicular. While this ideal is easily attained along the spherical LBL baseline when the vehicle is operating at a distance from the line of roughly half the spacing of the transponders (as in Figure 2-1), with hyperbolic LOPs this optimal area of nearly perpendicular crossings is much smaller, and deteriorates much more quickly as the vehicle moves outside the area confined by the transponder array[3].

2.1.2 Ultrashort Baseline Navigation

Unlike either of the LBL schemes discussed above, Ultrashort Baseline navigation (USBL) requires only one transponder to compute position. Range to the transponder is determined by the spherical method described above, and angle to the transponder is calculated by comparing the phase of the arriving signal at individual elements of a receiver array on the vehicle.

Figure 2-2 shows a simple USBL configuration in one axis. Typically, a vehicle would mount a two dimensional array with a second line of hydrophones perpendicular to those shown to determine both components of the angle to the transponder. Note that the array element spacing, d, must be less than half the wavelength of the carrier signal to avoid spatial aliasing. The bearing to the transponder in each plane is calculated according to,

$$\theta = \arcsin\left(\frac{\phi\lambda}{2\pi d}\right) \tag{2.5}$$

where ϕ is the phase angle difference between the hydrophones. Once the bearing and elevation of the incoming ray relative to the receiving array have been calculated, data from

[2]Many land based radio navigation systems operate by this method. For example, the hyperbolic LOPs generated by LORAN-C can be seen overprinted on most nautical charts, labeled with their corresponding time delays from the beacons of the local chain.

[3]It is possible to hybridize the spherical and hyperbolic systems if the receiver has some, but not perfect, knowledge of when the signals were sent in addition to the time delays between their arrivals. See [4].

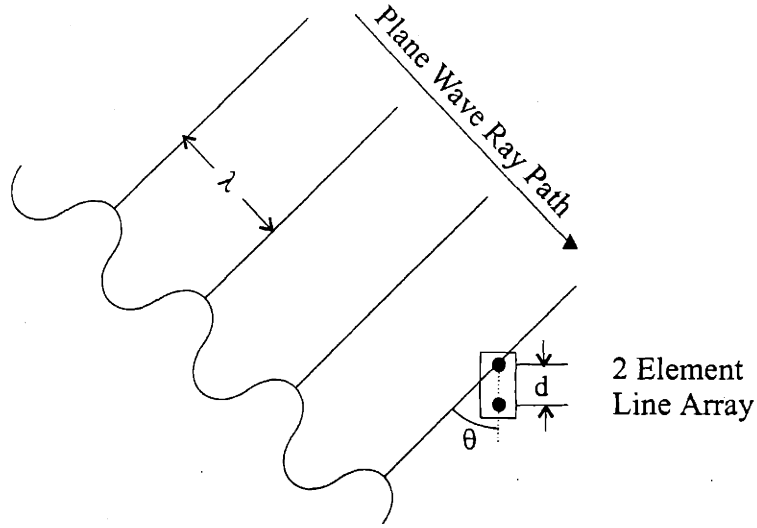

Figure 2-2: USBL Navigation in two dimensions

onboard attitude sensors is used to correct for the pitch, roll, and heading of the vehicle to give the range, bearing, and elevation to the transponder and generate a fix.

USBL is a useful technique for precision short range navigation such as docking where range and bearing to a specific target is important, or in situations where deployment of a more extensive array is not possible. However, accuracy of USBL degrades rapidly as range increases, as even a small error in the bearing measurement can lead to large uncertainty in position.

2.2 REMUS Basics

As the research described in this thesis was conducted within the framework of REMUS equipment and operations, some familiarity with the vehicle will help to explain the work which follows.

2.2.1 Vehicle Overview

REMUS is a relatively small, low cost AUV designed for coastal survey work. In its basic configuration, the vehicle is 52 inches long and 7.5 inches in diameter, weighs 68 pounds in air, and is easily handled by a crew of two from a small boat. Depending on the instrumentation and batteries installed, the vehicle can travel as far as 164 km on a single charge,

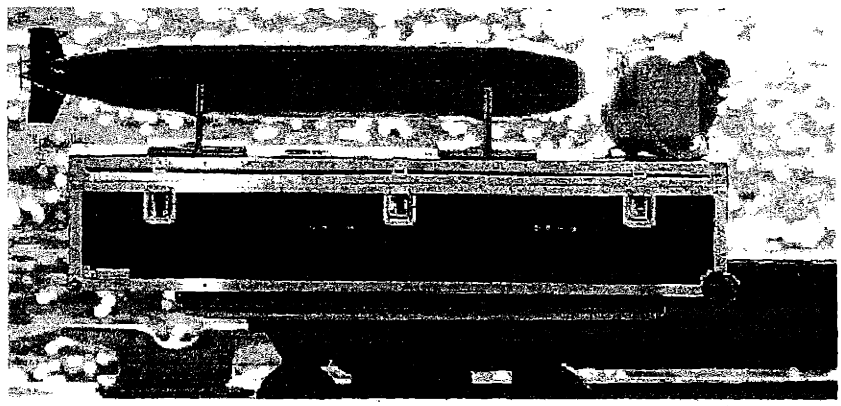

Figure 2-3: REMUS Vehicle with Transponder

and typically cruises at two to three knots. This endurance has been greatly increased by the development of a companion docking station, which allows REMUS to recharge its batteries, upload data, and download new mission programming without surface recovery.

REMUS is a flexible platform which can carry a wide array of instruments, including an acoustic Doppler current profiler (ADCP), conductivity-temperature-depth sensor (CTD), side scan sonar, and optical backscatter and dissolved oxygen sensors. The data is logged in the onboard computer system which is based on the PC-104 architecture. Once docked or connected to a computer on land, data is retrieved and the vehicle programmed via a graphical user interface which runs under Microsoft Windows.

2.2.2 Navigation Systems

REMUS can be programmed to navigate purely by dead reckoning, or by generating fixes from acoustic transponders in either LBL or USBL mode. The degree of accuracy required and the convenience of deploying additional equipment determine which navigation mode is used for a specific mission.

The REMUS dead reckoning system is essentially simple, but can be augmented by several more precise sensors. With the standard sensor package, heading and attitude are determined from a fluxgate digital compass with an integral tilt sensor, depth is measured by a pressure sensor, and distance traveled is estimated by counting turns of the propeller. On more extensively equipped versions of the vehicle, sonar altimeters can be installed to measure the distance to the bottom, speed over ground can be determined from the ADCP

Figure 2-4: REMUS Transducer and USBL Nose Array

data, and rate sensors report high frequency changes in vehicle attitude and velocity.

Acoustic navigation on REMUS has also evolved from a fairly simple system to one that is more complex and supports a wider array of modes. As first developed, REMUS relied solely on USBL navigation, using uncoded toneburst signals in a frequency band centered at 27 kHz. Figure 2-4 shows the transducer under the nose of REMUS which interrogates the transponder, and the nose array which receives the reply. Outgoing signals are encoded on an EPROM[4], and generated by an arbitrary transmitter board. Received signals are preamplified, and fed to a dedicated PC-104 Digital Signal Processor (DSP), which processes the signals as described in the next section. Later development included the introduction of the coded signals mentioned in Chapter 1 and discussed in detail below. Because the codes were originally introduced to improve the accuracy of the bearing estimate, the transponder transmits a coded waveform, but the interrogation pulse (which has no impact on the angle calculated) remains uncoded. As a result, while bearing and elevation estimates were improved, any error in the detection of the interrogation pulse at the buoy will still impact the range calculation despite the more accurate timing at the vehicle. Current plans to install the DSPs necessary to decode the spread spectrum waveforms on the transponders could cut range estimate error considerably, since it would allow round trip use of coded signals instead of one way.

As discussed previously, USBL navigation is most accurate at relatively short ranges, so for longer range missions, an LBL system was developed. This can operate in either the same frequency range as the original USBL system, or if even greater range is required lower frequency ($\sim$12 kHz) transponders and vehicle transducers can be used. (In this frequency

[4]Erasable Programable Read Only Memory

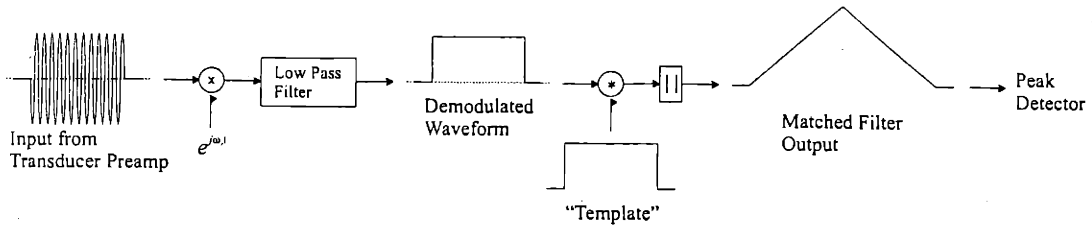

Figure 2-5: Complex Demodulator and Matched Filter

band, acoustic attenuation in seawater is approximately proportionate to the square of the frequency.) In addition, a LF USBL system is currently being implemented to complete the range of options available for REMUS navigation.

The transponders used for REMUS operations are all acoustically similar, but packaged and deployed differently. The first generation of transponders were the off the shelf units shown in 2-3. Each was programmed to a specific frequency and waveform. The spheres were slightly buoyant, so they were anchored at approximately the same depth as REMUS was to operate, and a surface float was attached for recovery. The transponders currently in use are much larger, with the electronics packaged in a box on the surface expression and the transducer suspended below. These transponders also have a radio link to shore, which allows them to interrogate the vehicle and report the range. In this way, the vehicle navigates relative to the buoys, but the buoys can also locate the vehicle, allowing operators ashore to track its progress.

2.3 Signal Processing

No matter what navigation scheme is employed, precise timing of arriving signals is critical for navigational accuracy. This section describes the process by which the arrival of signals from the transponder is timed, and the impact waveform shape has on this process.

2.3.1 Detection and Matched Filtering

Figure 2-5 shows a simplified and idealized block diagram of the processing by the REMUS DSP which gives the time of arrival of the incoming signal from the transponder. The first stage is a standard complex demodulation. The input signal (after several stages of gain and bandpass filtering which are not shown), represented here without any noise, is multiplied

by a complex sinusoid at the carrier frequency. After low pass filtering to remove the copies of the signal at twice the carrier frequency which result from the frequency domain convolution of demodulation, the output is the signal which originally modulated the carrier at the transmitter. The simple but unrealistic square "boxcar" waveform is shown in the example. Although it is shown as a real signal, in general it also has an imaginary component which is preserved throughout the complex demodulation and matched filtering process.

The output of the demodulator is then input to the matched filter, which is the key to accurate timing of the signal's arrival. Matched filtering relies on the convolution operation, which calculates the cross-correlation between two input signals. This allows the receiver to search for a specific waveform in the presence of noise, and is a more reliable arrival time estimator than a simple threshold detector.

The matched filter has two inputs: the demodulated signal, and a "template" waveform to be detected in the demodulated signal. This template is a time reversed (since the convolution operator flips the signal back) copy of the waveform which originally modulated the carrier at the receiver. Convolution slides the flipped template past the input signal, multiplying the two waveforms and integrating over time at each step. In the example shown, the matched filter input may be complex, so the absolute value of the output is calculated to combine the contribution of the real and imaginary parts. The result is a signal as long as the sum of the input and template, and has a maximum at the point in time of the two signals' greatest cross-correlation, which should be the time of arrival.

In linear systems theory, a matched filter is defined as follows[14]:

Given the following linear, time invariant system,

$$y(t) = x(t) * h(t) \tag{2.6}$$

where $h(t)$ is the impulse response of the filter, $h(t)$ is a matched filter if

$$h(t) = kx(\Delta - t) \tag{2.7}$$

where k and Δ are arbitrary constants. Carrying through the complex convolution operation in 2.6 and substituting from 2.7 results in the autocorrelation function of x:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h^*(t-\tau)d\tau \tag{2.8}$$

$$= \int_{-\infty}^{\infty} x(\tau)x^*(\tau-t)d\tau \tag{2.9}$$

$$= \phi_{xx}(\tau) \tag{2.10}$$

One property of a matched filter is that for a given input signal, it is the linear filter which results in the maximum signal to noise ratio at the output. For the output of a matched filter, it can be shown that the maximum SNR is given by [21, 14],

$$SNR_{out} = \frac{2E}{N_0} \tag{2.11}$$

where E is the energy of the input signal $s(t)$ of duration T,

$$E = \int_0^T s(t)s^*(t)dt \tag{2.12}$$

and N_0 is the power density of the input noise,

$$N_0 = \frac{E_N}{T_N BW_R} \tag{2.13}$$

where the noise energy, E_N is calculated according to 2.12 over the interval T_N, and BW_R is the bandwidth of the receiver input noise.

The ratio in 2.11 depends only on the energy of the signal and noise, and interestingly is not directly related to the bandwidth or peak power of the transmitted signal, so long as the filter is matched to its waveform. As discussed later in this chapter, however, these factors do affect the resolution with which a signal's arrival can be timed.

Further, the output and input SNRs are related by [14],

$$SNR_{out} = 2B_N T(SNR_{in}) \tag{2.14}$$

where B_N is the bandwidth of the matched filter, T is the duration of the signal, and $SNR_{in} = P_{in}/N_{in}$, the ratio of the input signal average power to the "in band" input noise density. Therefore, the matched filter improves the signal to noise ratio of the arrival detection by a factor of $2B_N T$.

In practice, this means that if the receiver has a copy of the waveform to be detected, a matched filter can calculate the cross-correlation between the received signal and the signal to be identified. The matched filter output will then be maximal at the time when the input most closely matches the "template", allowing the signal's arrival time to be resolved even with the addition of noise and other interference which would otherwise obscure the presence of the desired waveform. In the example shown in Figure 2-5, the arrival time of the signal in the input is obvious, but with the addition of noise or a more complex waveform, the output of the matched filter shows correlation which it would not have been possible to detect otherwise.

To actually determine this time of arrival, the matched filter output must be sent to a peak detector. In the ideal case, such as that shown, the time of arrival is simply the time of the maximum value of the output. However, with the addition of noise or other interfering signals, the matched filter can generate multiple peaks, the largest of which may not correspond to the time of the arrival of interest. The art of selecting the correct arrival time is discussed in Chapter 3.

2.3.2 Waveform Selection

Given the simplicity of the example described above, it may be difficult to imagine a case where a simple waveform approximating the boxcar would not result in accurate timing of the signal's arrival. However, the addition of noise and extraneous signals quickly degrades the performance of a system based on a toneburst waveform.

Figure 2-6 shows examples of matched filter output for three scenarios. The Matlab simulated input signals are 10 ms toneburst waveforms at 11 kHz and sampled at 50 kHz, identical to the low frequency system on REMUS. In the top line of plots, the input is an ideal signal without noise or other interference, which demodulates nearly perfectly.[5] The matched filter output is the clean triangle expected from the previous section (the autocorrelation), with its peak at 20 ms, the time of the signal's arrival.

The second line of plots also shows an ideal example. A delayed arrival approximately 3 dB down from the first is added to the original signal. This delay could be the result of multipath interference, mentioned in the introduction and discussed further in the next

[5]Note the small ripple on the square wave in the middle plot, a result of Gibbs phenomenon associated with the jump discontinuity. A truly square envelope such as this could not be generated as a real signal.

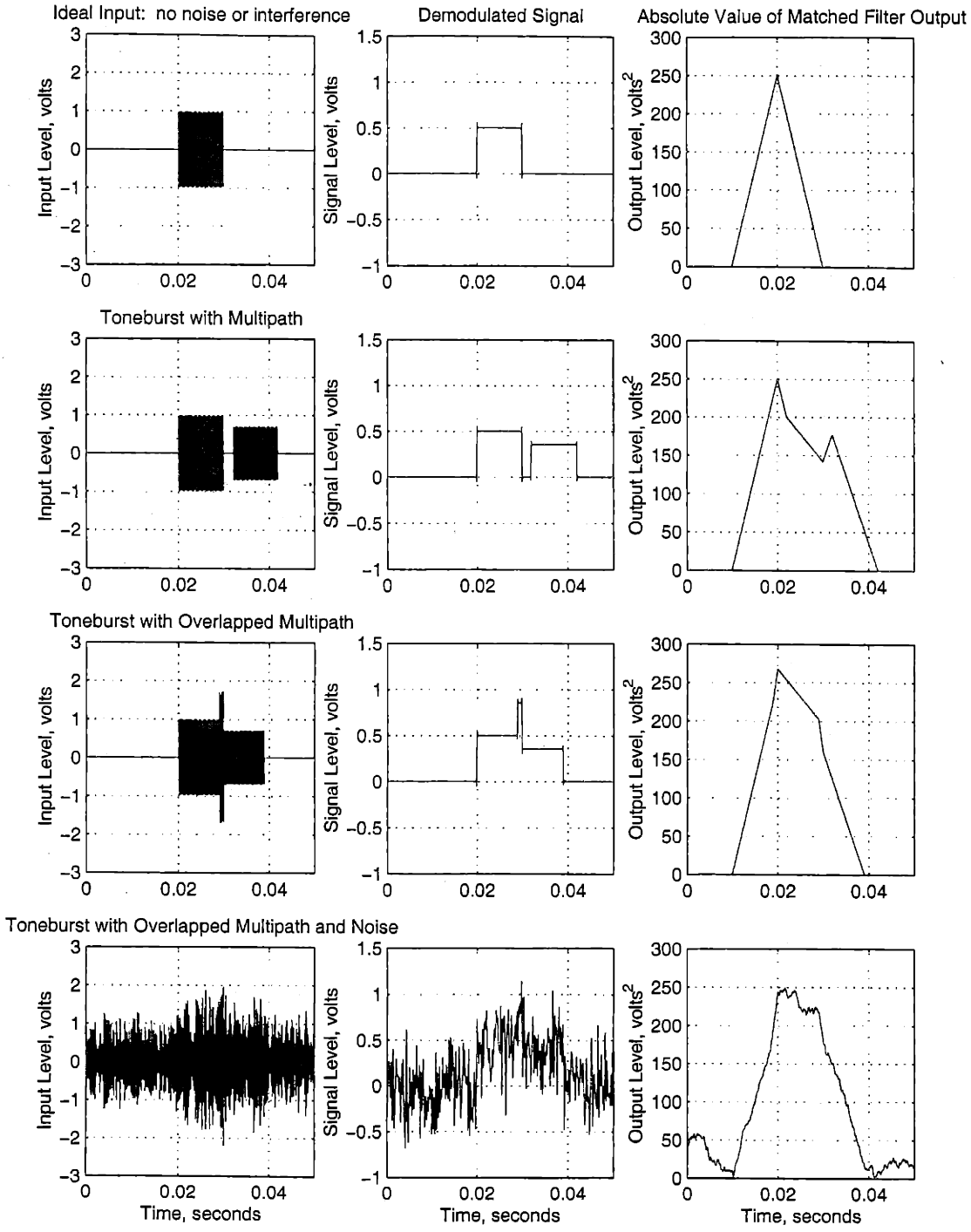

Figure 2-6: Simulated Matched Filter Outputs for 11kHz, 10ms Toneburst Signal

section. In this case, the second signal arrives 12 ms after the first, so there is no overlap. Again, the two signals demodulate cleanly, and while the matched filter overlaps the two triangles, the peaks are still distinct and correctly timed.

The following plots offer a more realistic example of signals which could be encountered in the field. The third line shows an example similar to that of the second, but with slightly less delay. The magnitude of the delayed arrival is the same as the second example, but it is only 9 ms behind the first, resulting in 1 ms of overlap.[6] In seawater, a 9 ms delay corresponds to approximately 13.5 m of additional pathlength, a realistic figure for a surface or bottom reflected ray at moderate range in relatively shallow water. The overlap is preserved in the demodulator output, and results in the loss of the distinct second arrival after matched filtering. However, the primary peak is still sharp enough to allow accurate timing of the first arrival.

Finally, the fourth example repeats the scenario of the third with the addition of noise. White noise across a 6 kHz band centered on the carrier frequency is added at a magnitude resulting in a signal to noise ratio (SNR) of approximately 3 dB. The effect on the matched filter output is obvious: the broad based triangle of the first three examples loses its sharp peak, making accurate timing of the first arrival impossible, and the shoulder corresponding to the second arrival in the third example is no longer identifiable. Accuracy would definitely suffer if this output was used for navigation.

Qualitatively, the solution to this problem appears to be shorter tonebursts. It is clear from the plots that shorter pulses would result in more narrow triangles in the matched filter output, giving sharper and better defined peaks for improved timing resolution in the presence of noise and multipath interference. This is the expected result from elementary Fourier theory: as a signal approaches an impulse in time, its spectral content becomes infinite, and therefore high bandwidth, short duration pulses result in the least ambiguity in time. However, shorter pulses of the same average power contain less energy, which according to 2.11 lowers the SNR and results in shorter effective range. Figure 2-7 demonstrates this with the same conditions as the bottom panel in 2-6, but with a toneburst one quarter the length. While the peaks in the matched filter output are distinct, their magnitude

[6] In this simulation the two signals are in phase because they were modulated by the same carrier, resulting in the constructive interference seen in the plot of the input signal. In the field this would obviously not always occur, as the phase of the delayed arrival would be randomized relative to that of the first by its additional pathlength and reflections.

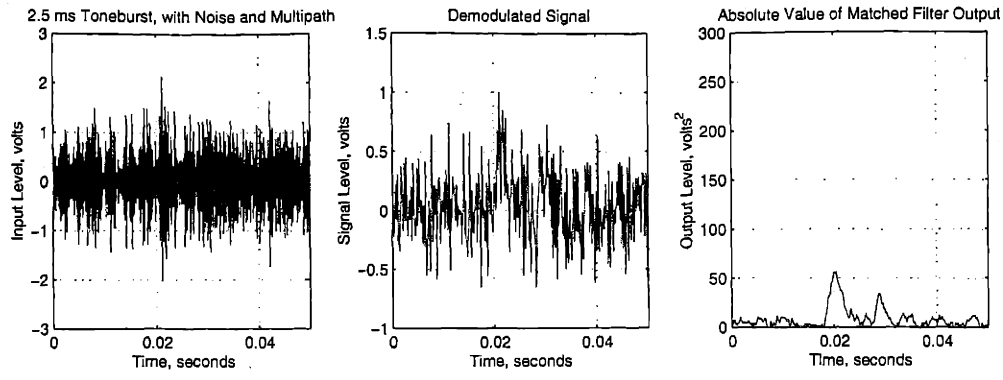

Figure 2-7: Simulated Matched Filter Outputs for 11kHz, 2.5ms Toneburst Signal

relative to the noise beneath is considerably less. It is this quandary which motivates the use of coded signals.

The trade off between range and resolution can be quantified by applying the Cramer-Rao bound, a fundamental limit on the precision of parameter estimation, to timing the arrival of signals. Parameter estimation theory states that the estimator (the system by which a value for the parameter in question is determined) which gives the unbiased result of minimum variance is unique, and the Cramer-Rao bound is the lower limit of this variance. For arrival time estimation, it can be shown that this limit has the following form [1, 21],

$$\sigma_\tau^2 \geq \frac{1}{\left(\frac{2E}{N_0}\beta_\tau^2\right)} \tag{2.15}$$

where σ_τ^2 is the variance of estimate of the arrival time in seconds squared, E is the energy of the input signal and N_0 is the noise density as given in 2.12 and 2.13, and β_τ is a measure of the bandwidth in radians defined as

$$\beta_\tau^2 = \frac{\int_{-\infty}^{\infty} \omega^2 |S(j\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |S(j\omega)|^2 d\omega} \tag{2.16}$$

in which $S(j\omega)$ is the Fourier transform of the input signal $s(t)$.

From 2.15 it is clear that assuming noise is constant, the variance of arrival timing, and therefore precision of range estimation, depends on the energy and bandwidth of the signal used. The trade off discussed above arises from the relationship between these two parameters. As expected from 2.12, the signal energy is directly proportional to the duration.

However, for the toneburst signals shown in the examples so far, bandwidth is proportionate to the inverse of duration, leaving the relationship

$$\sigma_\tau^2 \propto T \tag{2.17}$$

Therefore, in order to improve the accuracy of the range estimate, the energy, and therefore range, of the signal must be sacrificed. Coded signaling offers a way around this link between energy and bandwidth with waveform shapes whose bandwidths are not directly dependent on duration. This allows signals of long duration and high spectral content to be generated, resulting in low variance range estimates over long distances.

In general, coded sequences add bandwidth to a signal by modulating the carrier with a more complex waveform than the boxcar used for a toneburst (hence the term "spread spectrum"). Again, this is consistent with basic Fourier theory: any function in time can be expressed as the sum of sines and cosines, but a waveform with several transitions requires more frequency components than a simple boxcar gating function. Perhaps the simplest example of a wideband signal is the chirp waveform, which ramps frequency across the transmitter's bandwidth over the signal's duration. Other schemes separate the signal into "chips", segments which are differentiated from each other by phase (Binary Phase Shift Keying) or frequency (frequency hopping). In these sequences, dividing the signal into more chips further increases the bandwidth. In all cases, the goal is to produce a signal with an autocorrelation function which approaches that of an impulse, but retains the SNR of a long toneburst. The bandwidth of these coded sequences is not directly dependent on the duration of the signal, as it is for the tonebursts, so for these signals the relationship in 2.17 becomes

$$\sigma_\tau^2 \propto \frac{1}{T} \tag{2.18}$$

By using coded signals, the trade off between resolution and range is eliminated, and all aspects of performance are improved with longer signals.

The REMUS navigation system uses Binary Phase Shift Keying (BPSK), in which the toneburst signal is broken into chips, each with either 0 or 180 degree phase shift. The number of chips and sequence of these phase shifts is determined by a pseudo-random binary pulse compression code, a sequence of ones and zeroes selected for optimal autocorrelation

properties. These codes can be classified according to their peak autocorrelation sidelobe level. Sequences with a peak sidelobe level of 1 were investigated by Barker, who determined that 13 bits is the maximum length of a code with this property [2]. Likewise, Turyn found that the longest code with sidelobe level 2 is 28 bits [15]. The 13 chip Barker and 28 chip Turyn are the two sequences used for coded signaling on REMUS. Since the height of the autocorrelation main peak is proportionate to the number of chips, these codes of maximum length give the largest ratio of peak to sidelobe level.

$$\begin{aligned} \text{Barker} &= [0\,1\,0\,1\,0\,0\,1\,1\,0\,0\,0\,0\,0] \\ \text{Turyn} &= [1\,1\,1\,0\,0\,0\,0\,1\,1\,1\,0\,1\,1\,1\,0\,1\,1\,1\,0\,1\,1\,0\,1\,0\,0\,1\,0\,0] \end{aligned}$$

It has also been shown that significantly longer codes exist for higher peak sidelobe levels [6]. However, these long codes introduce a new tradeoff between range and resolution. Envelopes encoded with these sequences of the same length as those currently in use would have extremely short chips, extending the signal's bandwidth outside the range of the amplifier and transducer resonance curve. At the receiver, the additional bandwidth would again be clipped by the receiver input stages and the sampling rate of the DSP. The result could actually be degraded performance. On the other hand, extending the duration of the signal to stretch the chips to a more appropriate length would place a much heavier load on the vehicle power budget. Barker and Turyn Codes were selected because they most closely fill the available bandwidth without increasing power requirements.

To modulate a carrier wave, those chips corresponding to the zeros in the code are left unaltered, while those corresponding to the ones are reversed 180 degrees. This can be easily done by modulating the carrier wave with an envelope as long as the duration of the desired signal, with amplitude one for the chips equal to zero and minus one for the chips equal to one. Figure 2-8 shows the 13 chip Barker and 28 chip Turyn codes, and their autocorrelations, while 2-9 compares the bandwidth of the coded signals with that of the toneburst.

While Barker and Turyn codes offer superior autocorrelation properties compared to other forms of encoding, there exist only two (forward and reverse) for any given frequency and pulse length. This restricts the number of uncorrelated signals available for use at any one time, and would therefore limit the usefulness of these sequences in environments where

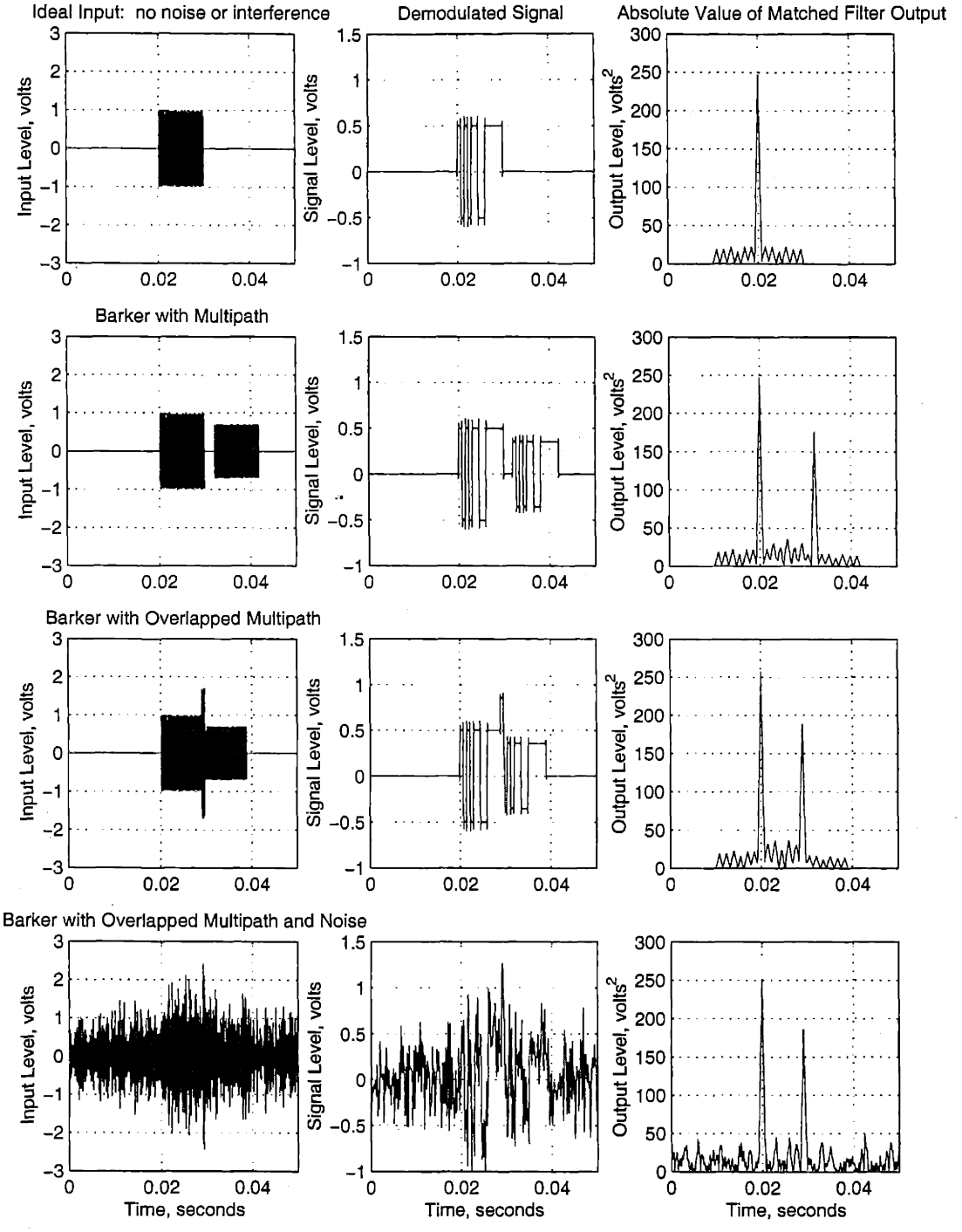

Figure 2-10: Simulated Matched Filter Outputs for 11kHz, 10ms 13 Chip Barker Coded Signal

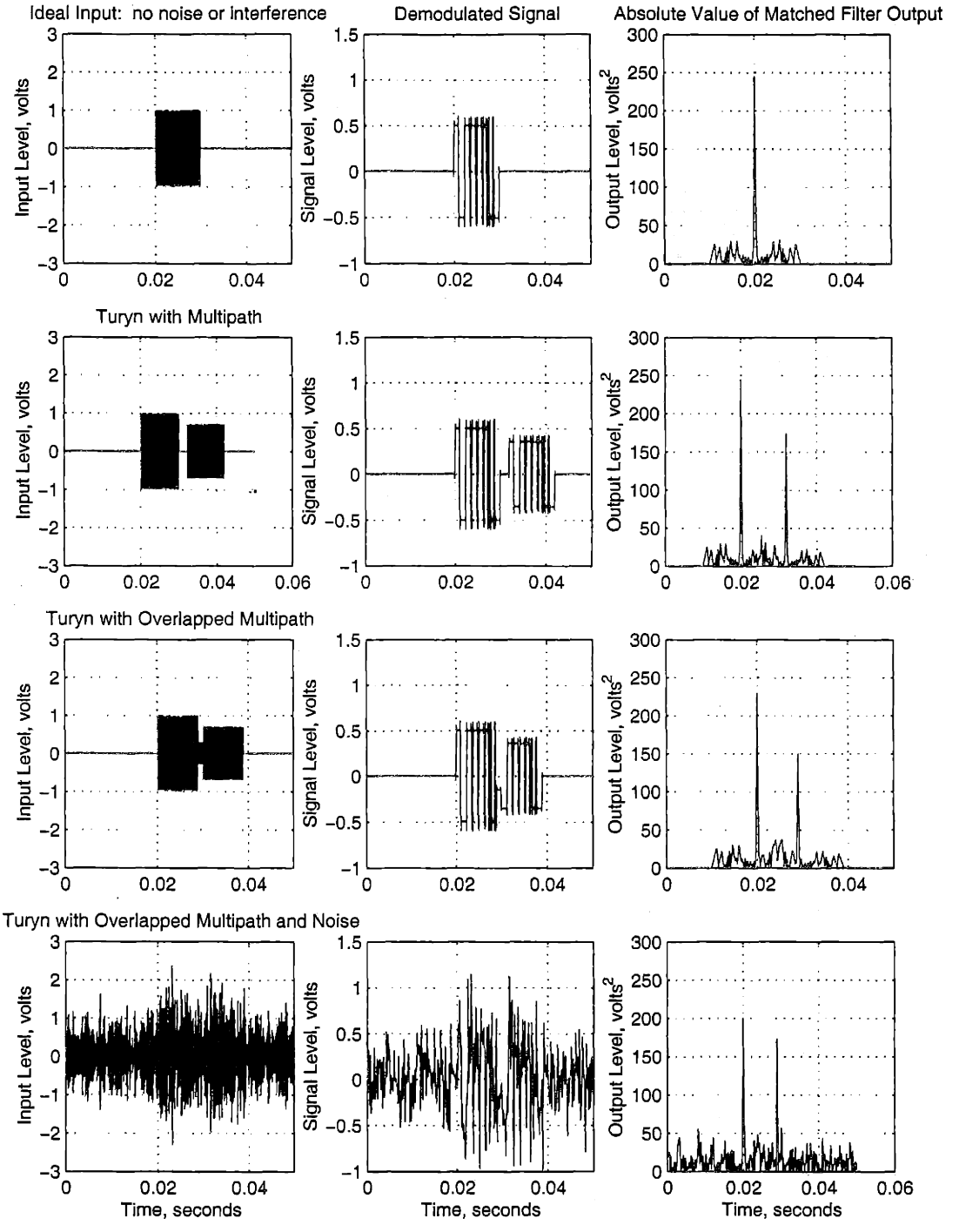

Figure 2-11: Simulated Matched Filter Outputs for 11kHz, 10ms 28 Chip Turyn Coded Signal

multiple sources and receivers were in concurrent operation. However, they have proven to be ideal for REMUS operations, which rarely involve more than one vehicle. The simplicity of generating BPSK modulated signals has made their integration into the existing hardware and software relatively straightforward.

Figures 2-10 and 2-11 demonstrate the improved performance these codes offer over tonebursts of the same length for the same conditions and scales shown in 2-6. Note the ease with which the matched filter identifies the arrival of the codes and the sharpness of the matched filter output in the presence of noise and multipath interference. This thesis seeks to quantify the extent to which these properties improve navigation in the field for REMUS.

2.4 Sources of Uncertainty

This section introduces the sources of potential error in arrival timing which have been identified. These factors can be separated into two basic categories: theoretical limitations on arrival time estimation, and environmental factors which cause error through changes in the acoustic propagation. The various components of these are discussed here, and the impact of each will be assessed as part of the analysis in Chapter 4.

2.4.1 Fundamental Constraints

Parameter estimation theory imposes theoretical limits on timing resolution as a function of the energy and bandwidth of the signal transmitted. As discussed earlier in this chapter, the Cramer-Rao bound defines the minimum parameter estimation error solely as a function of the signal transmitted. It is the "best case scenario" of performance, and assumes that the signal, with addition of noise, will arrive at the receiver as it was transmitted, and makes no account for errors in modulation, demodulation, or propagation. It is therefore a somewhat unrealistic target for the precision of a real navigation system, but does give an indication of the relative performance of different signals and the maximum precision which could be expected under ideal conditions. In this section we calculate the minimum variance of the signals used by REMUS according to this limit.

Computing the integrals required to arrive at a closed form, symbolic solution for 2.15 is difficult and time consuming, so the following results were computed numerically and graph-

ically under the same conditions as shown in Figures 2-6, 2-10, and 2-11. The bandwidth β_r of each signal was determined both by numerically solving 2.16, and by reading the width of the signal's spectrum 3 dB below the peak (the "full width at half maximum") from a plot of the Fourier transform. Both methods yielded reasonable and consistent results. The simulated noise was designed with a 6 kHz bandwidth, and the noise power density was computed according to 2.13.

	Arrival σ^2, seconds	Range σ, meters
Toneburst	3.55×10^{-9}	8.94×10^{-2}
Barker	1.44×10^{-11}	5.69×10^{-3}
Turyn	2.47×10^{-12}	2.36×10^{-3}

The above table gives the calculated values for the Cramer-Rao bounds of each waveform. The first column is the arrival time estimation variance, as given by 2.15, while the second column gives the equivalent standard deviation in range, assuming a sound speed of 1500 m/s. Since centimeter and millimeter level accuracy is far more precise than necessary, it is unlikely that acoustic navigation of REMUS will ultimately be constrained by these limits. These values do, however, quantify the relative accuracies of the available waveforms.

2.4.2 Environmental Constraints

Dispersion

Dispersion is a process common to many forms of wave propagation in which different frequency components of a signal travel at different speeds, causing the signal to arrive at the receiver "smeared" in frequency and time. This can be particularly problematic for wideband signals such as the Barker and Turyn codes used on REMUS, since accurate matched filtering relies on the presence of the signal's entire spectrum.

Intrinsic dispersion is caused by material properties of the transmitting medium and is most commonly observed spreading white light into its component colors in a prism. Fortunately, this form of dispersion is not a problem in water across the range of sound frequencies in question. However, geometric dispersion, which is a function of the waveguide geometry, is a serious concern in underwater acoustics.

The method of normal modes is a general technique for solving the wave equation for the acoustic field at some distance from a sound source. The depth and wavelength of the sound determine the number of modes of propagation the waveguide can support. Each of these modes propagates at a phase velocity C_n, which depends on wavelength λ, group velocity c, depth h, and mode number n [5]:

$$C_n = \frac{c}{\sqrt{1 - \left(\frac{(n-1/2)\lambda}{2h}\right)^2}} \tag{2.19}$$

Phase velocity varies slowly with frequency for small n, and typically is constant across a signal's bandwidth at low mode numbers. However, at higher mode numbers there can be a significant difference in phase velocity across the bandwidth of a signal, which spreads the frequency components in time as they propagate.

The method of normal modes can be approximated by the WKB[7] or ray method. In this technique, the modes are assumed to correspond to discrete acoustic paths, or rays. These rays undergo reflection and refraction at boundaries in a manner similar to light modeled by geometrical optics. The primary assumption of ray theory is that these paths are non-dispersive. From 2.19, it can be shown that variations in phase velocity across a given bandwidth are minimized for high frequency and a slowly varying sound velocity profile. Therefore, to neglect dispersion the following relationship must hold [5]:

$$\frac{1}{\omega}\left|\frac{dc(z)}{dz}\right| \ll 1 \tag{2.20}$$

Fortunately, the shallow coastal waters in which REMUS typically operates are well mixed by currents and surface action, resulting in little variability of sound speed in the water column. Specifically, during the winter months when most of the data presented in this thesis was collected, the sound speed varied no more than a few tenths of a meter per second over the thickness of the waveguide. For REMUS navigation signals, which are in the range 10 to 30 kHz, the left hand side of 2.20 takes values on the order of 10^{-7}, easily fulfilling requirements for the ray approximation.

This same result can be arrived at by assuming the normal mode solution, and calculat-

[7]Originally, a simplification of the solution of the wave equation for quantum mechanics, developed by Wentzel, Kramers, and Brillouin.

ing the difference in phase velocity across the signal's bandwidth for each propagating mode according to 2.19. For typical REMUS mission conditions, it is not until approximately the 100th mode that this phase velocity difference appears across a 3 kHz bandwidth. While these modes are supported by the waveguide, they correspond to rays whose angles of propagation are so steep that their arrival time is delayed considerably from the direct path and their signal strength weakened by reflection loss at the interfaces.

Demonstrating that dispersion is not a factor in REMUS navigation and that the WKB approximation is applicable also allows the use of acoustic propagation models which depend on the ray approximation. Ray theory generates the most intuitively simple results for acoustic propagation, and output from ray models will be presented later in this chapter and used for comparison with field data in Chapter 4.

Noise

Although the effect of ambient noise on the maximum precision of arrival time estimation is included in the Cramer Rao bound discussed above, anticipating and combating this error requires knowledge of the sources of noise and their relative strengths. This section categorizes noise by cause, focusing on sources dominant in coastal waters.

The strength of underwater sound, including ambient noise, is generally expressed in units of decibels relative to one micro Pascal (dB re 1 μPa). The decibel is a dimensionless ratio of signal intensity, and in underwater acoustics the reference is a plane wave whose RMS intensity (pressure/unit area) is 1 μPa. Noise level is therefore given by

$$\mathrm{NL} = 10 \log \frac{I_N}{I_{ref}} \tag{2.21}$$

in which I_N is the noise intensity, and I_{ref} the reference intensity as defined above. This measure of signal level will appear frequently in the discussions of noise and acoustic signal strength which follow.

Surface Action: The movement of water at the free surface where it interacts with the atmosphere is the most important natural source of ambient noise in the ocean. In deep, open water, most of this sound is at low frequencies, but in the coastal zone the range to the surface is short enough that higher frequency sounds propagate throughout the water column. The processes by which sound is generated at the surface are not completely un-

derstood, but appear to be well correlated with sea state and wind speed. At extremely low frequencies, hydrophones can be affected by the change in hydrostatic pressure as waves pass over or even as the water column height above cycles with the tide. At higher frequencies, on the order of tens to hundreds of Hertz, sound may be generated by the turbulent pressure fields transmitted from the surface as the wind blows across the water ("flow noise"). However, at the frequencies of concern to REMUS navigation, the primary source of surface noise is probably wave action. As the wind increases and white caps form, the breaking of these waves and the subsequent popping of the bubbles entrained produces broadband noise at tens of kilohertz. Studies undertaken for the defense of harbors during World War II and since have shown that the average noise level in the coastal zone is approximately 50 dB re 1 μPa at 10 kHz and 40 dB re 1 μPa at 25 kHz, and increases by 7.2 dB for each doubling of the wind speed [16].

Man Made Noise: In the deep ocean, ship traffic is the single largest source of sound between 50 and 500 Hz, where noise from ships can propagate for thousands of kilometers. While these frequencies are too low to affect REMUS navigation, at short range in coastal areas and harbors ships, and boats are responsible for a broad spectrum of noise which reaches into the tens of kilohertz and above. Machinery noise is transferred very efficiently to the water by steel hull plating, and many other acoustic sources, such as depth sounders and sonars, operate close to or within acoustic navigation frequencies. Since this noise has numerous sources, and ships by their nature are mobile, measurement of a constant background noise level at these frequencies in coastal areas is difficult.

Other Acoustic Sources: There are many other sources of noise in the ocean, but fortunately most have a small impact in the coastal zone at typical acoustic navigation frequencies. For example, seismic and thermal noise are both relatively low level sources which are only a factor at low frequencies in the comparative quiet of the deep ocean. Biological noise, including the sounds produced for communication and navigation by marine mammals, can be intense in the coastal zone at frequencies in the tens of kilohertz, but has never been observed to be a problem in the field for a REMUS mission.

Non-Acoustic Noise: Although different in nature from the sources of noise discussed above, non-acoustic noise has the same deleterious effect on signal processing and arrival

timing. Non-acoustic noise has many potential sources, but the primary cause in this application is probably radio frequency pickup on unshielded components and leads in the preamplifier. Johnson thermal noise may also contribute, but given the temperature of components, the bandwidth of signal, and the output impedance, this is probably small compared with the signal level and other noise sources. Quantization error during analog to digital conversion may also play a small part. These sources and the steps taken to minimize them will be discussed further in the methods chapter.

Transmission Loss

Signal level is the counterpart of noise in the precision of arrival time estimation. Because source level is generally limited by power consumption constraints, transmission loss, combined with noise level, becomes the limiting factor in determining range far from the source. Acoustic transmission loss is expressed as a decibel quantity for comparison with noise level as defined in the previous section:

$$\mathrm{TL} = 10 \log \frac{I_0}{I_1} \tag{2.22}$$

where I_0 and I_1 are the acoustic intensities one meter from the source and at the receiver, respectively.

Underwater acoustic transmission loss has two independent components, spreading loss and attenuation. Attenuation results from the combination of scattering from objects such as entrained bubbles and plankton, absorption due to the molecular interactions in water which converts some acoustic energy to heat, and leakage of the signal out of the sound channel into the air or sediment. Spreading loss is caused by the geometry of a signal propagating in all directions from its source.

Spreading Loss: In an unbounded, lossless medium with an omnidirectional source, the energy of a signal is spread evenly over the interior surface of a spherical shell whose radius corresponds to the time since transmission. Since this energy is constant but the area of the sphere increases as r^2, the flux through a unit area of the shell must decrease with increasing distance, resulting in the following expression for intensity as a function of range:

$$I = \frac{P}{4\pi r^2} \tag{2.23}$$

where P is the signal power (energy/time) and r is the range, or radius of the sphere on which that power is distributed. With a reference distance of one meter in 2.22, the constant factors cancel and leave

$$\mathrm{TL}_{sphere} = 20 \log r \tag{2.24}$$

The assumption of spherical spreading is appropriate in water where the range to the receiver is less than the depth of the ocean, so an interface is never encountered. However, for applications such as coastal AUV navigation, the range is typically far greater than the water depth. This gives rise to a similar situation as that described above, but instead of spherical shells, the acoustic energy is spread evenly on approximately cylindrical surfaces of radius corresponding to range bounded by the ocean's surface above and the bottom below. Cylindrical spreading modifies 2.23 to

$$I = \frac{P}{2\pi r H} \tag{2.25}$$

where H is the water depth, and 2.24 becomes

$$\mathrm{TL}_{sphere} = 10 \log r \tag{2.26}$$

Attenuation: Absorption is the dominant component of transmission loss due to attenuation at the ranges and frequencies typical of acoustic navigation. The logarithmic absorption coefficient, α, is a measure of the absorption per distance propagated given by

$$\alpha = \frac{10 \log I_1 - 10 \log I_2}{r_2 - r_1} \tag{2.27}$$

where I_1 and I_2 are the intensities at ranges r_1 and r_2.

There are three primary causes of absorption in seawater: shear, or ordinary, viscosity, volume viscosity which effects the time taken for water molecules to return to their usual positions in the crystal lattice as the compression wave passes, and ionic relaxation of magnesium sulfate and boric acid. A semi-empirical expression derived by Thorp [13] includes these three factors:

$$\alpha = 1.094 \times \left(\frac{0.1 f^2}{1 + f^2} + \frac{40 f^2}{4100 + f^2} + 2.75 \times 10^{-4} f^2 + 0.003 \right) \tag{2.28}$$

in which f is the frequency in kilohertz, and multiplying by 1.094 converts from the traditional units of dB/kiloyard to dB/kilometer. From the behavior of this relationship, the advantage of lowering the operating frequency of the REMUS navigation system is obvious. At 27 kHz, absorption is approximately 5 dB/km, while at 10 kHz this drops to about 0.5 dB/km.

Combining expressions for transmission loss due to absorption and cylindrical spreading, we have

$$\mathrm{TL} = 10 \log r + \alpha r \times 10^{-3} \tag{2.29}$$

The importance of transmission loss and effect of range on arrival time estimation is seen by returning to equation 2.14, which gives the SNR of the matched filter output as a function of the SNR of the input, the bandwidth, and the signal duration. From Figure 2-7, it is clear that the matched filter output resolution suffers as SNR declines. If we assume that source level is constant and noise level is not a function of range, the SNR of the matched filter output, and therefore the resolution of the range estimate, is proportionate to $-10 \log r$.

Sound Speed Estimate

As discussed in the section on basic acoustic navigation techniques, there are two essential pieces of information required to calculate range: travel time, and sound speed. While considerable attention has been paid so far to the problems in estimating the travel time of the acoustic signal, the propagation speed of that signal has been assumed to be a constant, known factor. This is not the case.

Sound speed is a material property of the propagating medium which is primarily related to density. Since sound propagates by compression wave, it is to be expected that the proximity of molecules to one another, and the manner of their interaction, will determine the speed at which sound travels. In water there are several semi-empirical expressions for sound speed which are optimized for specific regimes of temperature, depth, and salinity. From the available literature, the following formula was selected as most appropriate for

this work [11]:

$$c = 1449.2 + 4.6T - 5.5 \times 10^{-2}T^2 + 2.9 \times 10^{-4}T^3 + \\ + (1.34 - 10^{-2}T)(S - 35) + 1.6 \times 10^{-2}D \tag{2.30}$$

where c is the sound velocity, T is water temperature in degrees Celsius, S is salinity in parts per thousand, and D is depth in meters. The problem of accurately determining sound speed for REMUS navigation is simplified by the shallow, well mixed water in the coastal zone. Since salinity and temperature are generally constant throughout the water column, and the dependence on depth in 2.30 is fairly weak, the shallow water waveguide is essentially isovelocity. However, an accurate estimate of this single sound speed still requires knowledge of the temperature and salinity of the water. For instance, in water with a temperature near 20° Celsius, a measurement error of a degree results in a range error of approximately 2 meters a kilometer from the source.

At this point, it is appropriate to emphasize the meanings of *accuracy* and *precision*. While the two terms are often considered interchangeable in colloquial language, there is a subtle difference in their scientific definitions. Accuracy refers to the quality of an estimate in absolute terms, an accurate estimate being one which is close to the true value. Precision, on the other hand, is relative, stating a measurement's quality in reference to other estimates of the same parameter. Therefore, it is possible to have a highly precise but inaccurate measurement, where all the points are close together but far from the actual value. Likewise, a point which falls close to the actual value but is part of an ensemble of measurements which are scattered is accurate, but not precise. Obviously, the goal of any experiment is to generate data which is both accurate and precise, but in a scenario such as acoustic navigation, it is difficult to achieve since an actual value of transponder range is hard to determine by means other than those under test. As the title indicates, this thesis is primarily concerned with the precision of underwater navigation, as the absolute accuracy is difficult to measure.

In well mixed coastal water, the parameters which affect sound speed vary slowly enough that sound speed measurement for acoustic navigation is primarily a problem related to

accuracy rather than precision. For the length of sequences typically used in acoustic navigation, the time scale of the variability is greater than the length of the signal. However, these effects can be important to consider in acoustic communications, in which the signals sent can extend in length to the point where the properties of the propagating medium can shift while over the course of the signal's duration.

There are some dynamic effects which can have an appreciable impact on range measurement precision, as they vary not over the duration of individual signals but between range estimates. Vertical stratification is common in many coastal areas where water masses of differing temperature and salinity are brought into contact by the forces of runoff, tide, and wind. While these water masses are generally well mixed in depth, where they meet water of differing characteristics a front is formed across which these properties are nearly discontinuous. This causes an abrupt change in sound speed, which reflects and refracts acoustic rays in the same way that horizontal stratification, such as the thermocline, does in deep water. This range dependence of sound velocity can actually be used as a tool to track these fronts, but greatly complicates the task of navigating a moving AUV. While REMUS often operates in areas of frontal discontinuity, range dependence does not appear to be a significant factor in Woods Hole where the field work reported in this thesis was conducted.

Current velocity also has an impact on sound speed. Unlike electromagnetic radiation, which depends on no medium for propagation, sound moves through the frame of reference of the material through which it is transmitted. Therefore, if sound is transmitted through a moving water mass, the sound velocity is the vector sum of the sound velocity in the water's stationary frame, and the current velocity of the water. In many of the coastal areas in which AUVs operate, currents of several knots are not uncommon. Depending on the orientation of the sound rays relative to the current, a water speed of 2 knots (1 m/s) can result in a range error of almost a meter at 1 kilometer. Again, this phenomena is generally of concern to accuracy, but in areas such as Woods Hole where currents are both strong and variable, precision can be affected as well.

Transponder Location

Another critical but problematic step is localizing the transponder. Even if the range estimates are perfect, the vehicle's position cannot be found without knowledge of where the

transponders lie relative to one another and the earth. This is a more difficult proposition than might initially be expected.

The single greatest error in acoustic navigation results from the limitations on transponder positioning accuracy. While GPS is the most accurate means available, even with differential correction the receivers commonly available on the boats deploying transponders are at best accurate only to within a few meters. Fortunately, this error is acceptable for most AUV applications, but is the primary reason that field tests of the absolute accuracy of range estimates are not practical.

Precision of the range estimate is a function of the manner in which the transponder is moored. Again, this can be the single largest source of error in acoustic navigation. While systems requiring exact repeatability often have transponders attached directly to the seafloor, AUVs such as REMUS generally must be deployable more quickly than a bottom mounting would allow. Most AUVs use transponders attached to the mooring lines of surface buoys, and therefore the transponder's position is affected by the buoy's motion in the waves, wind, and current. Since acoustic navigation is not usually used to determine the vehicle's depth, the buoy's wave heave, which changes the transponder's location in the z-axis, is not a severe problem. However, the movement of the buoy with the wind and current in the x and y axes can introduce significant precision errors. Most light moorings require considerable scope for the ground tackle to hold, so the length of the anchor rode is usually greater than the actual depth of the water. This leaves the buoy free to move on the surface within a "watch circle" centered over the anchor, but with a radius approximately equal to $\sqrt{\text{rode}^2 - \text{depth}^2}$. The buoy's position on this circle is a function of the surface currents and wind, and its radius changes as the water depth follows the ebb and flood of the tide. Distinguishing this error from signal processing and acoustic precision error, which are the primary focus of this research, is discussed with the results and analysis in Chapter 4.

Waveguide Effects

In shallow, isovelocity coastal waters such as those in which the experiments reported in this thesis were conducted, refraction of acoustic rays as they travel from source to receiver is not an issue as it is in deeper waters offshore with greater temperature gradient. However, sound is reflected by interfaces with air and sediment very efficiently, which creates a

waveguide between the source and receiver bounded by the water's surface and the bottom. As discussed in the introduction, these alternative paths by which the sound can travel give rise to multipath interference. Selecting the direct path arrival, which gives the straight line range to the source, from many arrivals which can possibly be stronger and closely grouped, creates the potential for error in both the accuracy and precision of the range estimate.

Ray Tracing and Waveguide Dimensions: The shape of the waveguide plays a significant role in the multipath arrival structure. While a propagating medium bounded by two flat interfaces will support many alternative paths, a boundary with relief, such as the ocean floor, can selectively eliminate rays, including the direct path, if they are blocked from the receiver by bottom features. In addition, the thickness of the waveguide also influences the multipaths that are allowed. As the water depth changes with the tides, the waveguide is constantly redefined.

Figure 2-12 shows three examples of ray diagrams and arrival times for acoustic waveguides of dimensions typical for a REMUS mission. These traces were generated using a simple ray propagation model, assuming a well mixed environment in which the sound speed profile is not range dependent. For the purposes of demonstrating multipath interference, it is assumed that the surface and bottom are good reflectors. In these diagrams, the source is located on the left and the receiver is on the right, and both are 5 meters beneath the surface. The water is 20 meters deep and the range is 1 kilometer, so the aspect ratio of the plot is not to scale. To keep the plots simple, sound is assumed to propagate from the source with a 10 degree beamwidth. This is actually fairly realistic since higher angle rays are usually severely attenuated by multiple reflections, as discussed below. Although the transducer would emit a continuous beam of sound, for modeling this pattern is broken into individual rays, each of which propagates independently according to Snell's Law. Those rays which arrive at the receiver are called eigenrays, and the angles at which they are emitted are the eigenangles. In the plots which follow, only the eigenrays and angles are displayed.

The top panel represents nominal conditions. The plot on the left shows the paths taken by the eigenrays. Notice that even though the sound speed gradient in depth is minimal, some refraction does occur. On the right, the time for each eigenray to reach the receiver is plotted, showing that the arrivals of the eigenrays are very closely spaced in time. This

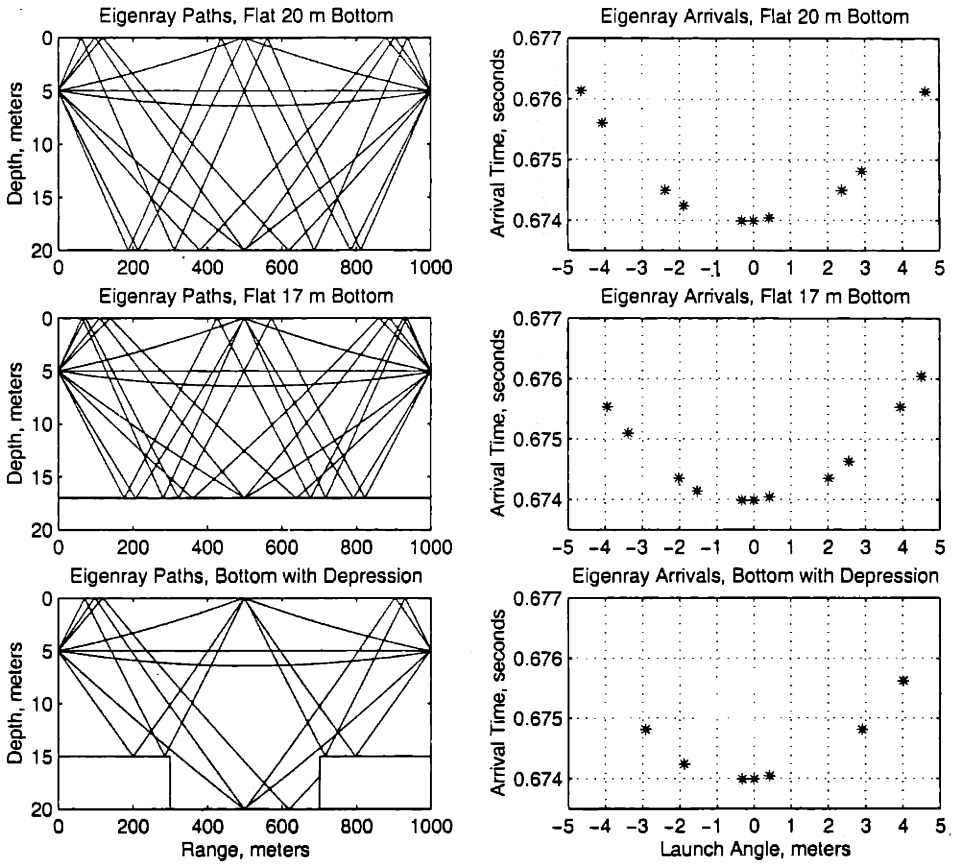

Figure 2-12: Simulated Acoustic Ray Paths and Arrival Times

is a result of the thin waveguide, which means that the direct path is only slightly shorter than those reflected by the top and bottom. This makes the arrivals difficult to distinguish. In addition, several reflected arrivals can arrive at the same time, as in the case of the rays launched at approximately +2.3° and -2.3°. While these rays will probably be attenuated by reflection and individually have less signal strength that the unreflected, direct arrival, the dynamic nature of the waveguide can create momentary conditions under which these weaker signals arrive at the same time and in phase, constructively interfering into a signal stronger than the direct path arrival. As a result, it is not sufficient to simply assume that since the direct arrival is unattenuated by reflection it will have the highest signal strength and will consequently correspond to the peak of the matched filter output.

The middle panel shows the output of the model for the same flat bathymetry profile, but with the water level 3 meters lower. There is some change in the multipath structure, but the direct path is unaffected and the arrival times of the reflected eigenrays remain approximately the same. Therefore, at long range in a thin waveguide, tidal fluctuation is probably not a major source of range estimate precision error.

The lower panel displays the results from a more realistic bathymetry profile. The range to the receiver is unchanged from the first two examples, but the depth is range dependent, as represented by the black line. In this case, there is a depression in the bottom between source and receiver. The eigenrays which do not reflect off the bottom are unaffected, but those which interact with this interface take different paths from those of the same eigenangle in the top panel. Some new modes of propagation are supported, and others (such as the ray plotted in red) which were allowed in the case of the flat bottom do not reach the receiver because they are trapped in the depression. Variable bathymetry also creates "shadow zones", regions which no ray can reach. In the example shown in the lower panel, the area just to the right of the drop off from 15 to 20 meters is blocked from the transmitter by the bottom of the waveguide. A vehicle here would be unable to receive a signal from the source location by any path. When deploying transponders for AUV operations, knowledge of the bathymetry is helpful in avoiding the inadvertent creation of these zones.

In all three cases, note the closely spaced arrivals grouped near 0° transmission angle. These arrivals are so closely spaced as to be indistinguishable in time. For simplicity and because of the test environment, range and depth dependence of the sound speed profile has

been essentially ignored in this discussion. However, it is important to note that within this group of arrivals, minute variability in sound propagation can cause slightly altered arrival times, potentially changing the structure of the constructive and destructive interference between the arriving signals. If this minute variability is time dependent, the signal strength of the direct arrival can appear to fade in and out, making consistent selection as the direct path arrival difficult.

The Lloyd Mirror Effect: Under the correct conditions, multipath interference can do more to cause error than simply confuse the problem of arrival selection. For a point source located near a reflecting boundary, a surface or bottom reflection can appear as a separate "image" source, creating a dipole radiation pattern. This effect is called the Lloyd Mirror, and in the far field the interference of the source and image rays causes the signal level to decay as $1/r^4$ as opposed to the usual $1/r^2$ in regions of standard spherical spreading, and $1/r^2$ in cylindrical spreading environments. Fortunately, signals coded in time are less likely to be effected by this phenomenon, since even if the arrivals of the two signals overlap at the receiver, the additional path length of the reflected ray causes its code to lag behind that of the direct ray.

Reflection Loss: The strength of the multipath signals is determined by the loss at each reflection, and the transmission loss associated with the additional path length traveled. Reflection loss at any interface is given by

$$\mathrm{RL} = 10 \log \frac{I_r}{I_i} \tag{2.31}$$

where I_r and I_i are the reflected and incident acoustic intensities. Since the difference in path length between reflected arrivals and the direct path is small for a thin waveguide, we will assume that the additional transmission loss is negligible and focus on reflection loss as the primary mechanism of multipath attenuation.

Reflection loss at the water's surface is highly variable. The impedance contrast between water and air is very high, so if the water is flat sound is reflected with essentially zero loss. However, if the water's surface is perturbed by waves, reflection loss increases as more energy is scattered by the angular surface. This loss increases with frequency, as the scale of the roughness relative to the wavelength of the sound is greater for shorter wavelengths.

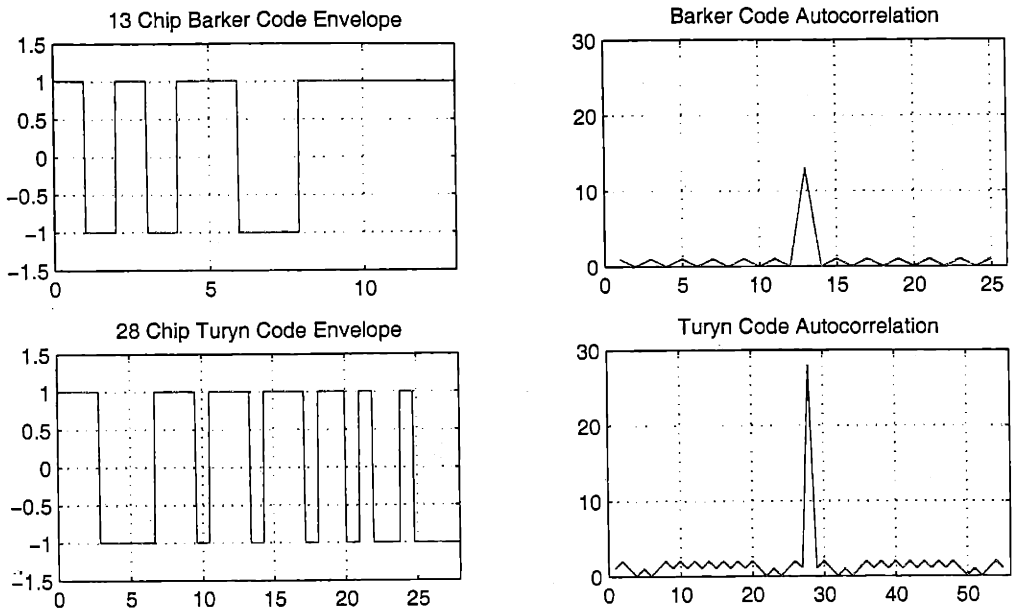

Figure 2-8: Barker and Turyn Codes Used on REMUS

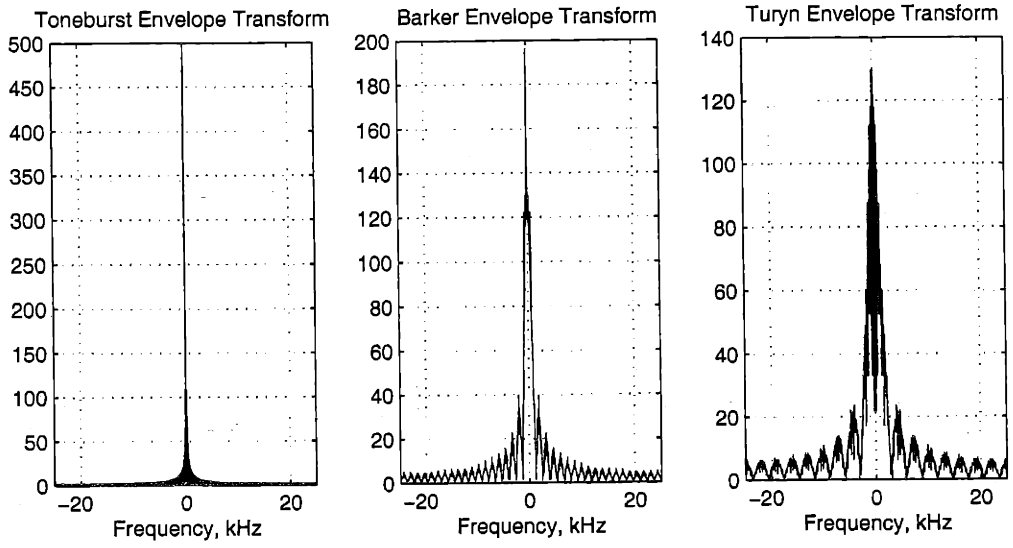

Figure 2-9: Fourier Transforms of the Toneburst, Barker Coded, and Turyn Coded Signals

The relationship between the scattering properties of a surface and the wavelength and grazing angle of the incident sound is given by the Rayleigh parameter,

$$R = kH \sin\theta \tag{2.32}$$

where k is the acoustic wavenumber of the incoming ray ($k = \frac{2\pi}{\lambda}$), H is the root-mean-squared (RMS) amplitude of the surface waves, and θ is the grazing angle (the angle formed by the sea surface and the incident ray). For $R \gg 1$, the surface scatters incident sound, while for $R \ll 1$ it acts as a good reflector. Calculating the Rayleigh parameter for the eigenrays generated by the model in Figure 2-12 for a range of RMS wave heights typical of coastal waters at the two REMUS frequency bands gives the following results:

Rayleigh Parameter Values

	1 Foot		2 Feet		3 Feet	
Incident Angle	10 kHz	27 kHz	10 kHz	27 kHz	10 kHz	27 kHz
0.3°	0.2	0.6	0.4	1.2	0.6	1.8
2°	1.5	3.9	2.9	7.9	4.4	11.8
3°	2.2	5.7	4.4	11.8	6.6	17.7
4°	2.9	7.9	5.8	15.8	8.8	23.6
5°	3.7	9.8	7.3	19.7	11.0	29.5

These values indicate that while the single surface reflected path (0.4° launch angle) will almost certainly propagate at 10 kHz, and probably at 27 kHz as well, higher launch angles at both frequencies will be considerably attenuated as wave height increases, especially those which undergo multiple surface reflections. Predicting reflection loss based on R is too inexact[8], but empirical data suggests that 1 foot waves will cause a 3 dB RL across this range of angles for the MF band, and presumably a somewhat lower loss at 10 kHz [17, 9]

Reflection loss at the bottom obeys the same basic principles as at the surface, but is by its static nature a much less dynamic problem which reduces its potential impact

[8]For electromagnetic wave scattering theory, it has been shown [3] that $\mu = e^{-R}$, where μ is the ratio of the coherent reflected amplitude to the incident amplitude. Applying this result to acoustics in which $(amplitude)^2 \propto (pressure)^2 \propto intensity$, RL can be written in terms of μ and therefore in terms of R. However, this results in values for RL which are much higher than observed, so it would appear that the assumptions on which μ is based do not hold for underwater sound.

on range estimate precision. However, while the dependence of reflection loss on incident angle, acoustic wavelength, and roughness is much the same at the bottom as the surface, the water/sediment interface is complicated by variable impedance contrast. Unlike the air/water interface where the impedance contrast is fixed, the bottom can be composed of many different types of materials, all with different sound speeds. This makes calculating the behavior of sound at this interface much more complicated. Porosity is a useful means by which to gauge a bottom type's reflection loss. The more porous a sediment, the more water it admits and the closer its density and sound speed come to that of water, reducing the impedance contrast at the boundary and increasing reflection loss. For this reason, materials such as mud and silt, which are highly porous, have relatively high reflection losses compared to materials such as hard packed sand and rock. In this thesis, variation in bottom type will not receive much consideration, as all the data was taken in environments with similar sediment characteristics.

The relationship between surface waves and reflection loss causes interesting conflicts in the forces working against accurate range estimation. In the case of multipath interference, it would appear that surface waves work in favor of increased precision, as they increase reflection loss and therefore decrease the number and strength of the multipath arrivals. This reduces the ambiguity in choosing the direct path, which is unaffected by these surface interactions. However, at the same time that surface waves increase RL, they also increase noise level, which we know decreases the precision of the arrival time estimate according to the Cramer-Rao bound. For the purposes of a "back of the envelope" calculation, consider a signal at 27 kHz propagating in a waveguide of dimensions similar to those shown in 2-12. If we assume a wind speed of 10 knots is sufficient to generate 1-2 foot waves, and further assume that each surface reflection at an angle of incidence of 5° will "cost" the ray 3 dB for this wave height, and each bottom reflection 4 dB (typical of a bottom of sand and rock [10]), an eigenray which follows a path involving 2 surface reflections and 2 bottom reflections will incur a total loss of 14 dB. At the same time, the increase in wind speed from flat calm (where reflection loss is essentially 0 dB) to 10 knots raises the ambient noise level by approximately 12 dB [20, 8, 12]. From 2.14, the SNR of the matched filter output is proportional to that of the input. Therefore, the matched filter peak corresponding to the direct path will be attenuated by a factor proportional to 12 dB, while the subsequent multipath arrival peaks are down by approximately 14 dB.

Another consideration is frequency: is the increased reflection loss at 27 kHz, and consequent decrease in multipath over 10 kHz, enough to make up for the signal strength lost as a result of the increased transmission loss at the higher frequency? Again, the result does not indicate a clear advantage either way. Increasing frequency from 10 to 27 kHz results in an additional transmission loss of approximately 4.5 dB over a kilometer, while the impact on reflection loss results in a gain of similar magnitude in multipath attenuation.

While the data on which these rough demonstrations are based are too inexact to carry any further, from the discussion of signal processing earlier in this chapter we know that the increased direct path SNR, and the decreased multipath eventually find their way through the matched filter and have opposite effects on the precision of the output and our ability to identify it. Do their effects cancel one another, leaving no net gain or loss for increased wave action or change in frequency, or does one or the other dominate? With no clear "winner" in these exercises, these are the types of questions this thesis will address by analysis of field data in the coming chapters.

Chapter 3

Experimental Method

The ability to transmit and receive coded signals has been a part of the REMUS system almost since the vehicle's inception. However, field testing of the vehicle using Barker and Turyn codes for navigation has not always yielded the improved performance expected. The series of experiments reported in this thesis was undertaken to provide the first quantitative data on the precision of this navigation system, and as much information as possible on the sources of the error observed. This chapter describes the methodology and apparatus employed in this investigation. A chronological approach is employed, tracing the evolution of the experiment from rough field data acquired at LEO-15 during the summer of 1998 to more carefully controlled, longer term experiments in Great Harbor in Woods Hole during the winter and spring of 1999. As the majority of the results presented in the next chapter are based on these last experiments, the equipment and procedures employed in Great Harbor are described in the most detail. The earlier field work is discussed in general terms and presented with the results which prompted the next step in the research.

3.1 July 1998: REMUS at LEO-15

During July of 1998, REMUS was deployed at LEO-15, the Long Term Environmental Observatory at 15 Meters located off the coast of New Jersey near the Rutgers University Marine Field Station in Tuckerton. As part of the 1998 National Ocean Partnership Program(NOPP), REMUS operated daily, conducting ADCP and CTD surveys along a 20 kilometer transponder baseline in water ranging from approximately 15 to 25 meters deep with a flat, sandy bottom. Because of the size of the area of operations, the low

frequency navigation system was used to extend the acoustic range and reduce the number of transponders needed.

During these operations, it appeared that the use of signals coded with Barker and Turyn codes provided no tangible benefit to the precision or accuracy of navigation. A more controlled test of the navigation system was conducted on July 29, 1998 to investigate this unexpected observation. A small boat was moored approximately 1500 meters from one of the transponders in the array, and a REMUS vehicle was suspended over the side, about 2 meters beneath the surface. Sea conditions were fairly quiet, with a small swell rolling through the area. The vehicle was programmed to find the range to the transponder.

As discussed in Chapter 2, REMUS navigates by interrogating the transponder with an uncoded toneburst, but the transponder can reply with toneburst, Turyn, or Barker coded waveforms. In this case, three signals were used: a 10 ms toneburst at 10.5 kHz, a 10 ms Turyn code at 11 kHz, and a 10 ms Barker code at 11 kHz. REMUS was programmed to save not only the calculated range, but the raw signal sampled at 50 kHz by the DSP and the matched filter output as well. Due to limited time and storage space, only four examples of each waveform were acquired.

Given the small sample size and conditions under which the data were collected, some error in range estimate precision was to be expected. Despite being anchored, both the transponder's mooring and the boat from which the vehicle was suspended were moving. Also, regardless of which code the transponder replied with, the interrogation ping was always a toneburst (since the transponders do not have the onboard DSP required for matched filtering to estimate arrival times of coded waveforms), effectively cutting any advantage gained from the coded signals in half. However, the 10 meter range of the results calculated by the vehicle, as shown in Figure 1.1 in the introduction, was a spread larger than could be easily accounted for by the source and receiver movement. Furthermore, the coded signals appeared to give no improvement whatsoever.

When the raw data files from REMUS were reprocessed on land, some errors in signal processing were discovered which were subsequently corrected on the vehicle. These corrections did little to improve the range estimate precision of the coded signals, however. Figure 3-1 shows three typical matched filter results from this data set.

Without a clearly dominant early peak, identifying and consistently selecting the direct path arrival for the range calculation is impossible. The poor quality of this output raised

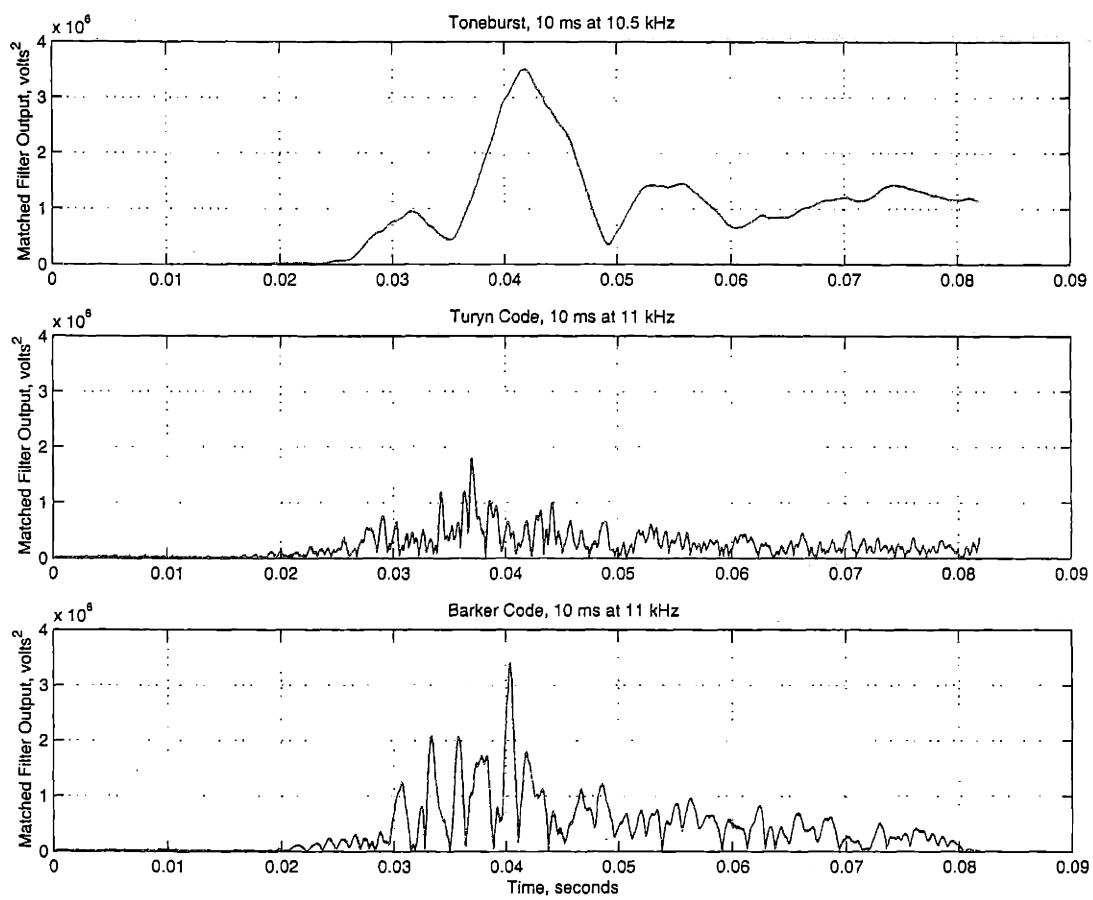

Figure 3-1: Matched Filter Output from LEO-15

several possibilities which demanded further investigation. While the long, thin waveguide of LEO-15 would be expected to allow stronger multipath propagation than other environments, it was possible that there were still errors in the signal processing hardware or software. One particular concern was that the bandwidth of the transducers was narrower than the coded waveforms, in particular the Turyn code. If the spectral content of the outgoing signal was altered, a matched filter at the receiver using the original template would not correctly identify the arrival time. After returning to Woods Hole, a more controlled set of experiments was undertaken to clarify this situation.

3.2 Fall 1998: Transponders off the WHOI Dock

The primary goal of the next experiment was to simplify the conditions so that the potential sources of error observed at LEO-15 could be systematically isolated and addressed. First, the two way acoustic path was eliminated by commanding the transponder to "ping" by either direct RS-232 cable connection or via the radio modem, rather than with an acoustic interrogation signal. Also, instead of using REMUS for the signal processing, the signals were recovered independently using the same hydrophone as is installed in the REMUS nose cone, preamplified, and displayed on a digital oscilloscope triggered by the outgoing command to the radio modem. The digitized data was uploaded to a computer, where the signals were processed in Matlab. This alleviated the storage space limitations encountered while storing data on REMUS at LEO-15, greatly increasing the number of signals acquired.

Over the course of several days in September, October, and November of 1998, data sets were recorded at a variety of ranges, with both low and medium frequency transponders. To verify that the bandwidth of the transducers was sufficient to accommodate the coded signals and that there were no additional errors in signal processing, several data sets were taken with the transmitting and receiving hydrophones at very short range (a few centimeters, but not in actual contact) in air. Because sound propagates so much less efficiently in air, it was expected that multipath arrivals could be eliminated to allow the experiment to focus solely on the signals and their processing. Similar data was recorded at short range in the well at the end of the WHOI dock, where the transducers were lowered to approximately the mid water depth, at least 5 meters from both the surface and bottom, again to minimize the interference of reflected paths. Figures 3-2 and 3-3 show examples

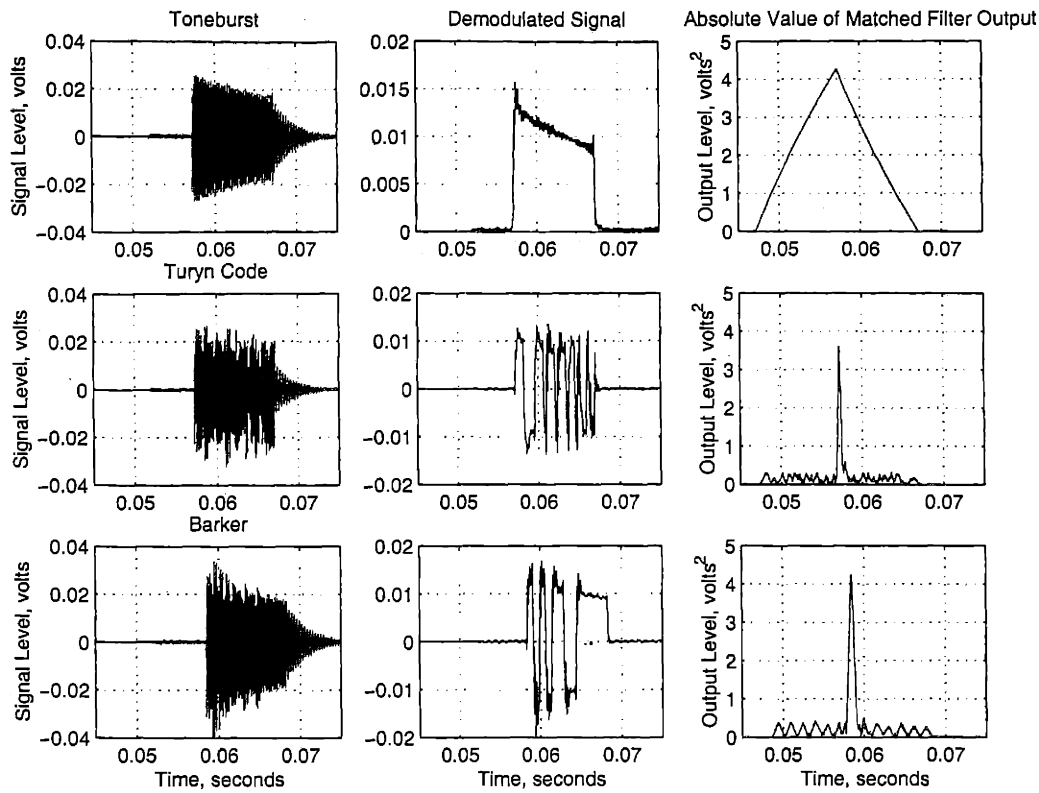

Figure 3-2: 10 ms 11 kHz Signals, Range 0 m (in air), 11-10-98

of the received signal, the demodulated signal, and the matched filter output from the LF and MF data sets. Comparing these plots with those of the simulated data in the last chapter, it appears that both transducers have sufficient bandwidth to reproduce the spread spectrum signals accurately. This was of particular concern for the MF transducer, which has a slightly narrower bandwidth than the LF transducer. While the transitions of the square codes show more ripple than did the simulation, the matched filter output is still quite good.

One other potential source of error was the radio modem link to the transponder. Because the shore side terminal commanding the transmitting radio modem, the transmitting modem itself, the receiving modem, and the microprocessor on board the transponder all had separate, asynchronous clocks, it was possible that the time between the ping command at the terminal and the actual acoustic transmission could be subject to variation which could affect the precision of the range estimate. From the data taken in close proximity in air, however, this lag was found to have an average value of 0.057 seconds, with a standard

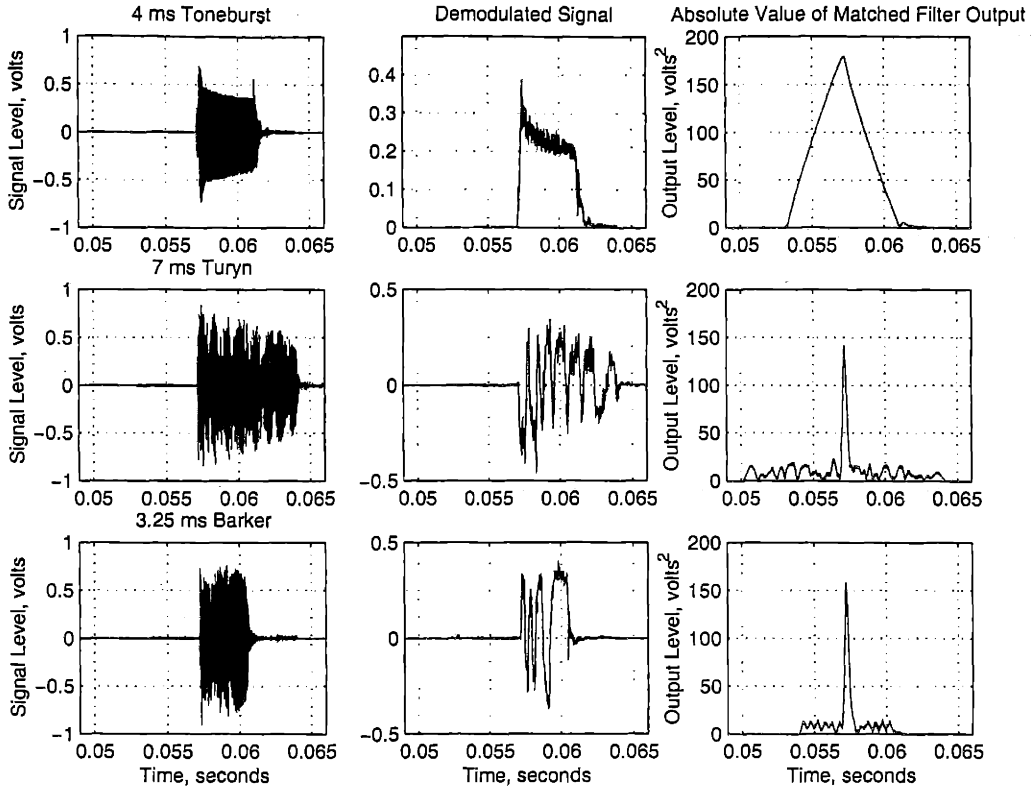

Figure 3-3: 27 kHz Signals, Range 0 m (in air), 11-24-98

deviation of 4.2×10^{-5} seconds, corresponding to a range error of ± 0.063 meters. As this is well within the precision required for navigation and below the level of error observed at LEO-15, it was concluded that "jitter" in the command timing was a factor that could safely be ignored.

Having ruled out two potential sources of error in the hardware, the transponders were moved into the harbor to begin the investigation of the environmental effects. Ten samples of each code at each frequency were recorded at ranges of approximately 0, 100, 200, 400, and 600 meters, as well as 1 kilometer for the MF transponder. Again, as the focus of this research was on precision rather than accuracy, no attempt was made to quantify the range in absolute terms, but the ranges calculated acoustically did agree well with the approximate DGPS position of the boat as the transponders were deployed. Figures 3-4 and 3-5 show the standard deviation of each group of ten range estimates.

The data here suggests that the coded signals offer significant improvement over the toneburst, and that the medium frequency signals are slightly more precise than the low

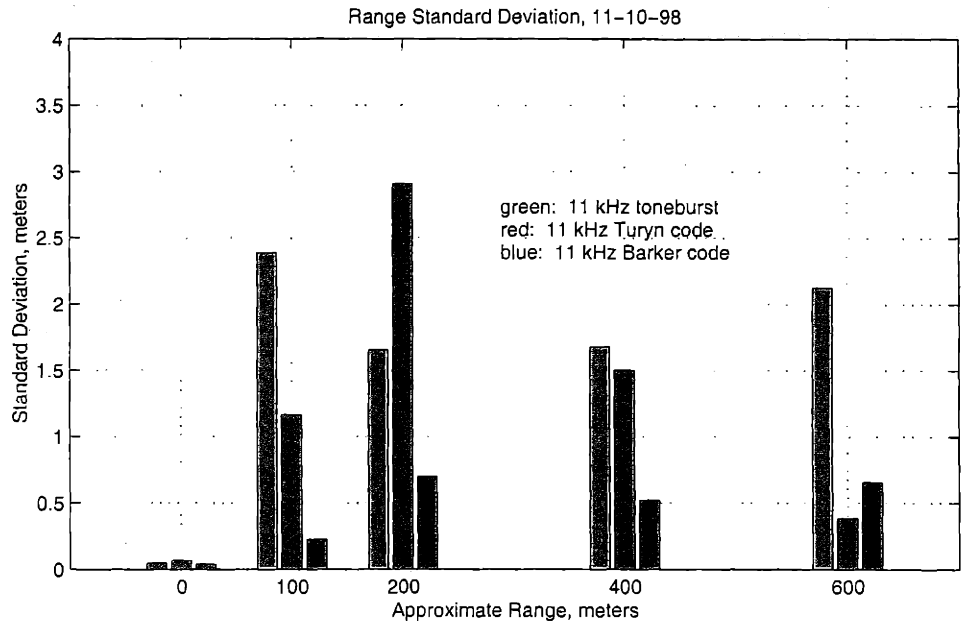

Figure 3-4: LF Range Standard Deviations at 5 Stations

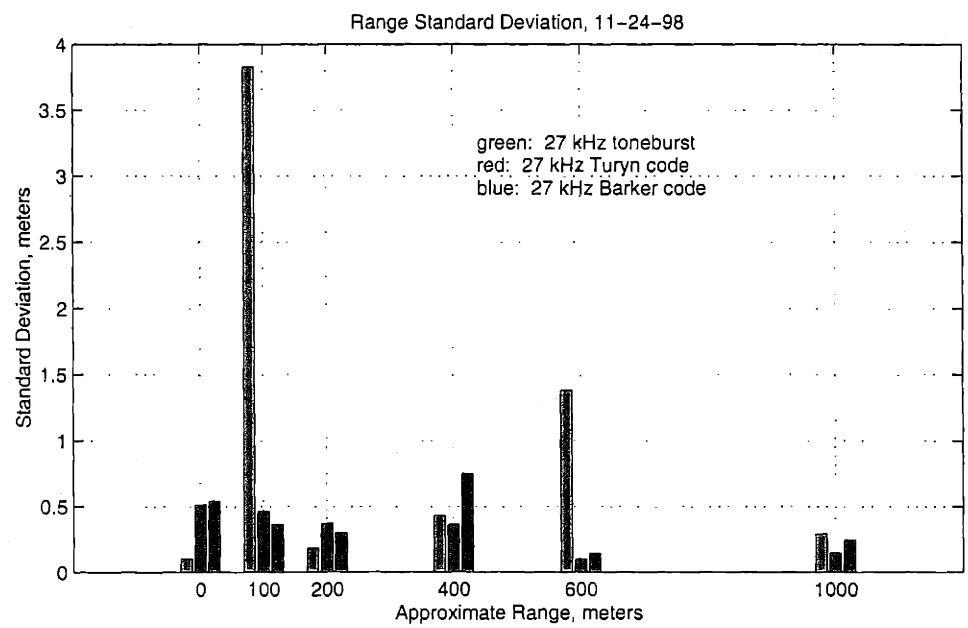

Figure 3-5: MF Range Standard Deviations at 6 Stations

frequency. However, 10 samples is not a large enough ensemble to be truly statistically meaningful. While the experiments in the fall of 1998 were sufficient to rule out several possible sources of error and narrow the field of study to the environmental factors affecting acoustic propagation, there was insufficient data to make any quantitative determination of the precision of the Barker and Turyn codes with respect to toneburst waveforms or each other, or the effect of frequency on precision. In order to further understand the properties of coded signals, and investigate their performance under a wide range of environmental conditions, a more long term installation capable of acquiring much larger data sets in combination with environmental information was required. The equipment and operation of this system is described in the next section.

3.3 Spring 1999: Transponders in Great Harbor

With the apparatus and signal processing proven sound, the investigation shifted to focus on the effect of the waveguide and dynamic environmental variables on the precision of range estimation. However, in order to generate the volume of data over time necessary to make a quantitative assessment of these factors, a more automated system was needed. The previous data were acquired manually, requiring human intervention to send the command to the transponder, upload the data to the computer, and save the trace in the correct format. For the next experiments, these operations were automated with a computer program which managed the entire data run, allowing unattended operation and the collection of thousands of data points. In addition, environmental data was collected from a variety of sources for correlation with the acoustic range estimates. This section describes in detail the automation system, the components controlled, and the data acquired, and is divided into six parts. First, the general physical layout of the experiment is described, followed by a discussion of the equipment used for transmission and receipt of the acoustic signals. Next are described the data acquisition system, the sources of environmental data, and the signal processing which was conducted after the data were collected. The final section describes some of the specifics of the data acquisition procedure.

3.3.1 General Layout

Long term study of the effect of environmental factors on acoustic propagation and range estimate precision required that the experiment be sufficiently controlled so that the influence of other variables could be minimized or accounted for. To this end, a fixed acoustic path was established across Great Harbor in Woods Hole, eliminating the effects of variable range or changes in bathymetry from the analysis. A receiving station was established on the dock at WHOI, and a low frequency and mid frequency transponder, both with onboard radio modems, were moored across the harbor at a range of approximately 600 meters. Except to change batteries and perform other routine maintenance, the transponders were left undisturbed for the four months they were deployed. While it would have been preferable to use a truly stationary source (the moored transponders being subject to watch circle error, as discussed in Chapter 2), the use of these "off the shelf" REMUS system components made deployment much simpler and allowed time which would have been spent on system development to be used for the collection and analysis of data.

Figure 3-6 is a section of the chart of Great Harbor, with the approximate position of the two transponders off of Ram Island, the acoustic path, and receiver location superimposed. While greater range might have been desirable, this was the longest path available with a fixed receiver at WHOI and unobstructed by land or bathymetry. This path does, however, run perpendicular to the prevailing current along the far side of Great Harbor, minimizing the effect of variable water velocity on sound speed.

Figure 3-7 shows the bathymetric relief relative to mean low water (MLW) as sounded at 50 meter intervals along the track from the transponders to the receiver. Note that the depth and range in this plot are not on the same scale, so once again the waveguide is essentially long and thin, as shown in the modeling examples in Chapter 2. The bottom along this path is primarily sand, with some rocks and rubble. These conditions are generally favorable to multipath propagation, though some reflected sound may be lost in the bathymetric depressions or scattered from the less smooth areas of the bottom.

3.3.2 Acoustic Subsystem

The REMUS transponders were the single most important component in the experimental system. As described above, a low frequency and a mid frequency transponder were deployed

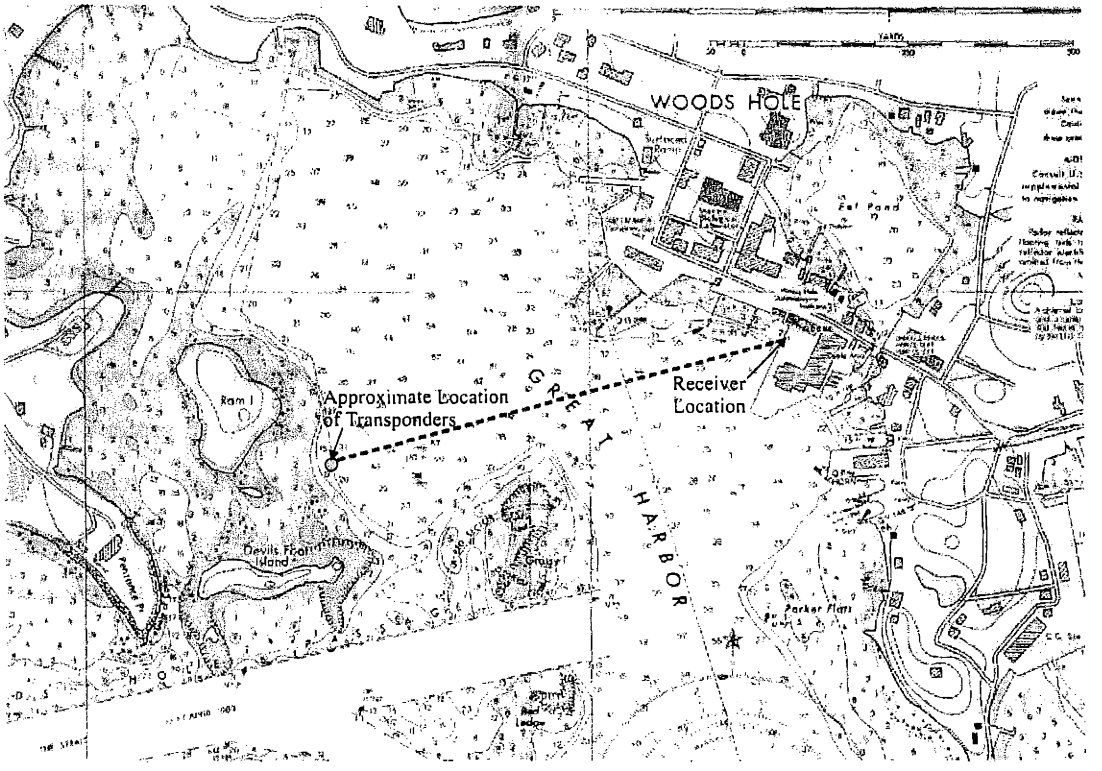

Figure 3-6: Great Harbor, Woods Hole, Massachusetts (from NOAA chart #13235)

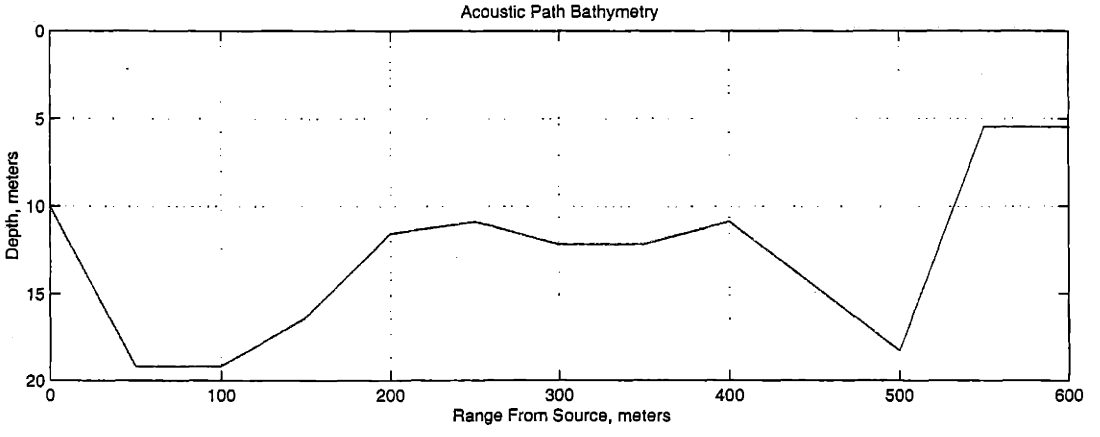

Figure 3-7: Depth Along Acoustic Path

in Great Harbor for four months in the spring of 1999. Although the MF transponder was packaged in a smaller, lighter mooring designed for easy launch and recovery, functionally they were nearly identical, with the exception of the difference in frequency.

The transponders were set up with the electronics in a watertight compartment on the surface expression, and the actual transducer attached to the anchor rode approximately 5 meters below the buoy. The Danforth anchor securing the LF transponder required 17 meters of rode and chain. Since there was approximately 2 meters of chain at the anchor which was probably too heavy to be moved by motion of the surface expression, an effective anchor rode length of 15 meters produced a surface watch circle with a radius of approximately 11 meters. At the transponder, which was 10 meters from the bottom section of chain, this watch circle shrank to approximately 7 meters (ignoring the catenary shape of the anchor rode). The MF transponder, being smaller and lighter, was moored with a small mushroom anchor requiring considerably less scope than the Danforth. This reduced its maximum watch circle to a few meters at the depth of the transducer.

The key to the experiment was the radio modem installed on each transponder. In a REMUS field operation, this permits the transponders to calculate range to REMUS, and report back to shore where operators can track the vehicle's progress. For this experiment, the modem was used simply to command the transponder to transmit one of the eight waveforms it was programmed with, allowing the one way acoustic travel time to be measured. The radio modem acted as a completely transparent serial link to the transponders, providing direct communication with the onboard microprocessor.

In order to keep the number of variables to a minimum, the signals were selected to be as similar as possible, varying only in the code with which they were modulated. From the available REMUS codes, the following were selected:

Transponder	Frequency	Duration	Waveform
LF	11 kHz	10 ms	Toneburst
LF	11 kHz	10 ms	Turyn
LF	11 kHz	10 ms	Barker
MF	27 kHz	4 ms	Toneburst
MF	27 kHz	7 ms	Turyn
MF	27 kHz	3.25 ms	Barker

The LF acoustic signals transmitted had a nominal source level of 193 dB re 1μPa @ 1m, while the MF signals were slightly weaker at 186 dB re 1μPa @ 1m. The transducers produced a toroidal beam pattern, with a beam width of about 90° for LF and 45° for MF. With transmission loss calculated according to Equation 2.29, this gives an acoustic signal level at the receiver of approximately 166 dB re 1μPa @ 1m for the LF transponder and 156 dB re 1μPa @ 1m at MF. At the receiver, the same transducers were suspended approximately 3 meters below the surface. They added a small amount of gain due to their directivity. The sensitivity at the receiver was -188 dB re 1 V/μPa, and 32 dB of gain was taken by the preamplifier before the signals were sent to the oscilloscope for digitization.

3.3.3 Data Acquisition

The acoustic signals were acquired from the preamp by a Hewlett-Packard 8 bit digitizing oscilloscope, which displayed the time series and uploaded the data to the computer where it was saved to the hard drive. Although the scope had a maximum sample rate of 100 MHz, the data was digitized at a lower rate corresponding to a window of useful length for signal processing. The low frequency signals, which had a length of 10 ms, were digitized at 40 kHz giving a window length of 50 ms, while the shorter mid frequency signals were sampled at 100 kHz for a window length of 20 ms. Both these sample rates satisfy the requirements of the Nyquist theorem.

The data acquisition process was directed by a program written in QBASIC running on the storage computer, an 80486 processor based DOS PC. Via an RS-232 serial link to the oscilloscope and the radio modem connection to the buoys in the harbor, this program selected the correct transponder and waveform, commanded the transponder to ping, set the scope to the correct mode and triggered it to acquire data, and uploaded and saved the time series. Figure 3-8 is a slightly simplified diagram of the program's operation.

For the purposes of this program, each complete data set is divided into "data series", groups composed of a toneburst, Turyn code, and Barker code time series acquired sequentially. The three signals composing each data series are taken immediately following one another, so it is assumed that they propagated under nearly identical environmental conditions, and therefore vary only in waveform. Originally, it was hoped that not only signals of different waveforms could be included in the same data set, but signals of different frequencies as well. Unfortunately, it quickly became apparent that the radio traffic between

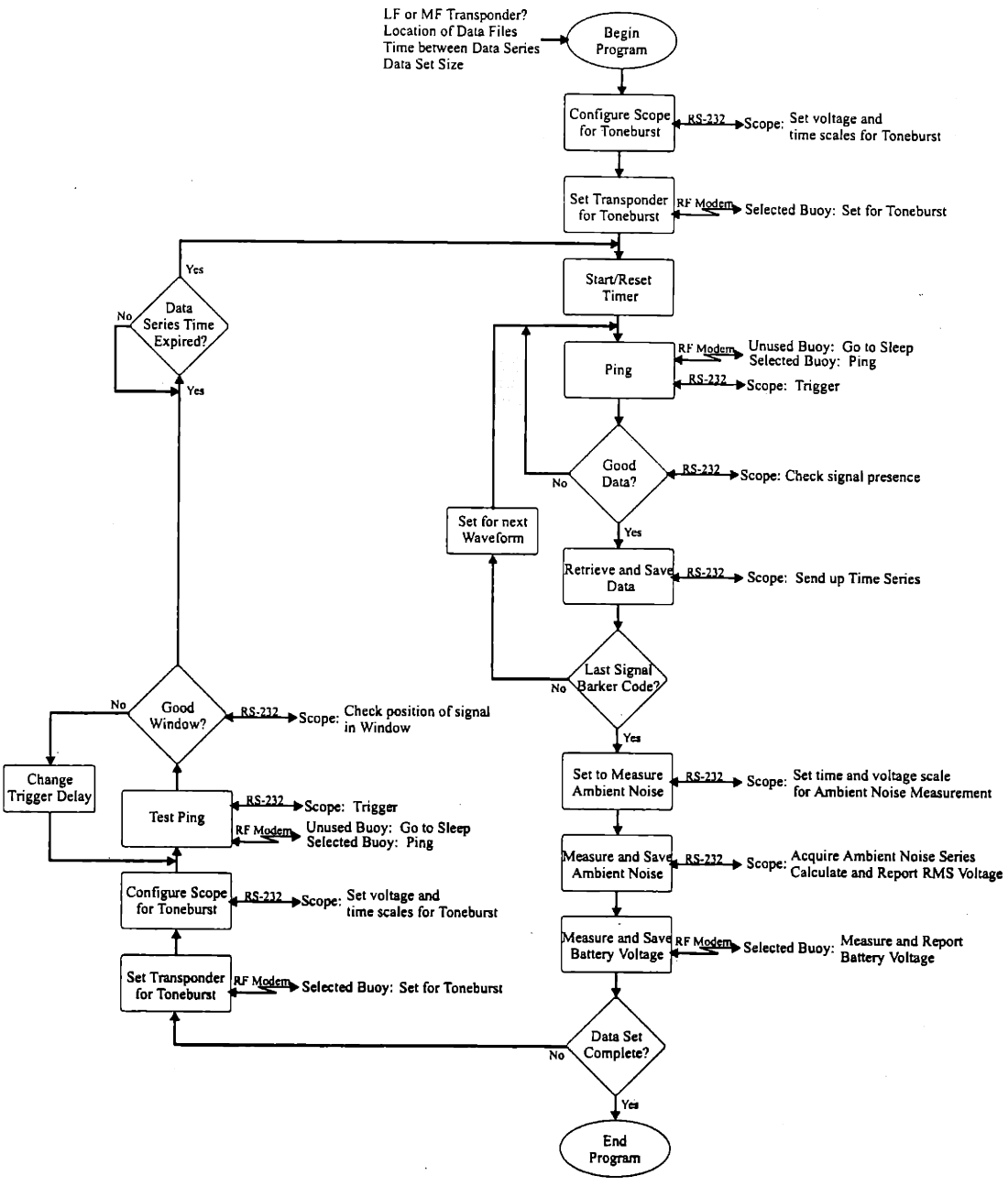

Figure 3-8: Flowchart of Experiment Control Program

three operational radio modems, even though addressed directly from the receiving station to only one of the transponders, caused unacceptable timing error. As a result, each data set contains examples of all three waveforms, but at only one frequency.

To execute, the program requires the user to specify which frequency will be used for the data set, the length of time between data series (typically 5 minutes), the location on the storage computer's hard disk where the data is to be saved, and the number of data series to be acquired. Once this information is given, the program begins by commanding the oscilloscope to recall its preset voltage and time scales and trigger delay for a toneburst signal at the specified frequency. These settings are saved in the oscilloscope's memory by the user before running the program. It is assumed that the user has correctly configured the scope for the current propagation conditions, so no testing of these settings is done before the first data is acquired. Next, the transponder corresponding to the selected frequency band is set to transmit a toneburst waveform at the desired frequency. At each step requiring the computer to interact with the transponder or oscilloscope, the program waits for acknowledgment of its command before proceeding. Once the configuration of the oscilloscope and transponder are verified, the program starts the data series timer.

At this point, the program is prepared to acquire data. To send a ping, it was found to first be necessary to put the unused transponder "to sleep." The transponders shut down to save power if they receive no traffic for 15 minutes. After going to sleep, they wake up for 15 seconds every 15 minutes to listen for traffic. If they hear none, they go back to sleep, but if there is traffic, they come back up to full power. Since it was observed that more than two radios in operation at the same time led to timing errors, the unused buoy is explicitly put to sleep before every ping command, to avoid the possibility that the command should fall in the unused buoy's 15 second listening window.

After the ping command is sent, the oscilloscope is triggered by the reply from the buoy confirming the acoustic transmission. Before uploading the data, the scope scans the received trace to be sure that a signal did in fact arrive within the window. Despite their general reliability, the radio modems are subject to occasional mistiming, which necessitates this error checking. If the signal has not arrived in the window, a new ping command is sent. If the data is good, it is uploaded to the computer and saved with a header identifying the time, date, and signal type. These time series are saved in a format compatible with Matlab.

The program then checks to see if the data series is complete. Time series are acquired in the order Toneburst, Turyn code, Barker code. If the last signal sent was a Barker code, the series is complete and the program continues. If not, the scope and transponder are set for the next waveform in the series, and the process of data acquisition repeats.

Once the data series is complete, two other pieces of data are collected. First, the scope is set to acquire a 50 second ambient noise sample. Rather than uploading the entire time series, the oscilloscope calculates the RMS noise level internally, and reports this value to the computer where it is saved, again with an identifying stamp. Finally, the transponder in use is commanded to report its battery voltage, which is also recorded.

Assuming that the number of data series to be acquired has not yet been reached, the program then enters the return loop, and a test ping is sent to check the position of the signal within the scope window. As the environment changes over the course of the data set, the signal can wander too close to the front or back of the window, cutting off desired segments of the time series. By examining the voltage level in several time ranges at the beginning, middle, and end of the window, the program determines if the signal is centered. If it is, the program proceeds, but if not, it adjusts the trigger delay to compensate and another test ping is sent to confirm the correction.

After the window check, the program waits for the data series separation time to expire. This time is usually set to much longer than the time required to receive the three pings of the series so that there is enough time for the window to be checked and adjusted before the next series starts. This keeps the data series evenly spaced, making later spectral analysis of the ranges in time much simpler. Once the time has expired, the timer is reset and a new data series is acquired. This outer loop continues until the data set is complete, at which point the program terminates.

3.3.4 Environmental Data

In order to investigate the performance of coded signaling under a range of environmental conditions, accurate data on the physical state of the acoustic waveguide was required for the same periods as the time series data. Without direct instrumentation in place, much of this data was inferred from other measurements. This section describes which variables were measured and how the data was obtained.

Sound Speed

As mentioned in Chapter 2, Great Harbor is a body of water well mixed by the forces of tide and wind. As a result, several surveys using a hand held YSI meter which measured salinity and temperature showed no dependence of these variables on range, and little dependence on depth. However, water temperature, especially at the surface, did rise perceptibly with solar exposure. Given these conditions, sound speed for each data set was determined based on several casts of the CTD off the WHOI dock over the course of the acquisition of acoustic data. This data allowed the calculation of a time dependent sound speed if variability of surface temperature with time of day was observed.

Tide Height and Current

While tidal height probably had little effect on the propagation properties of the waveguide (see the section on waveguide effects in Chapter 2), it had a potentially large impact on the size of the watch circles swept by the transponders. The tidal height in Great Harbor is measured by a NOAA tide gauge, the data from which is available on the World Wide Web. This information was retrieved and archived for analysis with the range estimates.

A related, but probably more important effect was that of the current acting on the transponder buoys. The currents through Woods Hole are very strong, and produce eddies in the area of the harbor where the transponders were moored. Unfortunately, there is no permanently deployed current meter to track the magnitude and direction of these flows, and none were available for temporary deployment. However, the current in the harbor is linked to the current through the Hole, which has been well documented. The Eldridge tide tables [22] contain predictions of current direction and magnitude in Great Harbor as a function of tidal level. Using the NOAA tide gauge data, it was possible to estimate the current velocity at the transponder buoys.

Wind Speed

Again, this is a somewhat indirect measurement of the desired quantity. As discussed in Chapter 2, surface waves are a source of noise in the ocean, but also degrade the strength of surface reflected paths. In general, wave height is a function of the wind speed, as in protected coastal areas surface action is almost always generated locally (as opposed to open

ocean swell). There was no wave gauge available for use during the acoustic experiments, but a qualitative measure of the surface action was taken from visual observations made at the time data was taken. In addition, the wind velocity, is recorded for several stations in the surrounding area, including the WHOI Meteorological Tower and the Buzzards Bay Tower. This is posted to sites on the Internet, from which the data at the time of the experiments was retrieved. However, the WHOI tower is protected by land from some wind directions, and the Buzzards Bay tower is approximately 20 miles away surrounded by open water, so neither offer perfect data on the conditions in Great Harbor. Since the WHOI tower was out of service for most of the period during which these experiments were conducted, Despite its shortcomings in combination with the visual observations of wave height, this data did prove useful.

3.3.5 Signal Processing

Once saved on the storage computer's hard disk, the data was transferred to more powerful machines for processing. The signals were demodulated and matched filtered in Matlab using a script which implemented the technique described in Section 2.3.1. While frequency domain processing often runs faster, for ease of examination of intermediate results all work was done in the time domain.

As discussed previously, one of the most difficult aspects of the range estimation problem is consistently selecting the direct path arrival from matched filter output which may contain many strong peaks, the strongest of which may or may not correspond to the direct path. While it is possible for a human to apply experience and often correctly identify the direct path by eye, automating this selection process is inexact at best. For the data collected here, a version of the same algorithm as is used on the REMUS vehicle was implemented. This technique depends on the assumption that the direct path will generate a fairly high, if not the highest, peak near the beginning of the matched filter output. Therefore, the direct path is identified as the first arrival within some fraction (usually 0.6-0.7) of the maximum peak. The matched filter output is scanned from the beginning to that maximum peak, and the first peak to satisfy the criteria is selected as the direct path arrival. Figure 3-9 shows the steps of the process in detail.

The top panel of 3-9 shows the raw absolute value of the matched filter output of a typical acoustic signal, in this case one modulated with a Barker code waveform. For this

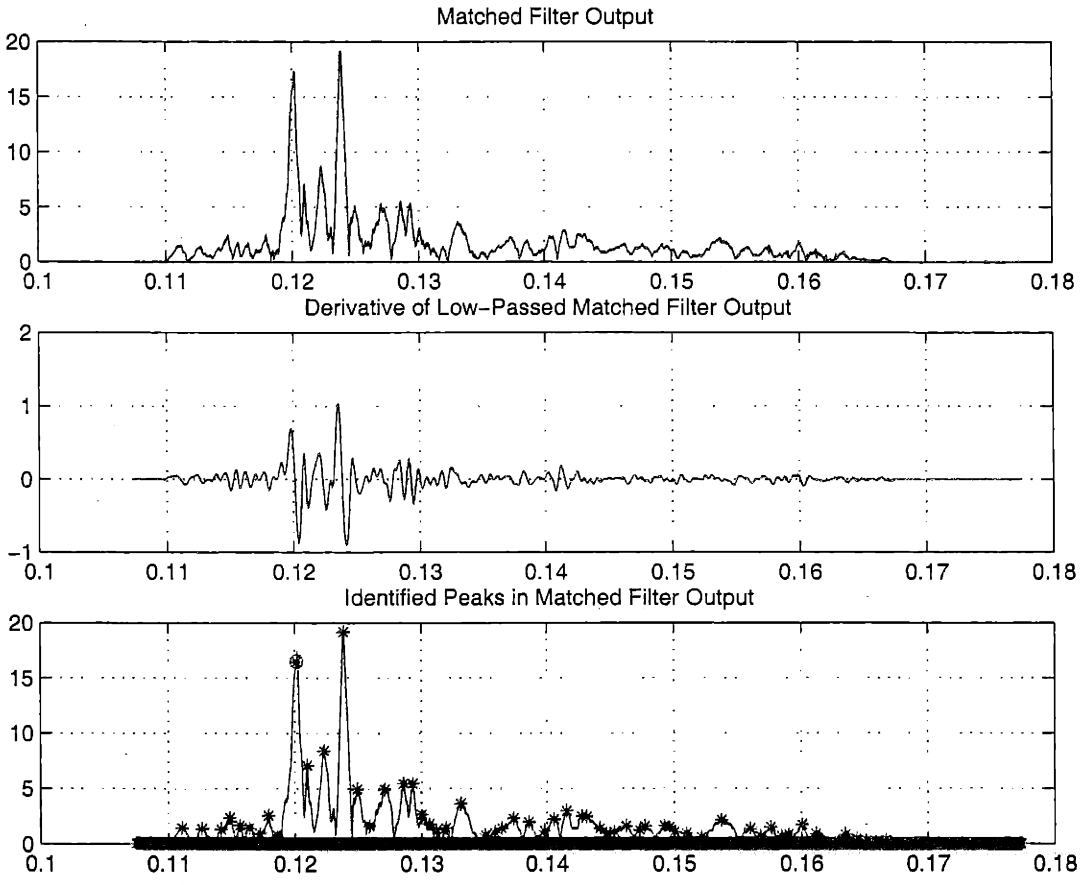

Figure 3-9: Direct Path Arrival Detection

data, it appears almost certain that the first large peak is the direct path arrival, since there is nothing approaching its magnitude preceding it. However, there is another even stronger peak following it. To implement the process described above, the matched filter output must first be low pass filtered to remove the ragged local maxima of the output which would otherwise be misidentified as independent peaks. This is done first forwards, then backwards, to avoid the phase delay caused by one way low pass filtering. The low passed data is then differentiated, as shown in the second panel. The peaks in the original data are identified as the points where its slope changes sign from positive to negative. These zero crossings are found, and associated with their corresponding ranges. In the third panel, the matched filter output is again shown with the true maxima marked by asterisks. The peaks are then arranged chronologically, and scanned. The first peak whose magnitude is above the detection threshold is identified as the direct path arrival, and its corresponding range is taken as the range to the transponder. In this case, as shown by the circled asterisk, the algorithm correctly selects the peak identified by eye as the direct path arrival.

3.3.6 Procedure

With the system described above in place, the experiment could be left unattended to acquire data in theoretically limitless amounts. In practice, once the program was bug free, data was typically acquired for about 12 hours, to include the full tidal cycle. Attempting to take longer data sets usually resulted in a program crash, as the QBASIC interpreter was not particularly effective at reallocating freed memory, and the physical memory available in the storage and control computer was limited. Data sets were taken under a range of environmental conditions, the primary variable being wave height as a function of wind speed. To minimize the influence of man made noise sources, most data was acquired at night when ship and boat traffic through the harbor was least. The data sets acquired are described in the next chapter, along with the processed results, the environmental data collected, and analysis.

Chapter 4

Data and Analysis

4.1 Data Catalog

Data was acquired over the course of four months, from January to April of 1999. While the apparatus was constantly being refined to make operation of the system simpler, no changes were made to the actual data collected which would make comparison across this period invalid. The following is an inventory of the acoustic data sets collected:

Date	Frequency	Duration	# of Data Series	Wind	Waves
01-28-99	LF	5 hours	400	NE 8-10 kts	some whitecaps
02-19-99	MF	5 hours	400	NE 8-10 kts	some whitecaps
02-23-99	LF	14 hours	1000	SSW 4-8 kts	small waves
03-18-99	MF	21 hours	477	SW 12-15 kts	1-2' whitecaps
03-31-99	LF	16 hours	384	W 20-25 kts	2' breaking waves
04-01-99	MF	10 hours	241	SSW 10-12 kts	whitecaps
04-02-99	MF	6 hours	141	N 1-3 kts	ripples
04-07-99	MF	3 hours	56	W 10-12 kts	whitecaps

As described in Chapter 2, each acoustic data set is composed of data series, which consist of a toneburst, a Turyn, and a Barker ping recorded in succession. Also, each data series has a corresponding ambient noise measurement. For each set of acoustic data, the wind recorded at the Buzzards Bay Tower and posted to the National Weather Service website was downloaded for later analysis. These wind velocities are archived as ten minute direc-

tion and speed averages. In addition the Woods Hole tide gauge reading was recorded over each data set's duration, and the current at the transponder estimated from the Eldridge tables.

The matched filter output, wind, and ambient noise data for the 03-18-99 and 03-31-99 data sets are plotted on the following pages, with the remainder in Appendix A. As noted above, the wind speed data is low pass filtered at the source, and as shown here the noise data collected has also been filtered to match. In the matched filter plots, the data from each waveform is compiled into an image. Each horizontal band corresponds to a single matched filter output, the color of which varies with the output's absolute value according to the colorbar at right. The time series are organized by their chronological position within the data set, labeled by actual clock hours starting on the day on which the experiment was started. (Thus, for an experiment started in the evening and running over night, 0200 becomes 2600.) The time scale is converted to meters for easier range estimate precision analysis. Overlaid on the image plot are the ranges calculated by the direct path arrival detector (solid black) and the maximum matched filter output (dotted black). Although the different positions of the two transponders mean that the ranges are slightly different, the range windows are the same width for both frequencies. Likewise, the time scaling of the y axis is kept approximately the same for data sets of different lengths or different series spacing.

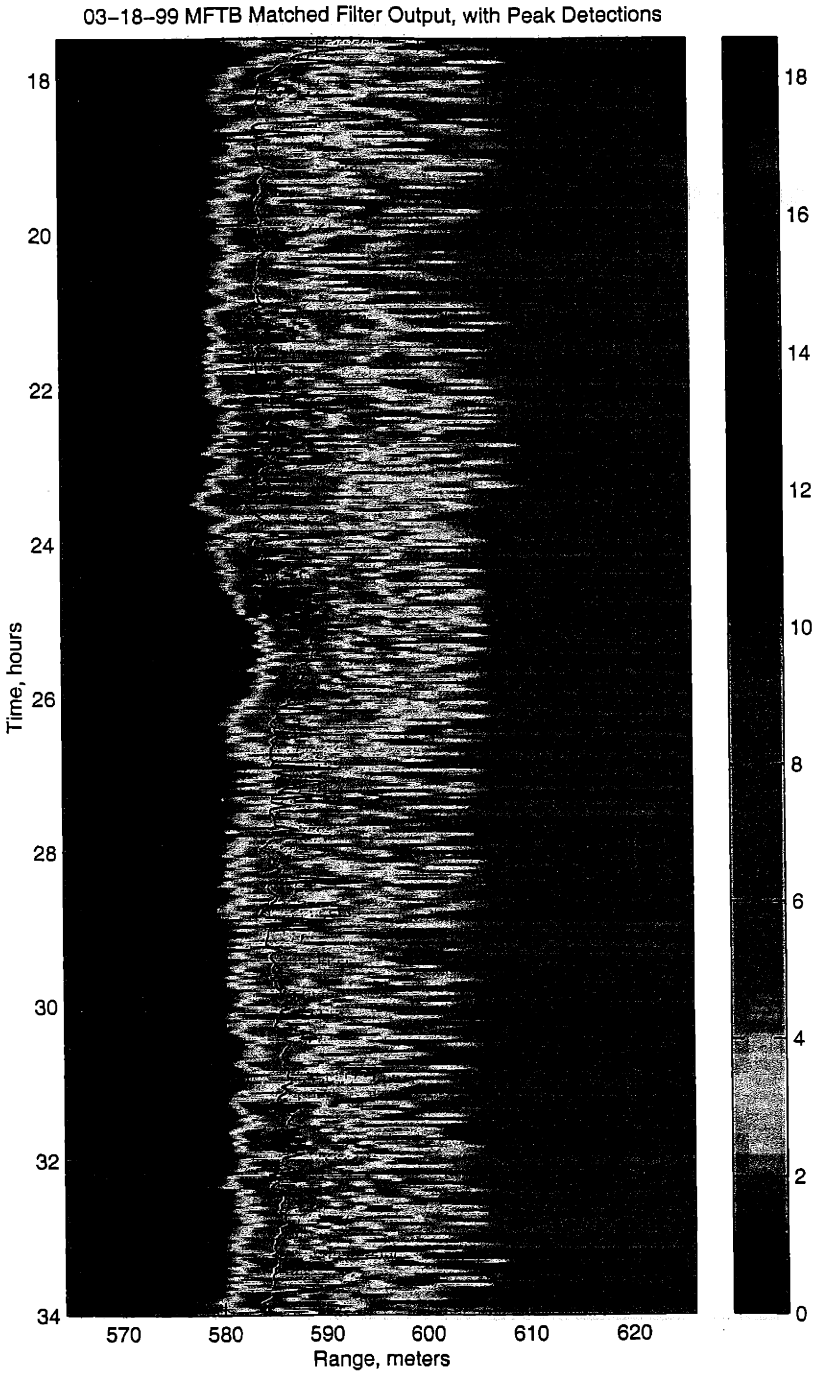

Figure 4-1: 03-18-99 Mid Frequency Toneburst

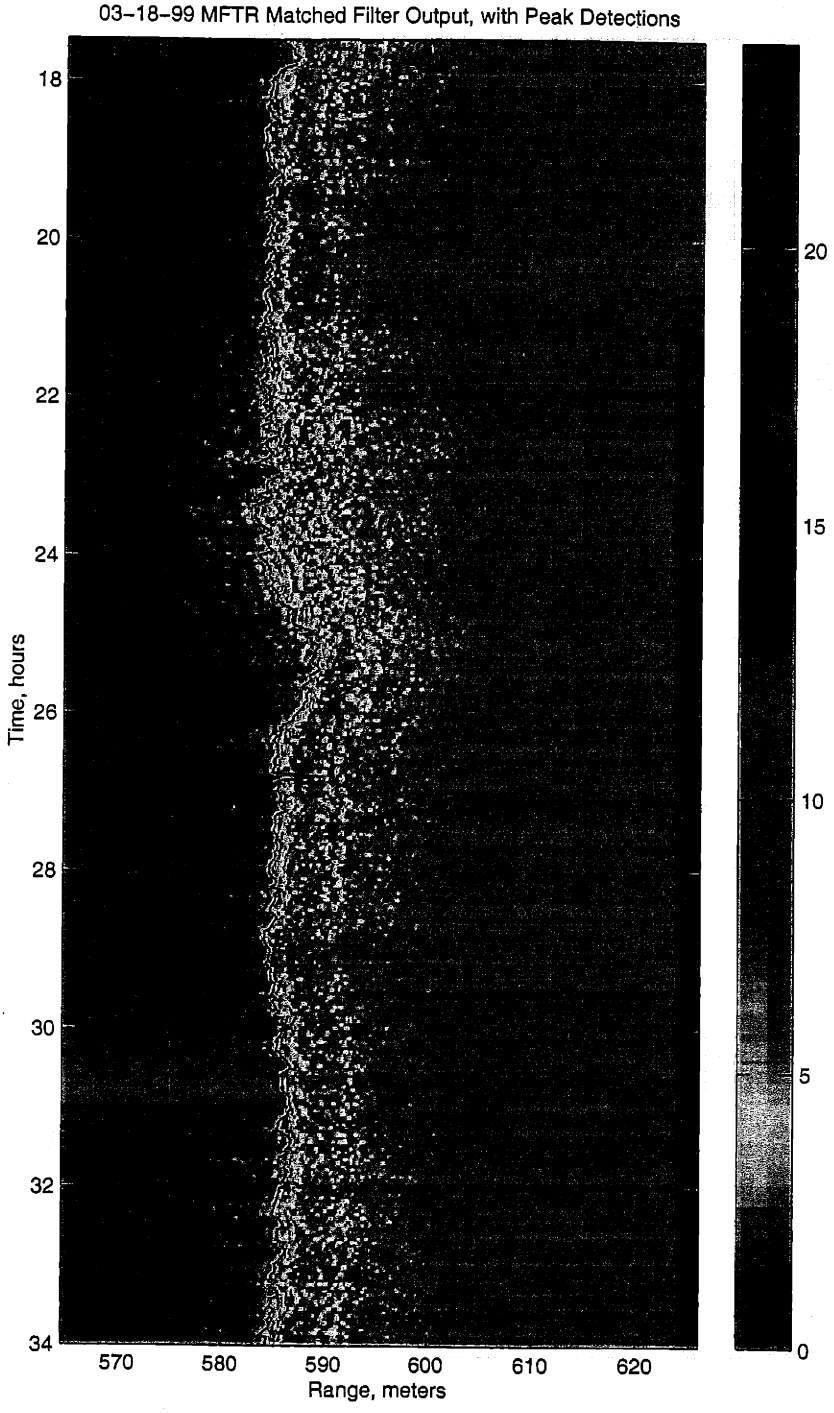

Figure 4-2: 03-18-99 Mid Frequency Turyn Code

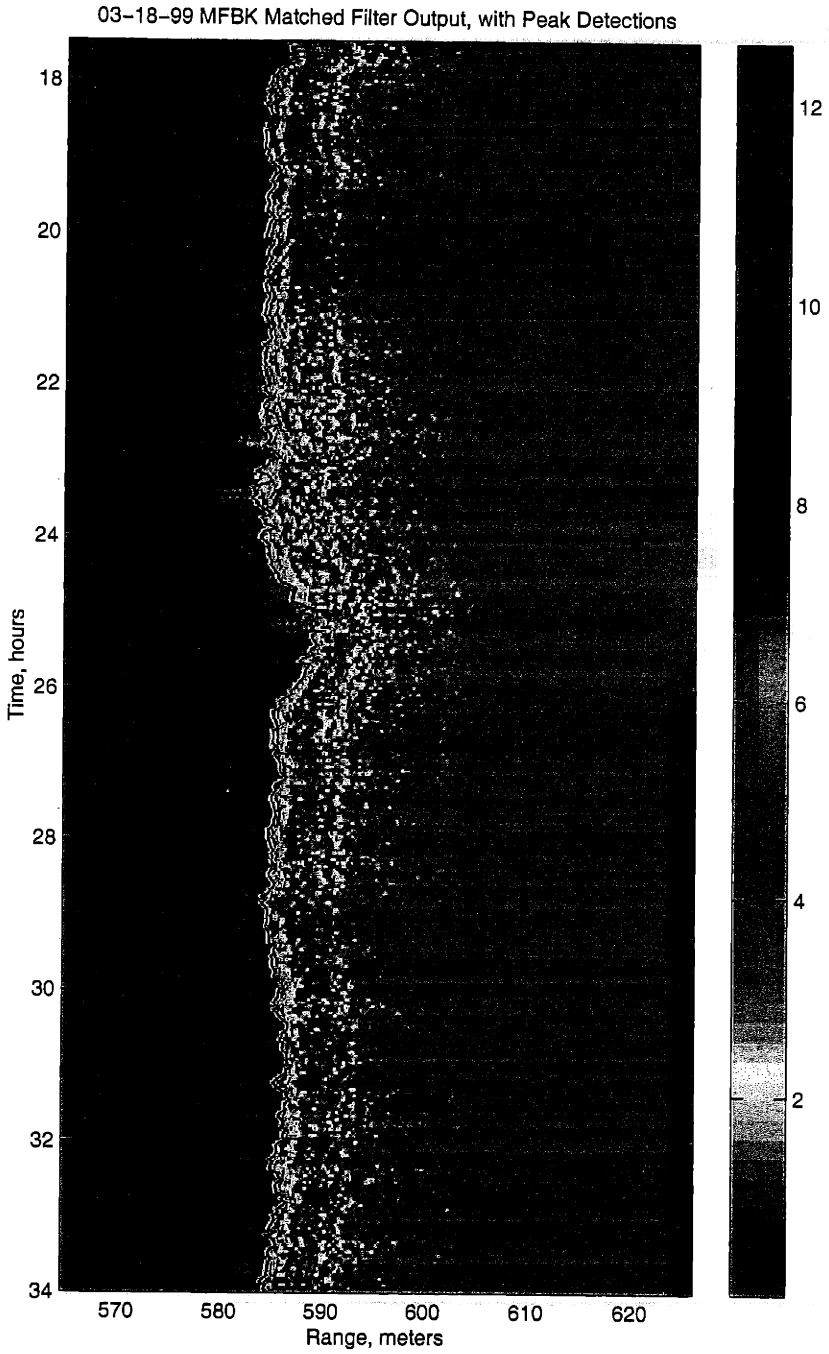

Figure 4-3: 03-18-99 Mid Frequency Barker Code

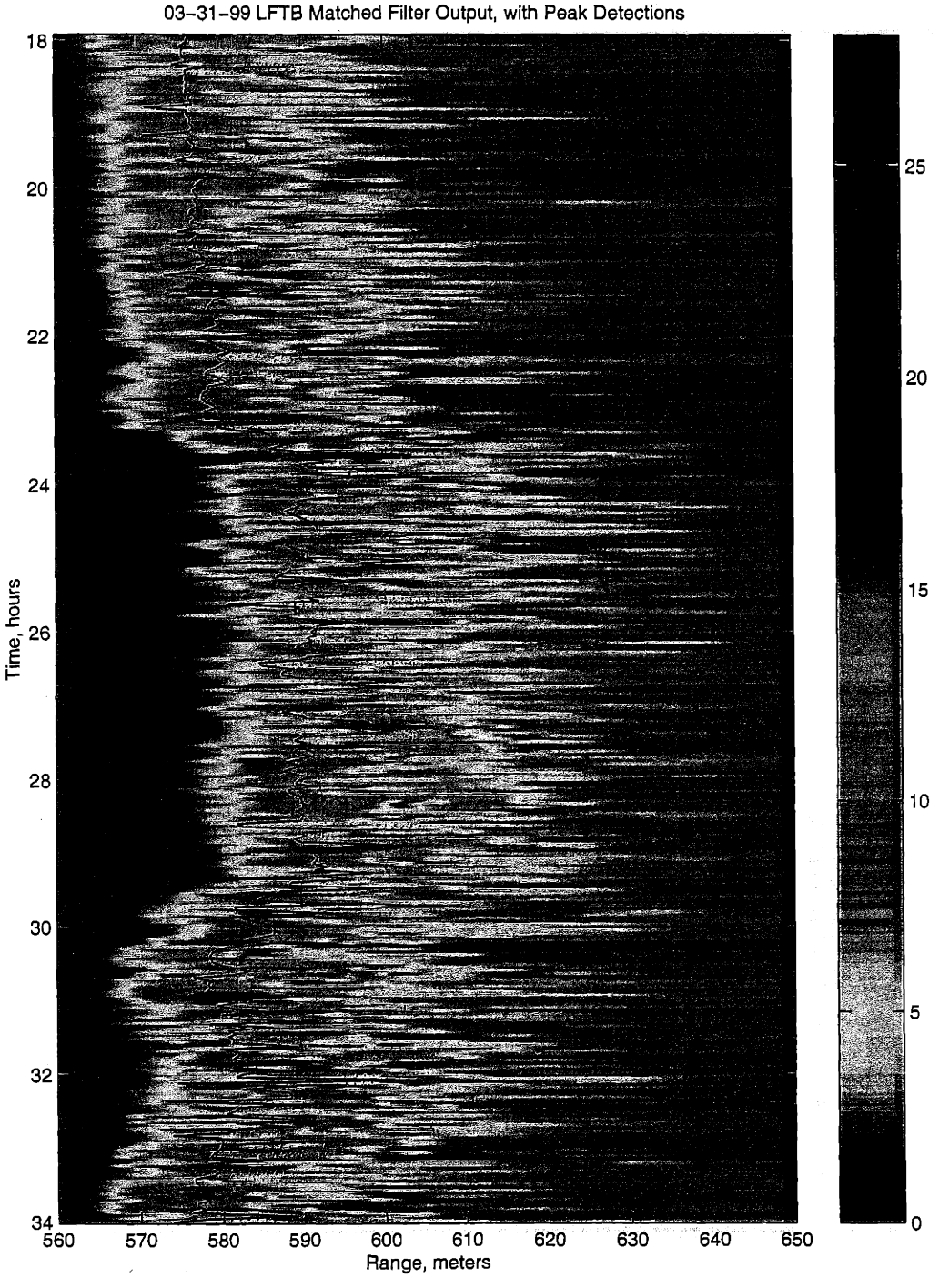

Figure 4-5: 03-31-99 Low Frequency Toneburst

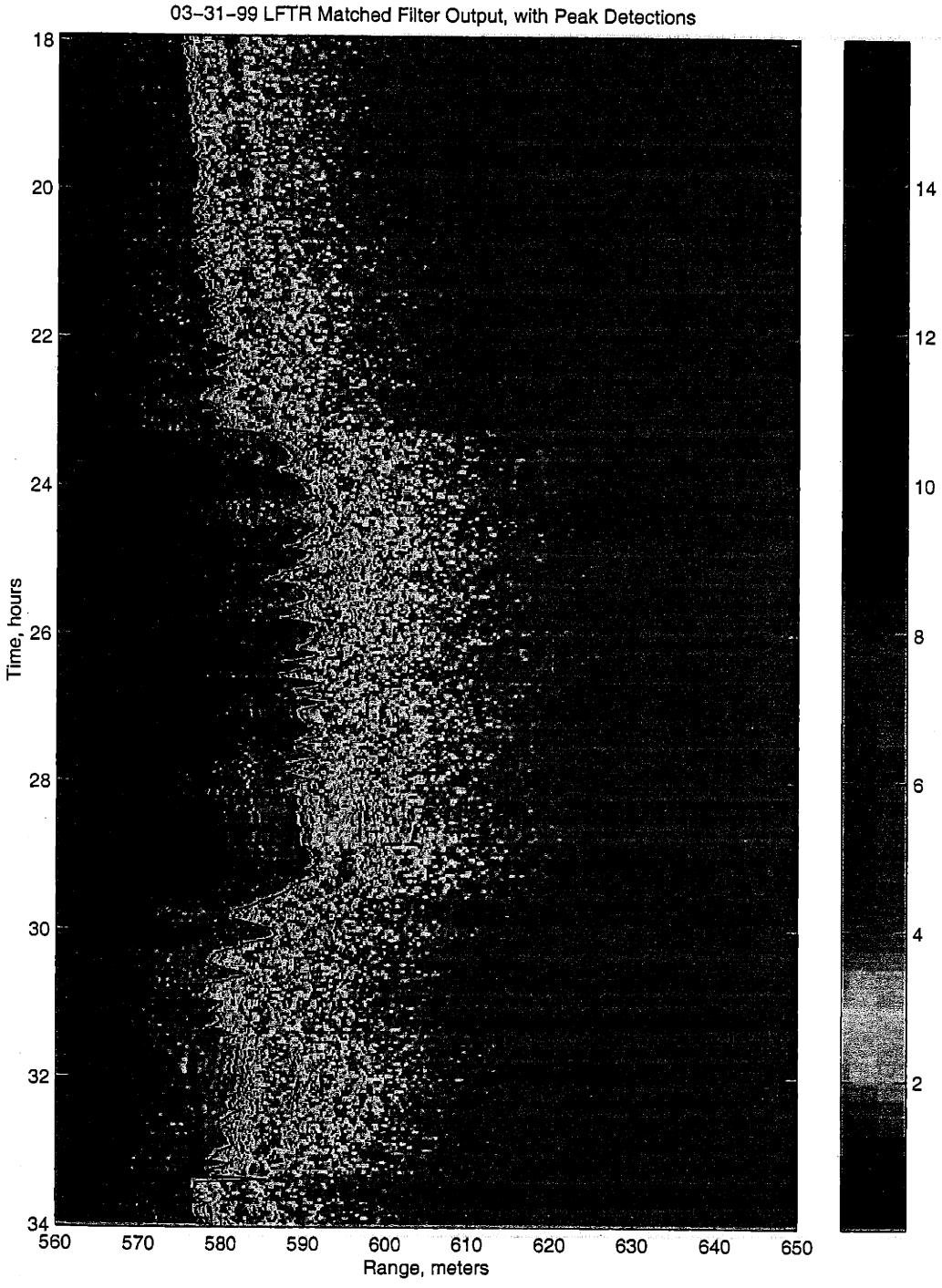

Figure 4-6: 03-31-99 Low Frequency Turyn Code

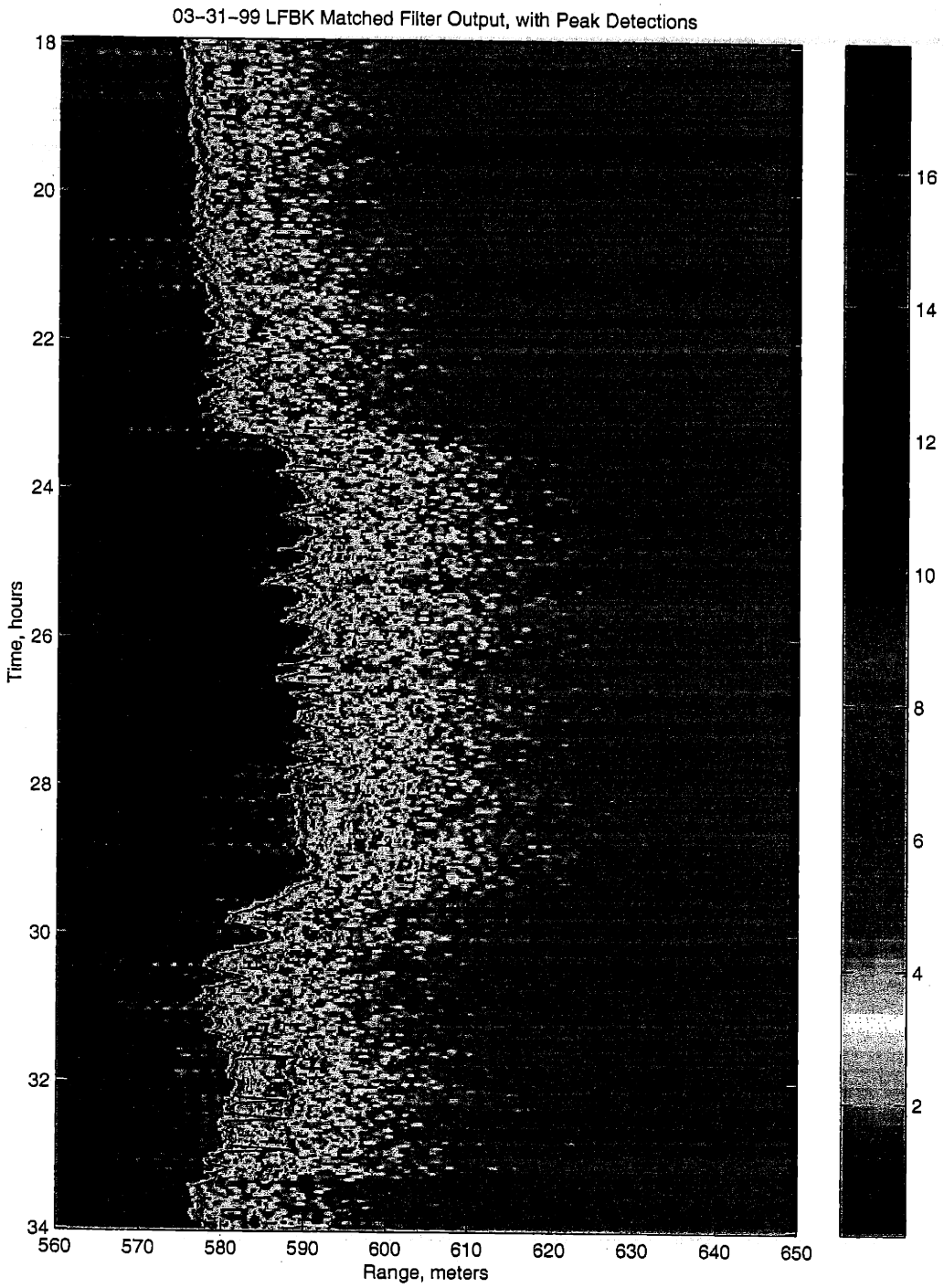

Figure 4-7: 03-31-99 Low Frequency Barker Code

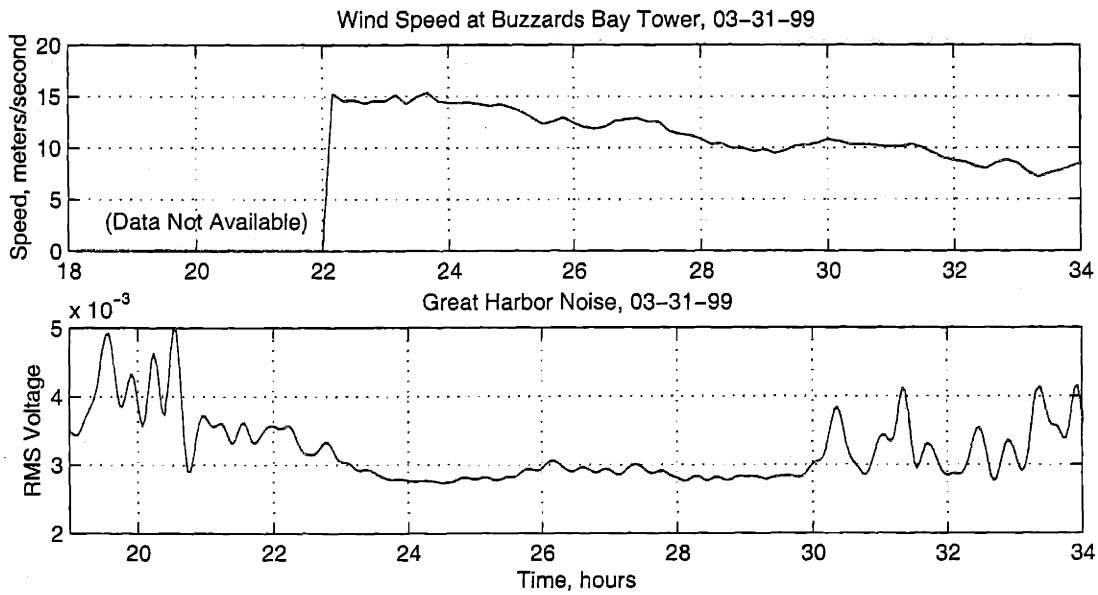

Figure 4-8: 03-31-99 Wind Speed and RMS Noise Voltage

Even without further processing, several preliminary results are evident. First, the precision afforded by the coded signals over the tonebursts is clear for all the data sets, regardless of frequency or environmental conditions. It would also appear that the Turyn code offers the finest resolution, as expected from the discussion of Cramer-Rao bounds in Chapter 2. The matched filter outputs from the MF signals have finer temporal resolution than those of the LF signals, another expected result. Strong delayed arrivals are seen in many of the data sets, and it would seem that the direct arrival detector works well. With just range estimate data, it would be impossible to tell if the fluctuations in the measurement were actual range shifts, or if the detector was simply hopping from one peak to another in the matched filter output. From these plots, which show the delayed arrival tracking separately from the first peak, it is clear that this is not the case.

Another interesting observation is the apparent lack of significant correlation between noise level and wind speed. This is a surprising result, given the strong dependence of noise on breaking waves that was discussed in Chapter 2. There are several conceivable explanations, including the possibility that the noise measurement is dominated by other sources, perhaps non-acoustic, which are stronger than whatever contribution is made by surface action. Without a relative measurement of noise level (as described in Chapter 2), it is not possible to compare this amount of noise with that predicted for coastal waters given the wind speed observed. Another possibility is the Buzzards Bay tower is too far away for wind speeds recorded there to be directly applicable in Woods Hole. Regardless, this observation suggests experimental error.

There are, however, some points of interest in the wind speed and noise data which are perhaps too strong to be coincidental. For example in the data collected over the night of March 31 to April 1, 1999, wind speed remains fairly strong and constant, but there is a significant decrease in noise level between 2200 on the 31st and 0600 the next morning. While Great Harbor tends to be fairly quiet in the winter, there is an active fishing fleet which transits the harbor each morning, passing directly over the path between the transponders and the receiver.

4.2 Tidal Influence

In the data sets covering longer time periods (for example, 02-23-99, 03-18-99 and 03-31-99), the range estimates exhibit some longer term periodicity than the typical ping to ping jitter. The magnitude of this oscillation and its low frequency suggest that it may be correlated with the forces of tide and current. This section investigates this relationship.

The data set collected on March 31 contains a particularly large range oscillation. Figure 4-9 shows the range estimated by the direct path arrival detector for the Barker coded signal, along with the water level and approximate current magnitude through Woods Hole. Qualitatively, it would appear that the abrupt range shift is connected to the tide and current. In the section on multipath ray propagation in Chapter 2, the results of the ray code demonstrated that the effect of tide height change on arrival time is minimal at this range and depth. However the 10 meter range of the error observed is consistent with the size of the watch circle for the low frequency transponder mooring as described in Chapter 3. Therefore, the most probable explanation is that eddie currents from the flow through the Hole caused the buoy to shift position to one side of the watch circle, and when the current switched direction, the buoy shifted to the other side. The buoy effectively converts the sinusoidal current signal to a square wave, as it stays pinned to one side until the current turns, and then immediately swings to the opposite edge of its range.

It is also interesting to note that while the 03-31-99 data set exhibits this large range shift, data collected over a time period of more than 12 hours on February 23 does not. This is initially puzzling, as one would expect any tidal effect to be evident over a period of 12 hours or more. However, the result is not surprising when the phase of the moon, full on March 31st and half on February 22, is considered. Since the tidal current velocity is directly related to the tidal range, the current is greatest at a spring tide and least at a neap tide. This correlation lends further evidence to the conclusion that the range shift in the 03-31-99 data set is the result of a tidal signal effecting the buoy location, and not a higher order acoustic phenomenon.

Since the purpose of these experiments was to investigate the sources of error in range estimation resulting from environmental effects on the properties of the acoustic waveguide, for best results error from other sources should obviously be minimized. Ideally, a fixed transponder would have been used to eliminate the possibility of this type of extraneous

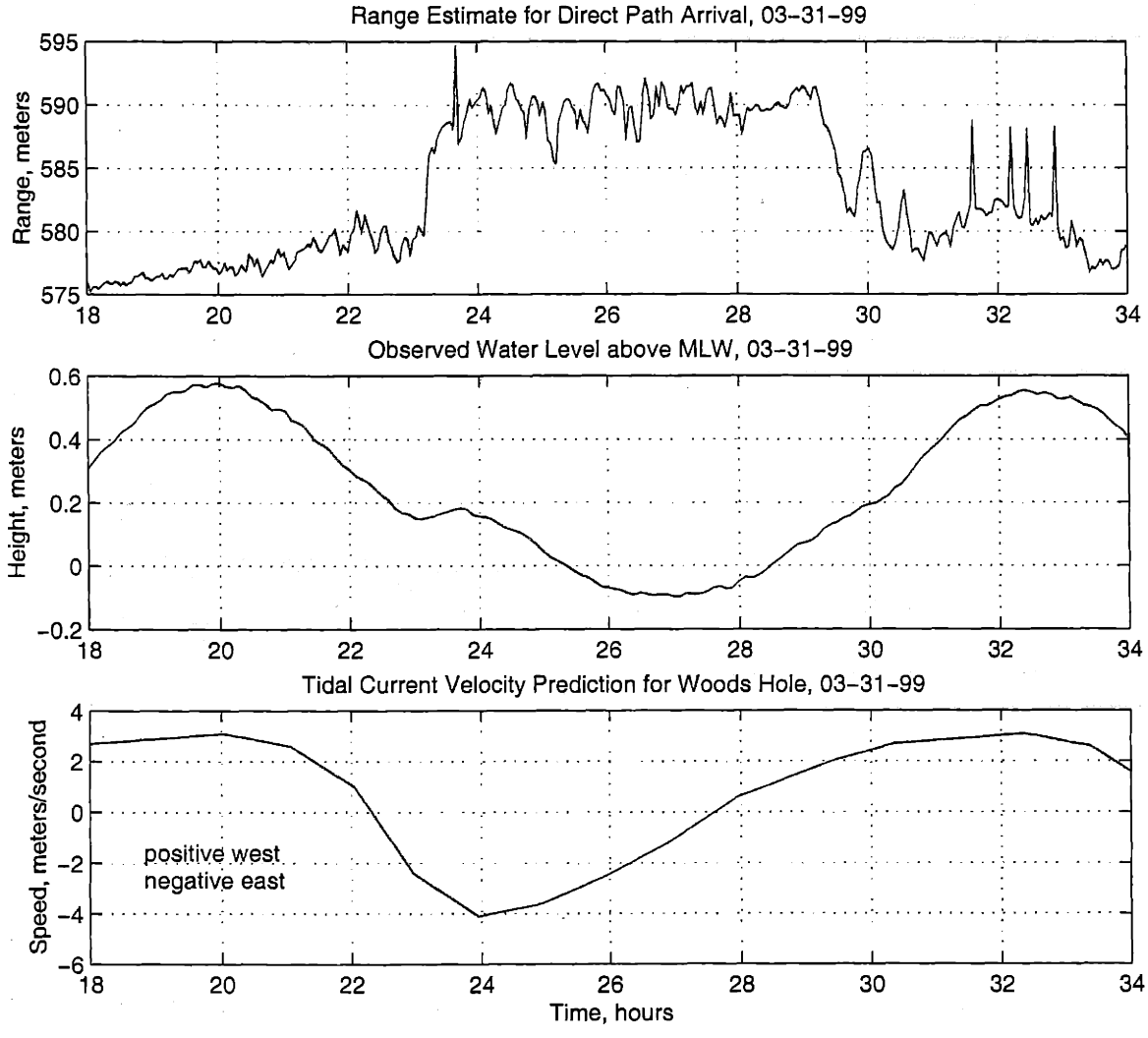

Figure 4-9: Range, Tide Height (from NOAA tide gauge), and Current Velocity (from Eldridge) for the 03-31-99 Data Set

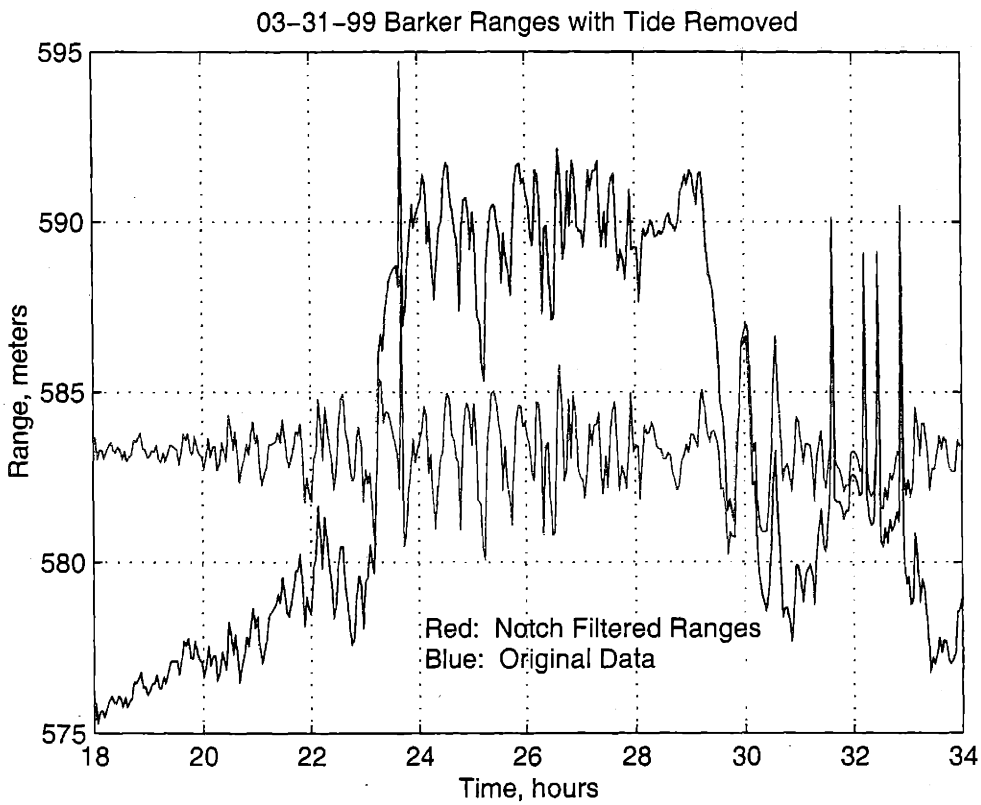

Figure 4-10: Notch Filtered 03-31-99 Ranges

error. Fortunately, the periods of the tidal components are well known, so they can be removed from the data by filtering. Figure 4-10 shows the range series in time for the 03-31-99 Barker codes after notch filtering to remove the M_2 (12.42 hour lunar) and S_2 (12 hour solar) tidal components. The original range data is overlayed to show that while the shift is gone, the higher frequency variability remains. While there are additional signals present in the tide at lower frequencies, at the spring tide the M_2 and S_2 combine for the greatest impact. For the purposes of range estimate statistics which follow, all data sets had these frequencies removed. While many of the other range estimation variances are probably also linked to the movement of the transponder rather than changes in the propagation of the acoustic signal, this strong tidal effect is the only one that can be positively identified.

4.3 Signal Precision

After removing the tide error, which we can positively identify and is unrelated to the characteristics of the acoustic waveguide under examination, the standard deviations in range for the data are contained in the following table, and plotted in Figure 4-11. The observed conditions of each data set are included from the table at the beginning of the chapter.

These results strongly support the original premise that coded signaling offers significant improvement in precision over toneburst waveforms. For the range of these experiments, Barker and Turyn codes offer an improvement of approximately a factor of 3 over the standard waveforms, and the Barker and Turyn codes are related by a factor of approximately 1.4. As discussed in Chapter 2, the Cramer-Rao bounds provide an absolute lower limit on the precision of the estimation of a parameter for a given noise level, and give the fundamental relationships between the precision of the codes. In general, these relationships appear to be preserved in the data.

Data Set	Frequency	Wind	Waves	Waveform	σ, meters
01-28-99	LF	NE 8-10 kts	some whitecaps	Toneburst	3.0159
				Turyn	1.4969
				Barker	0.7672
02-19-99	MF	NE 8-10 kts	some whitecaps	Toneburst	0.6692
				Turyn	0.4393
				Barker	0.4865
02-23-99	LF	SSW 4-8 kts	small waves	Toneburst	3.2138
				Turyn	0.8113
				Barker	0.9100
03-18-99	MF	SW 12-15 kts	1-2' whitecaps	Toneburst	1.3131
				Turyn	0.4055
				Barker	0.5647
03-31-99	LF	W 20-25 kts	2' breaking waves	Toneburst	3.5262
				Turyn	1.0939
				Barker	1.2624
04-01-99	MF	SSW 10-12 kts	whitecaps	Toneburst	2.3193
				Turyn	1.0136
				Barker	1.0208
04-02-99	MF	N 1-3 kts	ripples	Toneburst	2.5729
				Turyn	0.9121
				Barker	1.1202
04-07-99	MF	W 10-12 kts	whitecaps	Toneburst	1.2347
				Turyn	0.5707
				Barker	0.8328

Some interesting insight is gained from comparing the standard deviation in range of the data with the standard deviation calculated by the Cramer-Rao bounds, as shown in Figure 4-12. This ratio gives a figure from which the performance of a given waveform under a given set of conditions can be estimated. The lower the ratio, the closer to the best case performance. If the errors inherent to the field applied equally to all three signals, the three ratios would be equal in all cases. As shown above, it would appear that the toneburst signal comes closest to its ideal, and the Turyn is furthest, despite offering the tightest raw

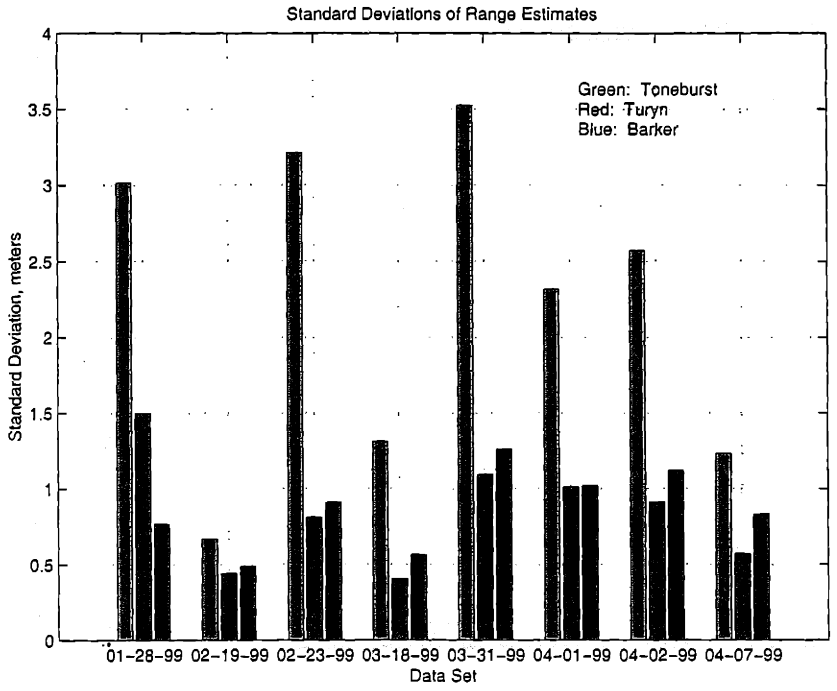

Figure 4-11: Standard Deviations in Range, Grouped by Data Set

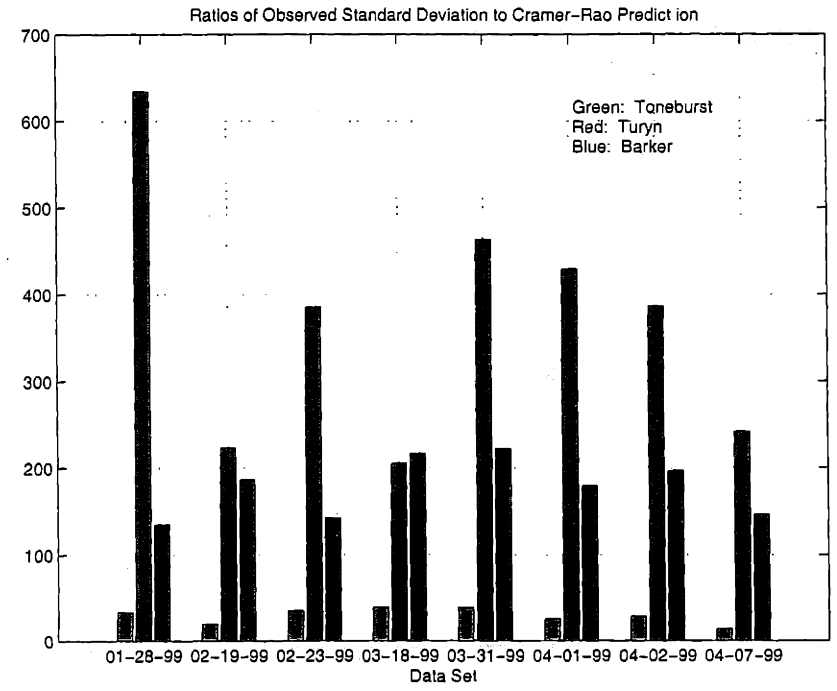

Figure 4-12: Ratios of Observed Standard Deviations to those predicted by Cramer-Rao Bound

standard deviation. This is not surprising, given the broad bandwidth of the Turyn code, which is slightly limited by the bandwidth of the transducers, output stages, and receivers. Clipping of a waveform's spectrum has a negative impact on the quality of its matched filter output, and since the Turyn code is the widest of the three, it is the first to be degraded by the narrow resonance curve of the transducers. In the same way, the toneburst, which has the most narrow band signal, has the least to lose from limited bandwidth. The Barker code falls between the two, as it is a wider signal than the toneburst and as such more affected by "real world" constraints, but not to the extent of the Turyn.

Again, there is insufficient data to draw conclusions on the effect of surface conditions on performance. However, the three ratios do appear to track together (with the exception of the 03-18-99 data set), which suggests that the same factors influence each.

4.4 Identifying Multipaths

Careful examination of any of the plots at the beginning this chapter reveals a strong multipath signature behind the direct path arrival, and in some cases this arrival is so strong that it draws the direct path arrival detector off the first arrival and onto the second. Throughout the data, there is one particularly strong arrival which appears to be responsible for majority of the mis-identifications. This arrival is delayed from the direct path by an equivalent distance of approximately 10 meters, and while it is often quite strong, there are also periods where it fades away almost completely. For example, in the 03-31-99 data set, there is a very strong occurrence near hour 32, but by hour 34 it has almost completely disappeared. Figure 4-13 shows the matched filter output from these two closely spaced but dissimilar time series.

One means by which to identify the sources of the remaining error in the range estimates is to follow the track of the rays which arrive at the receiver, and identify those ray paths which do not propagate. In this case, knowledge of the surfaces with which the ray in question interacts could offer some insight into its fading. As utilized in Chapter 2, a simple ray propagation model is a useful tool with which to identify the eigenrays, or allowed paths. Figure 4-14 shows the allowed rays for the situation of these experiments. The source is located four meters beneath the surface, the receiver three meters below, and the bathymetry and range match that along the acoustic path in Great Harbor. Given

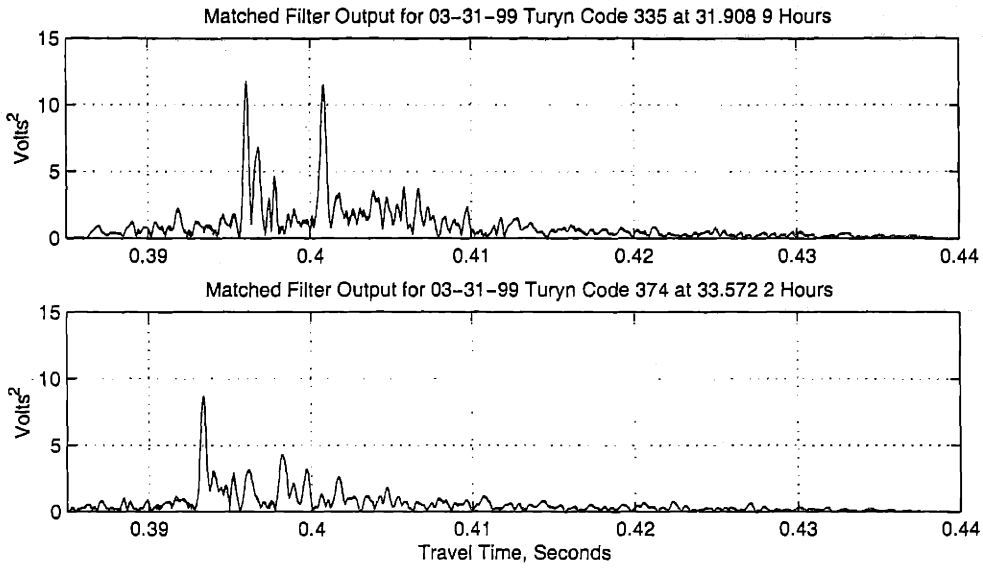

Figure 4-13: Strong Delayed Path Arrival in 03-31-99 Data Set

the relatively wide beamwidth of the transducers, a fan of rays 60° across was allowed to propagate. In this environment, without any consideration for transmission or reflection loss, 16 discrete eigenrays can propagate from source to receiver.

Although this plot shows the physical path along which the rays travel, for identification of rays without an array at the receiver to resolve their angles, the time of each ray's arrival is more useful. Figure 4-15 shows the arrival times for Great Harbor bathymetry.

From this, it is possible to assign matched filter output peaks to eigenrays. Close examination of Figure 4-13 reveals a group of slightly delayed arrivals at the base just after the first peak. These most likely correspond to the group of eigenrays which propagate with shallow positive angles, making one or two reflections on their way to the receiver. Since the pathlength of these eigenrays in a long shallow waveguide is only slightly longer than the direct path, the individual peaks corresponding to these arrivals are almost completely obscured.

The first peak and the second are separated by approximately 5 ms, for which there is a clear match in the eigenangle diagram. Two arrivals which leave the receiver at an angle of approximately -9° arrive at the receiver nearly simultaneously and 5 ms behind the first arrival. These would appear to be the only eigenrays which meet the arrival time criteria. After the second strong peak, there are several smaller peaks which correspond to eigenrays

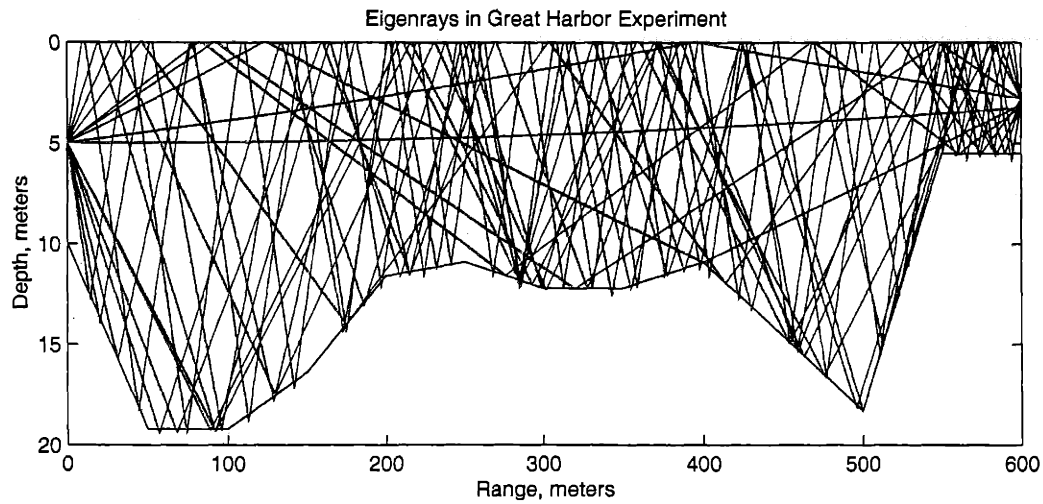

Figure 4-14: Allowed Eigenrays in a 60° fan from Source to Receiver in Great Harbor

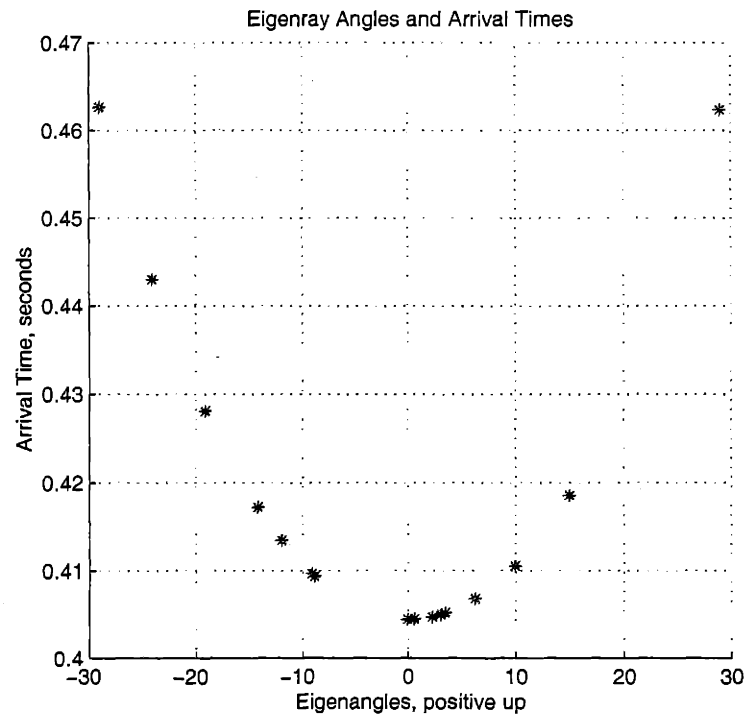

Figure 4-15: Allowed Eigenangles in a 60° fan in Great Harbor

of increasing angle. However, the matched filter output of these rays deteriorates rapidly, as the higher angles necessitate more reflections, each of which degrades signal strength.

Having identified the probable path by which the arrival in question traveled, the question of what causes it to fade in and out remains. The second panel of 4-13 shows the same peak structure, but the delayed arrival is considerably weaker. Since the demonstration in Chapter 2 showed clearly that small fluctuations in receiver or source position in a long thin waveguide have little impact on the multipath structure, it is unlikely that this delayed arrival's propagation is the product of a momentary alignment of bottom bathymetry as the transponder mooring moved in the waves and current. It is possible, however, that the abnormal strength of this arrival is the result of momentary constructive interference between the two, as the phase relationship necessary for this to occur would be dependent on minute changes in position or propagation. It is also likely that the conditions at the reflecting interfaces play a part in the signal's strength, as discussed in the section on reflection loss in Chapter 2. While the bottom reflection characteristics are fixed for the time scale of this experiment, the surface is dynamic, which can lead to reflection characteristics and modes of propagation which are dependent on surface conditions for their survival.

4.5 The Effect of Wind and Waves

As discussed at the end of Chapter 2, the effects of surface action act in opposition on the precision of range estimation through ambient noise and multipath propagation. While increased wave activity increases reflection loss and therefore decreases the propagation of multipath arrivals, simplifying the arrival selection process, it also adds ambient noise to the water, decreasing the range resolution according to the Cramer-Rao bound. The In-Band and Out-of-Band Signal to Noise Ratios are two quantities which can help to separate these two effects.

For the purposes of this investigation, the IBSNR is a measure of the multipath arrival strength. It is calculated by taking a window of several signal durations around the direct path arrival in the matched filter output, counting the number of peaks which exceed a threshold (here, 10%), and dividing the window width by that number. The units of such a value are arbitrary, but give an estimate of the strength with which reflected paths propagate.

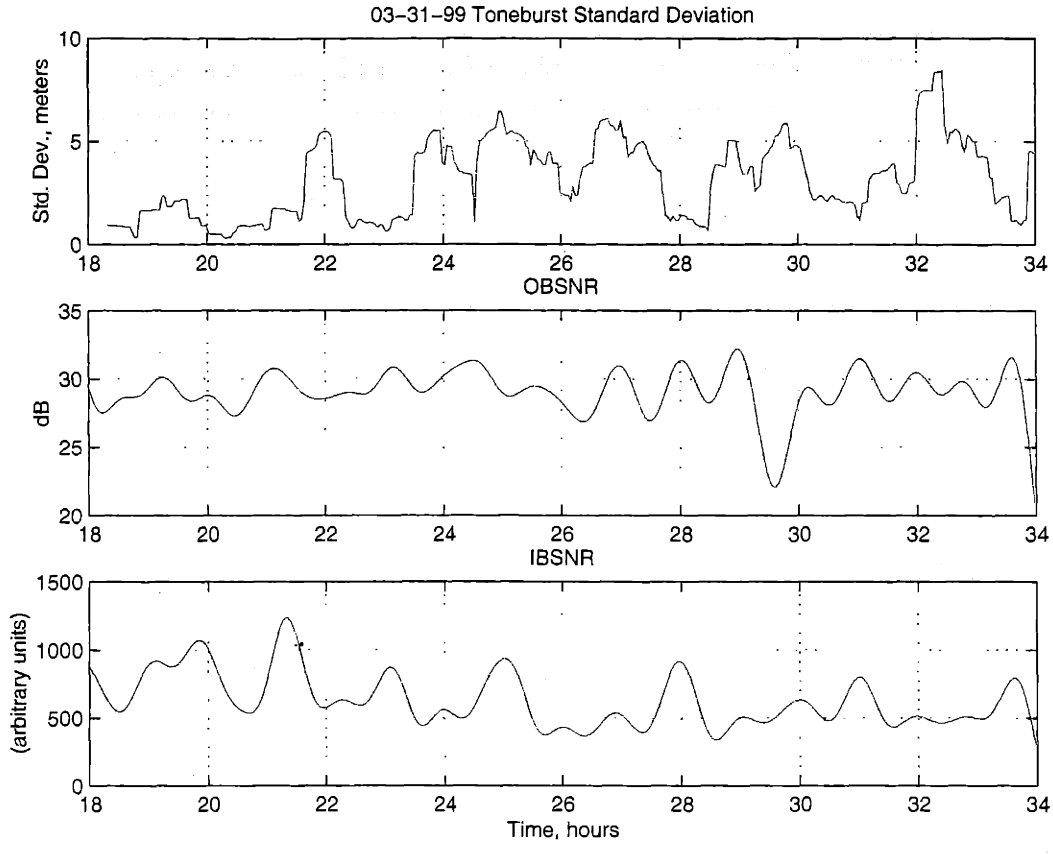

Figure 4-16: Standard Deviation, In-Band, and Out-of-Band SNR for 03-31-99 Toneburst

OBSNR, calculated much the same way as the usual SNR, is found by taking the logarithmic ratio of the direct arrival peak to the mean value of the matched filter output before the arrival of the signal. In noisy environments, the SNR of the output of the matched filter must also decrease, so this value should follow the impact of ambient noise on the matched filter output.

Figures 4-16, 4-17, and 4-18 show the relationship of these values to the precision of the range estimate for the 03-31-99 data set. In these plots, the running standard deviation in range over 11 samples is calculated and plotted in the top panel.

From this data, in conjunction with the wind speed data shown in Figure 4-8, the potential usefulness of these parameters begins to emerge. While the long term trend is difficult to spot in the Toneburst case, specific events in the range estimate precision record correspond with similar changes in the IBSNR and OBSNR. In the Barker and Turyn codes, however, there is a definite shift in influence between the two parameters as the wind speed

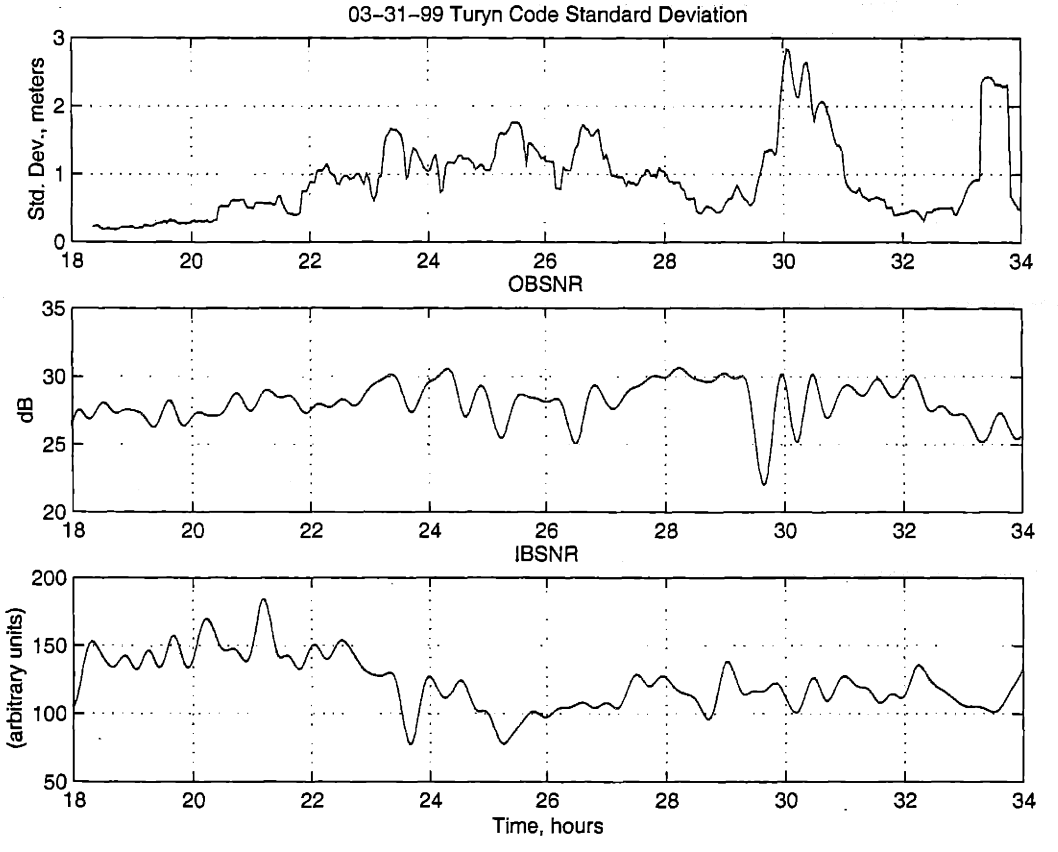

Figure 4-17: Standard Deviation, In-Band, and Out-of-Band SNR for 03-31-99 Turyn

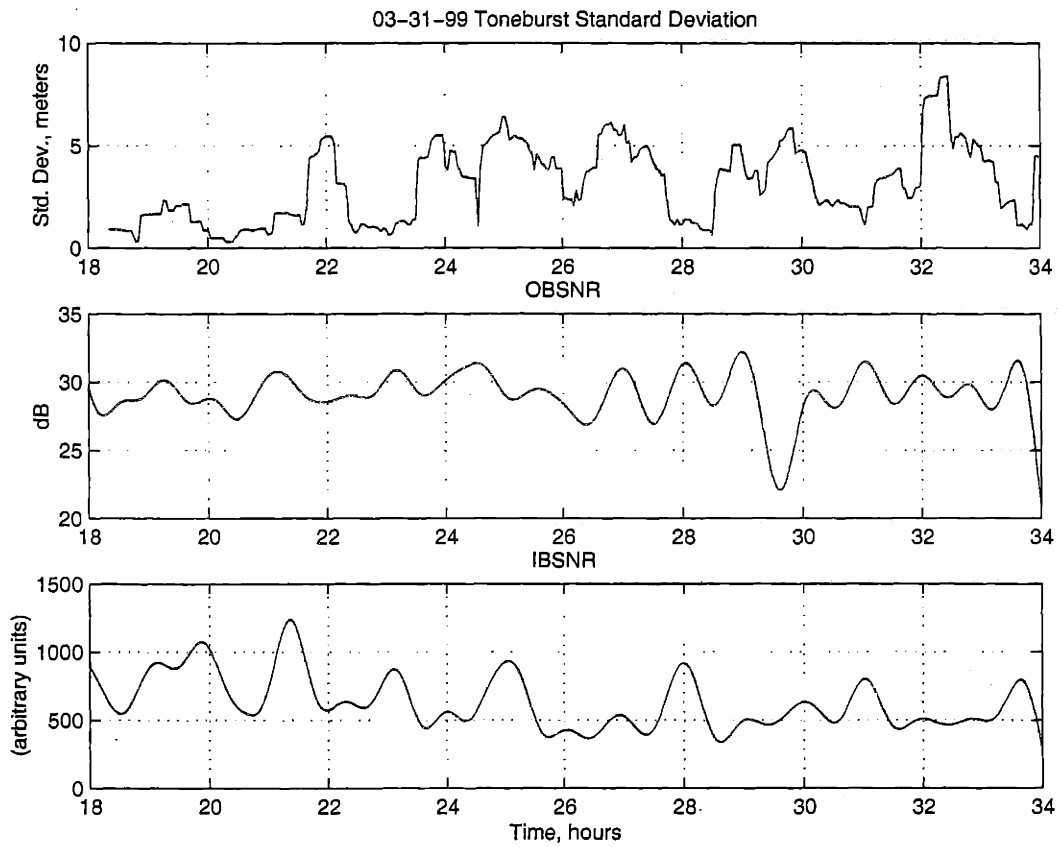

Figure 4-18: Standard Deviation, In-Band, and Out-of-Band SNR for 03-31-99 Barker

changes. For roughly the first four hours, wind speed was steady (this was observed at the time the data was collected, although data from the Buzzards Bay was unavailable) at approximately 15-20 knots. For this time, we see that the OBSNR is low, indicating high ambient noise, and the IBSNR is relatively high, corresponding to poor multipath propagation. The standard deviation during this time is quite good, indicating that for this combination of wind speed and frequency, the positive impact of low multipath level outweighs the loss of precision associated with high ambient noise. However, as the wind speed decreases, the situation reverses: IBSNR declines as multipath propagation improves, while OBSNR improves as the noise level drops. For these more quiet, flat conditions, short term changes in standard deviation appear to be primarily related to changes in the OBSNR, such as the event at 30 hours.

For the mid frequency case, similar correlation was observed. However, the overall connection between IBSNR and the standard deviation of the range estimate appeared to be significantly less, suggesting that the higher frequency was less susceptible to multipath interference under all surface conditions.

4.6 Signal Selection

Though it has been conclusively shown that for the conditions in Great Harbor coded signals offer tangible benefits in precision over tonebursts, the question of which code and which frequency to use under different environmental conditions remains. Ultimately, this is an issue of the tradeoffs between the resolution and robustness of the signaling and range estimation systems.

Figure 4-12 provides the best indication of the relative robustness of the signals tested, while the resolution shown in Figure 4-11 indicates the resolution and the results from the previous section hint at the causes of loss of precision under different conditions. With this data, it appears that the MF Turyn code provides the most consistently precise range estimate, but while this signal resists multipath well (due to its frequency, at which reflection loss is high, as noted above), it is susceptible to ambient noise. While the bandwidth of this signal gives it the finest precision, the loss of sampling accuracy and bandwidth clipping of the hardware reduces the energy recovered. Therefore, while its matched filter output is narrow and relatively free of multipath arrivals, the OBSNR is high. While this is an

acceptable tradeoff in the situation tested in Great Harbor (as SNR is inversely related to the variance of the estimate, while bandwidth is related by its inverse squared, according to the Cramer-Rao Bound), at longer range or in a noisier environment this signal's lack of robustness would hurt performance. On the other hand, the toneburst signal is extremely robust at both low and mid frequency, coming the closest to its Cramer-Rao limit of the three codes, but this limit is inherently lacking in precision. In the middle, the Barker code gives resolution nearly as good as that of the Turyn, but as discussed above, with a smaller loss in accuracy as ambient noise increases and less loss due to limited hardware bandwidth, allowing it to come closer to its theoretical maximum performance than the Turyn, but with greater resolution than the toneburst.

Chapter 5

Conclusion

While a great deal of work was required to generate the data and results presented in this thesis, there are certainly ample areas requiring attention in the future. In this final chapter, both noteworthy results and shortcomings of this research are highlighted.

5.1 Accomplishments

First and foremost, this thesis accomplishes its stated goal of providing a quantitative study of arrival time estimation of coded signals. As described in the results, Barker and Turyn coded signals do offer the improvement advertised, at least under the relatively benign conditions of Great Harbor.

In addition, some progress was made in identifying the sources of error which affect these coded signals. The relationship between the maximum resolution, as given by the Cramer-Rao bound, and practical results was addressed, and comparison of the properties of Barker and Turyn codes was made. The IBSNR and OBSNR parameters were introduced, and demonstrated to be somewhat useful in comparing the relative effects of multipath interference and ambient noise on the overall precision of range estimation.

Finally, the relationship of the codes and toneburst was explored in terms of the tradeoff between robustness and resolution. Without more data at various signal strengths, noise levels, and multipath structures, an exact metric by which to decide the optimal combination of code and frequency for the given conditions cannot be derived. However, from the data available it appears that while the Turyn code at the mid frequency gives the highest resolution, it will quickly be degraded by the combination of noise and transmission loss at

longer range than that tested here. The Barker code, in both frequency bands, seems to offer the best all around combination of resolution precision and signal robustness.

As noted throughout the previous chapters, the conditions of the experiments in Woods Hole represent only one of any number of possible field environments. The shallow, well mixed water allowed many issues of range and depth dependent sound speed to be safely ignored, but these factors are very much present in many field operations. Fortunately, while the environment may complicate and degrade performance, the basic relationships between codes and tonebursts discussed in this thesis should remain valid.

5.2 Areas for Future Work

While the theoretical background of the possible causes of error in coded signaling were described, the field data did not clearly support any of these scenarios. This was probably due to the fact that the effects were more subtle than anticipated, and thus required more sensitive and direct instrumentation. First and foremost, the error introduced by the moving transponder buoy made isolation of other causes of range estimation error difficult, even once the easily removed tidal signal was addressed. Any future work should utilize a transponder system fixed in space, or with a sufficiently accurate positioning system that absolute range to the receiver is known more accurately than the required acoustic signaling resolution.

In order to assess the impact of environmental variables such as wind, waves, and current on range estimation, better knowledge of these quantities is required. At the inception of this research, it was thought that the available information on wind speed would be sufficient to extract the effect of waves on acoustic propagation through reflection loss and noise. Any new work in this area should incorporate a direct measurement of wave height, or at the very least, local wind velocity.

Ideally, an experiment such as this would also address range dependence of the transponder signal. In the data taken off the WHOI dock in November of 1998, it appeared that precision was actually degraded as range decreased. This result may be explained by the fact that at short range, multipaths have a greater path length difference from the direct path than at long range, and are also less attenuated by transmission loss and reflection loss. This could create a situation with a number of strong signals spread behind the direct path. However, without more than 10 time series at each range and waveform, no meaningful

conclusions could be reached.

Finally, part of the difficulty with the system as implemented was that the maximum range was constrained by the physical dimensions of the harbor and the need to have the receiver on land. This left no means by which to alter the signal strength at the receiver. Since many of the errors in range estimation occur when at the "edge of the envelope", with no way to create marginal conditions the results generated were too good for these issues to be explored. If greater range were available, it would be possible to create the situations which are of most interest in pushing back the boundaries of arrival time estimation.

Appendix A

Additional Data

The plots which follow display the data taken in Great Harbor in the spring of 1999 which was not shown in Chapter 4.

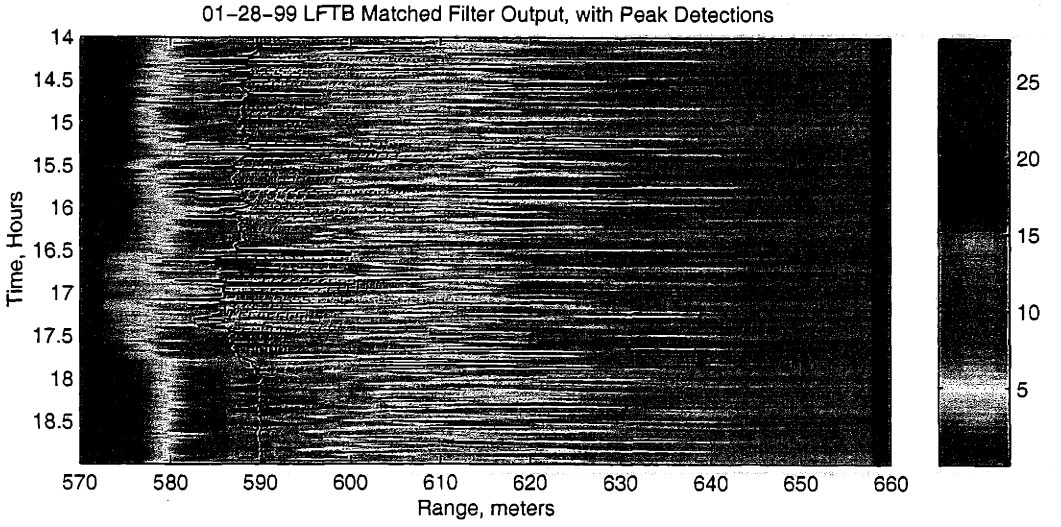

Figure A-1: 01-28-99 Low Frequency Toneburst

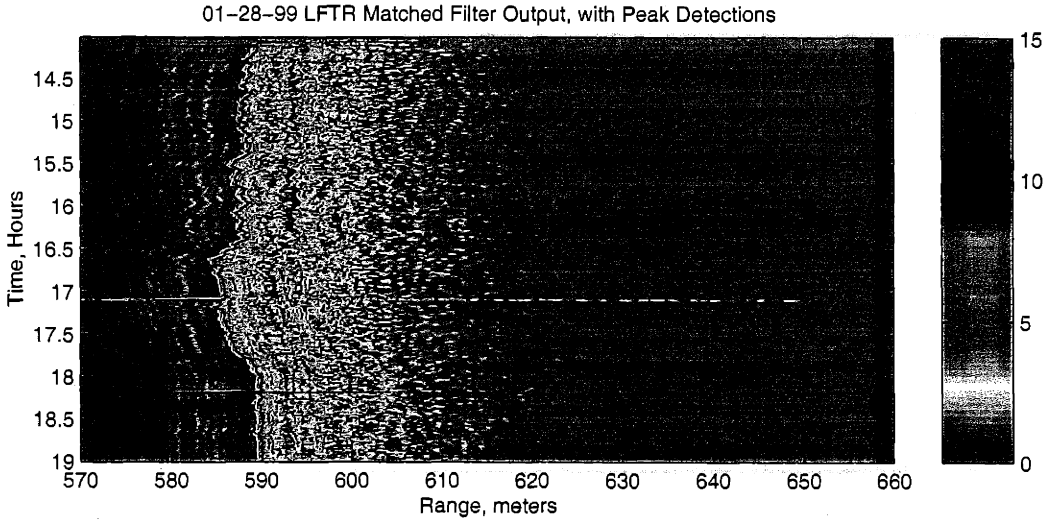

Figure A-2: 01-28-99 Low Frequency Turyn Code

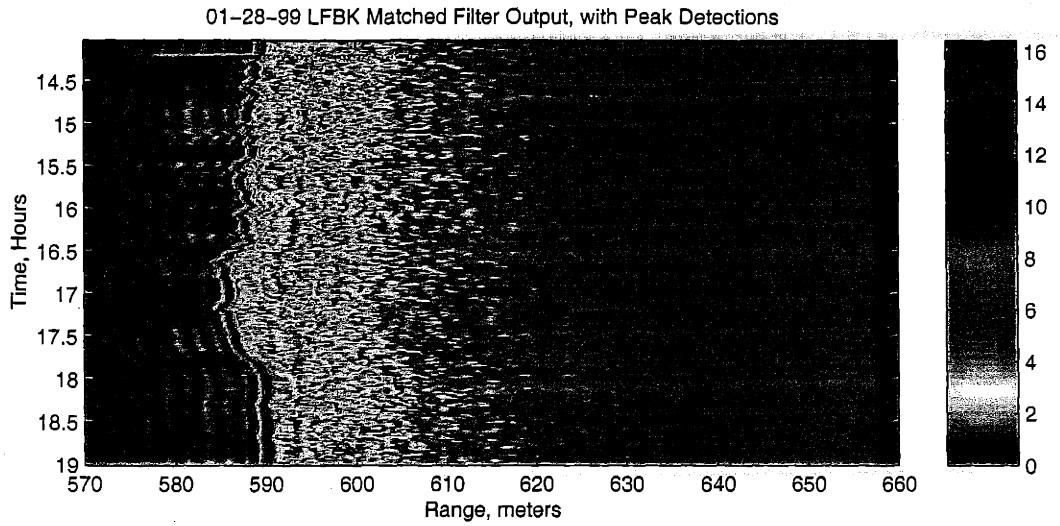

Figure A-3: 01-28-99 Low Frequency Barker Code

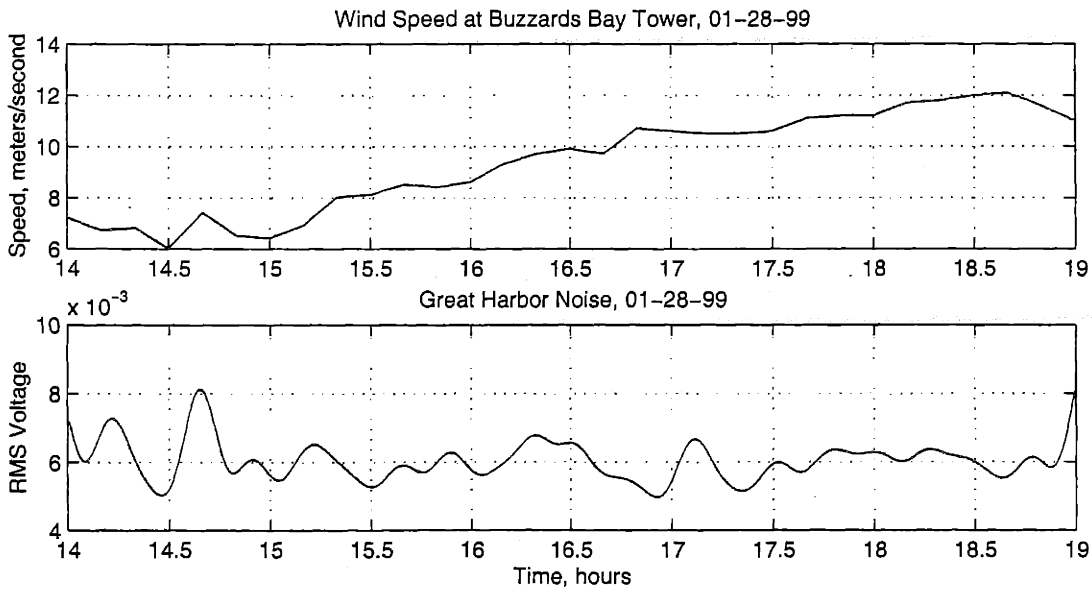

Figure A-4: 01-28-99 Wind Speed and RMS Noise Voltage

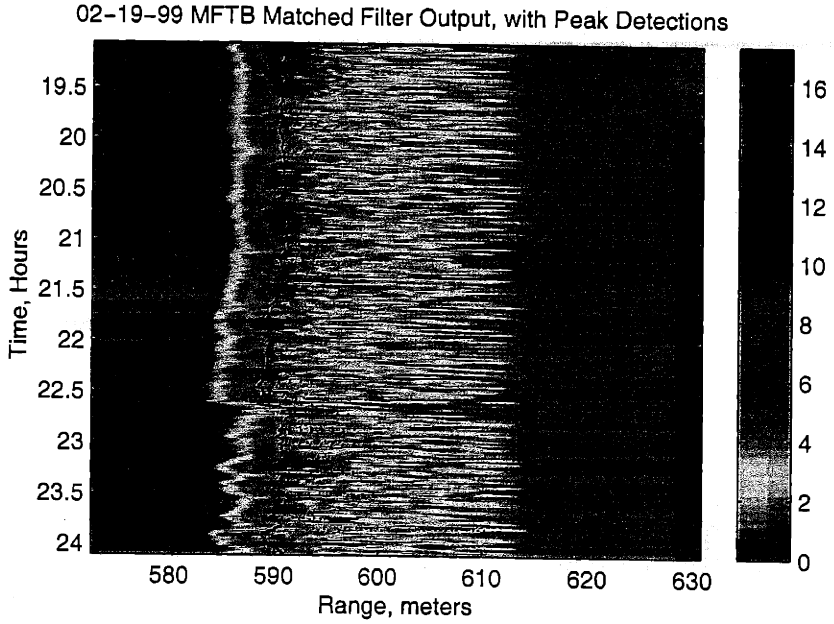

Figure A-5: 02-19-99 Mid Frequency Toneburst

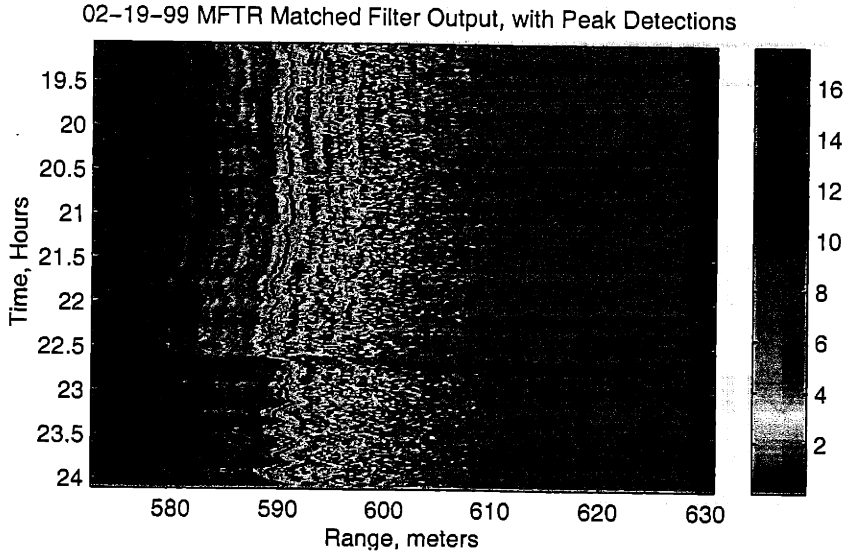

Figure A-6: 02-19-99 Mid Frequency Turyn Code

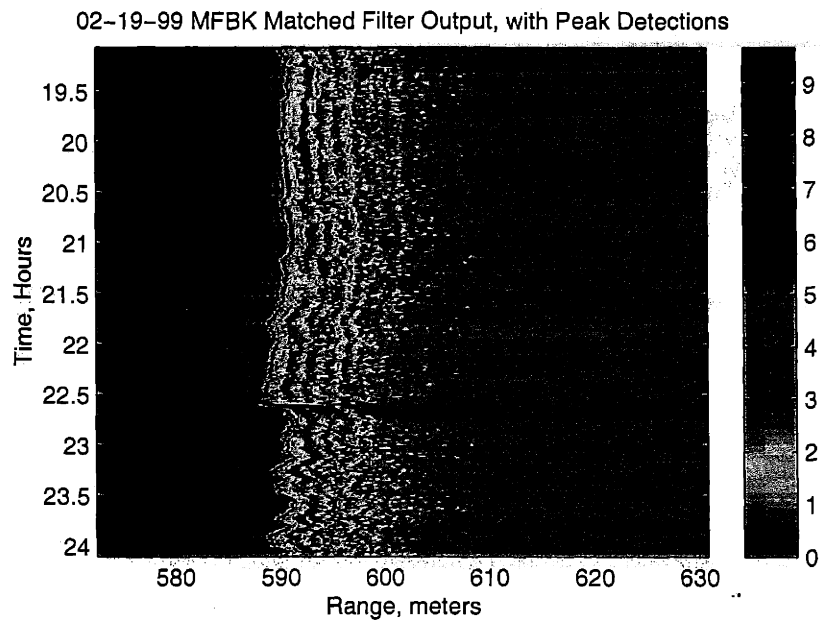

Figure A-7: 02-19-99 Mid Frequency Barker Code

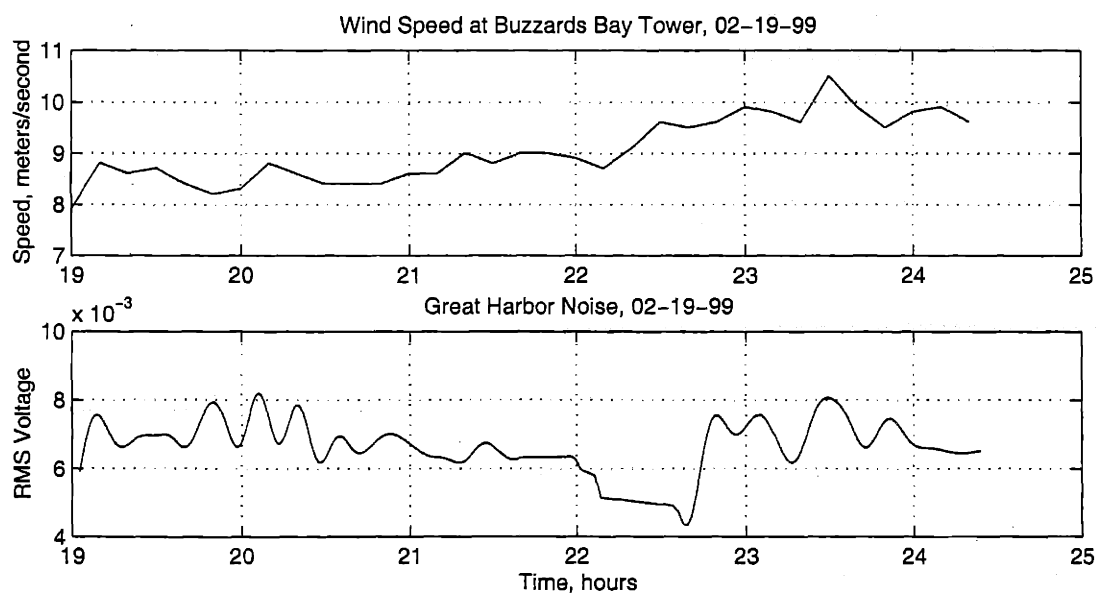

Figure A-8: 02-19-99 Wind Speed and RMS Noise Voltage

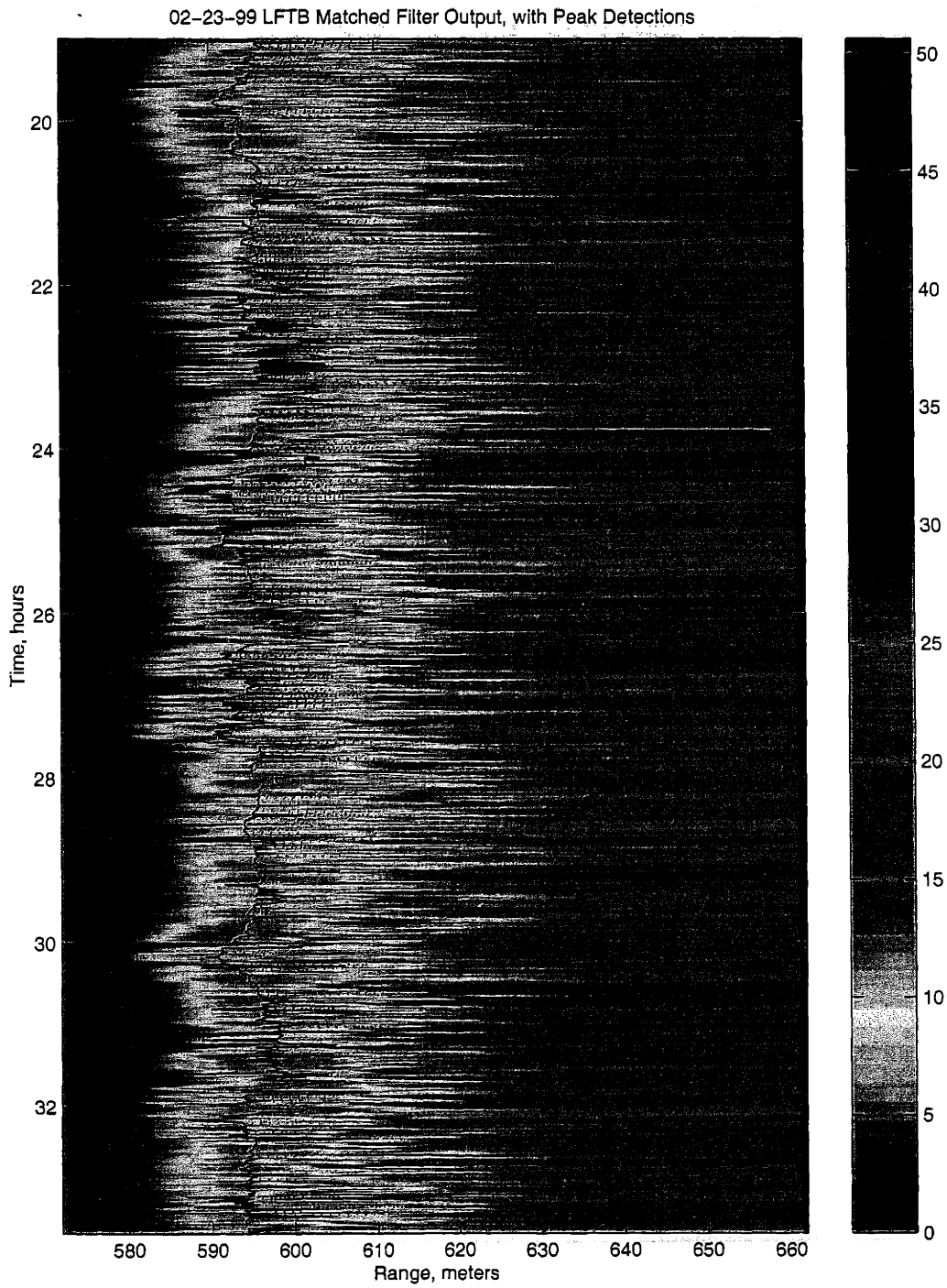

Figure A-9: 02-23-99 Low Frequency Toneburst

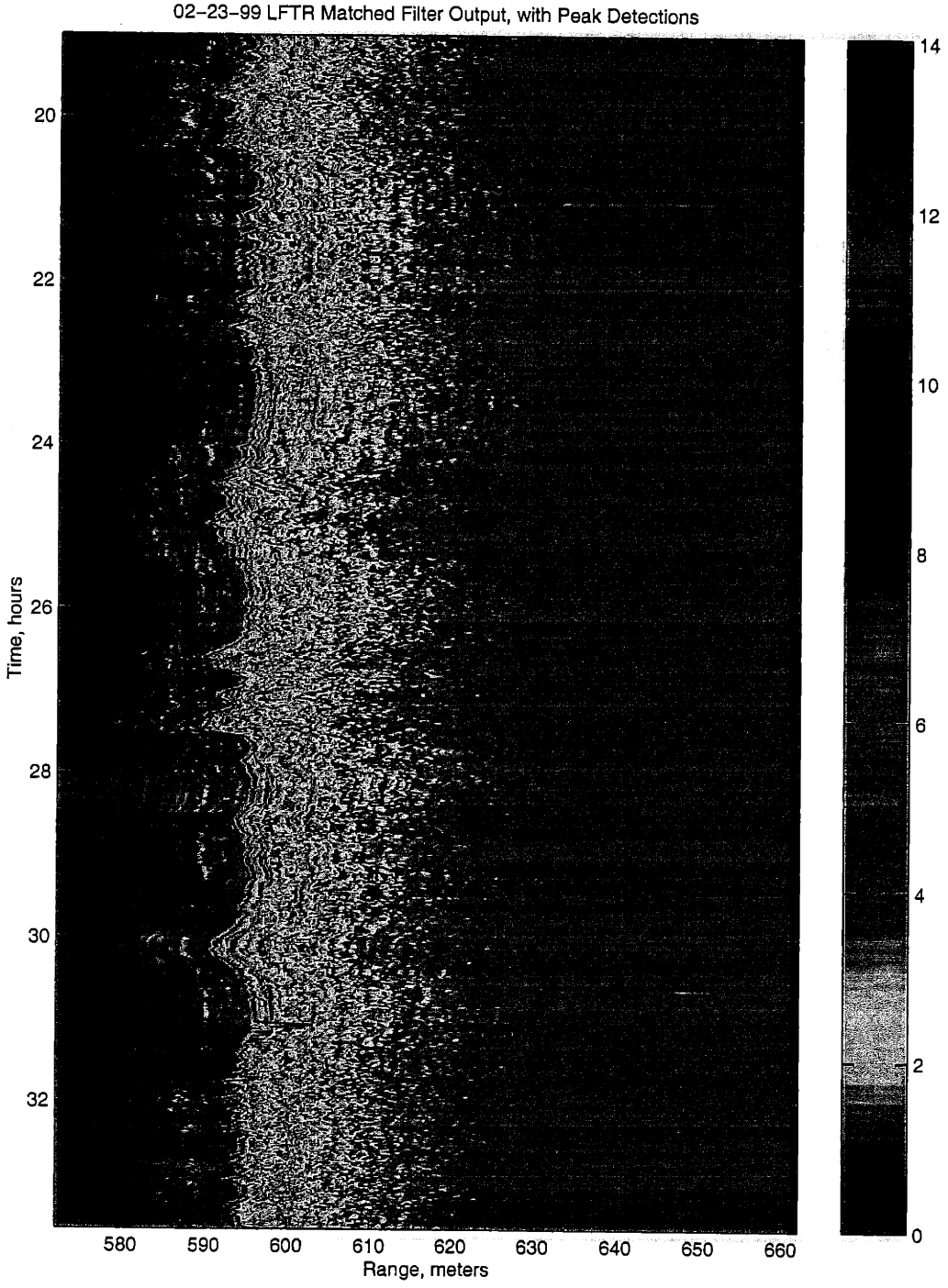

Figure A-10: 02-23-99 Low Frequency Turyn Code

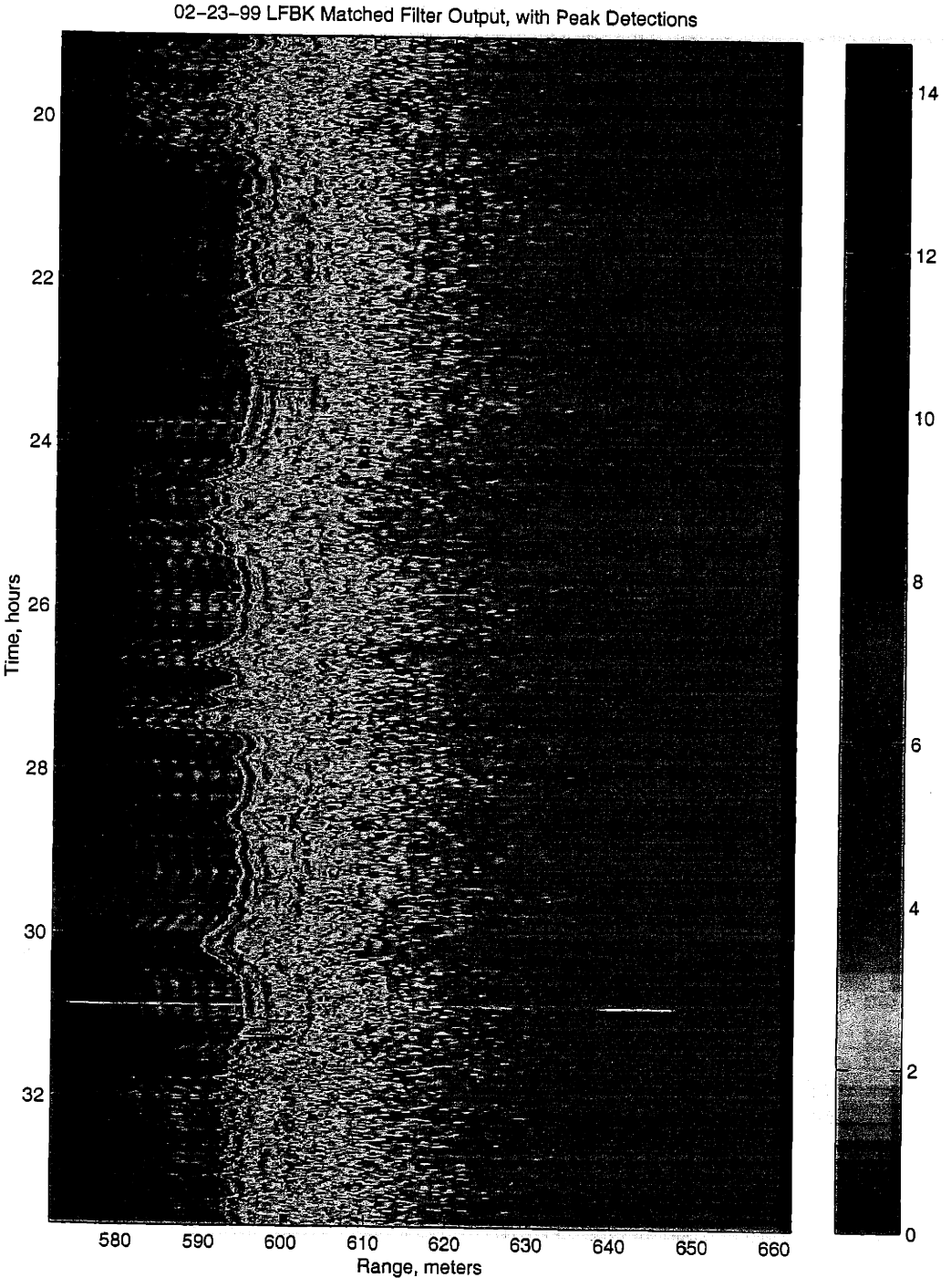

Figure A-11: 02-23-99 Low Frequency Barker Code

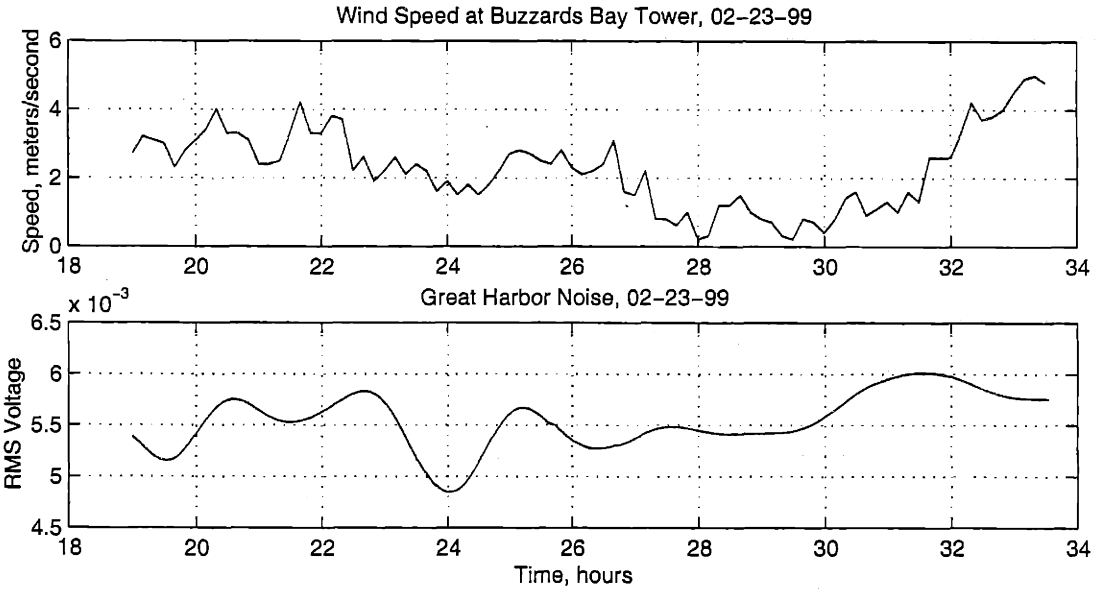

Figure A-12: 02-23-99 Wind Speed and RMS Noise Voltage

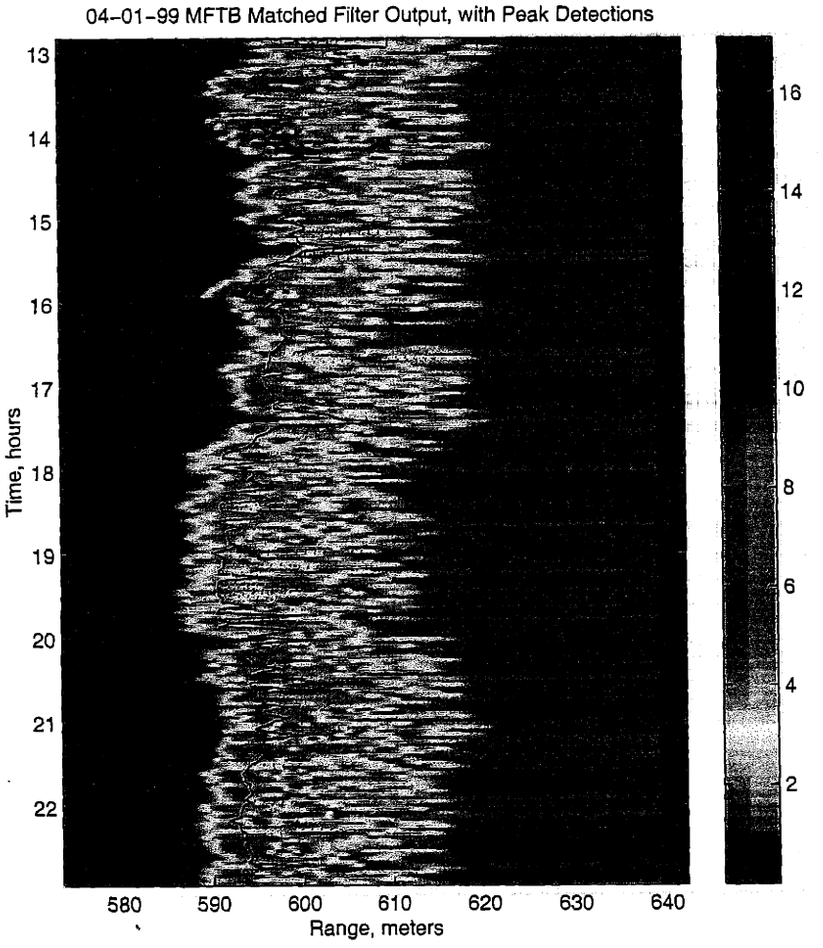

Figure A-13: 04-01-99 Mid Frequency Toneburst

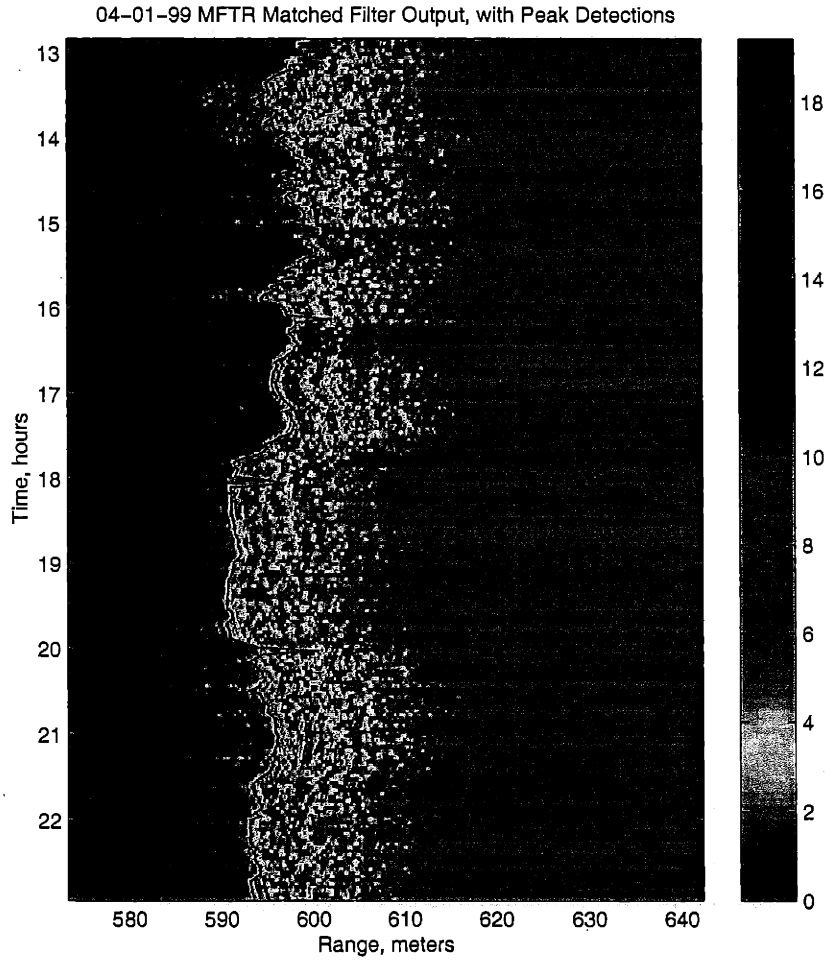

Figure A-14: 04-01-99 Mid Frequency Turyn Code

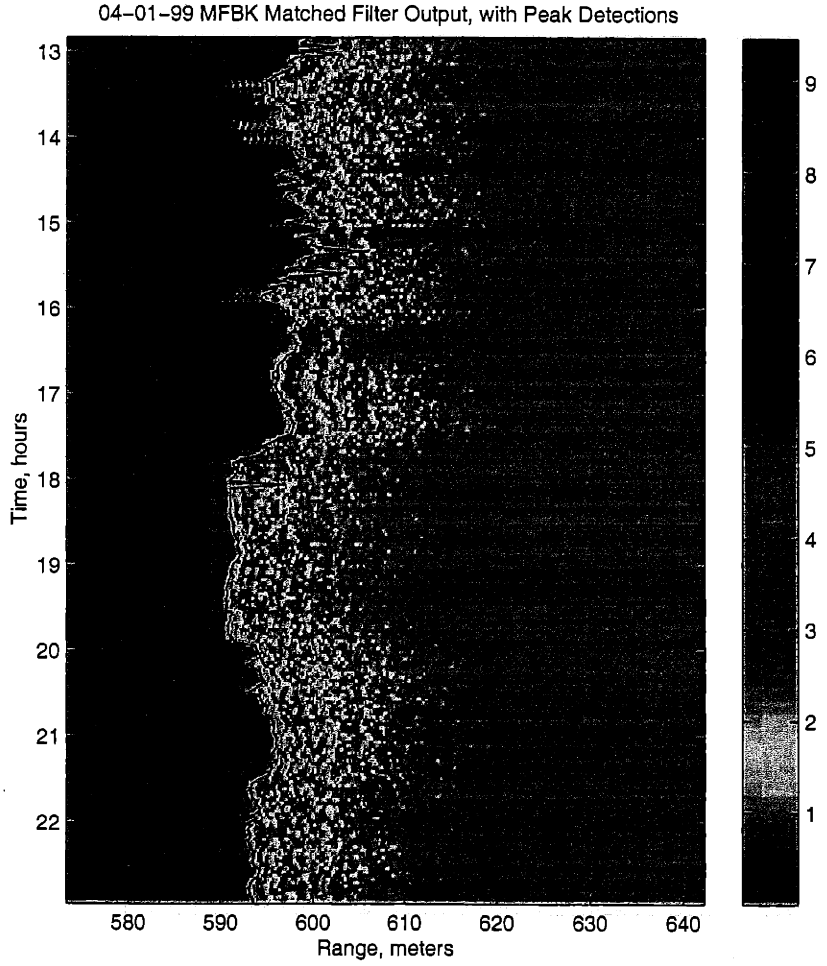

Figure A-15: 04-01-99 Mid Frequency Barker Code

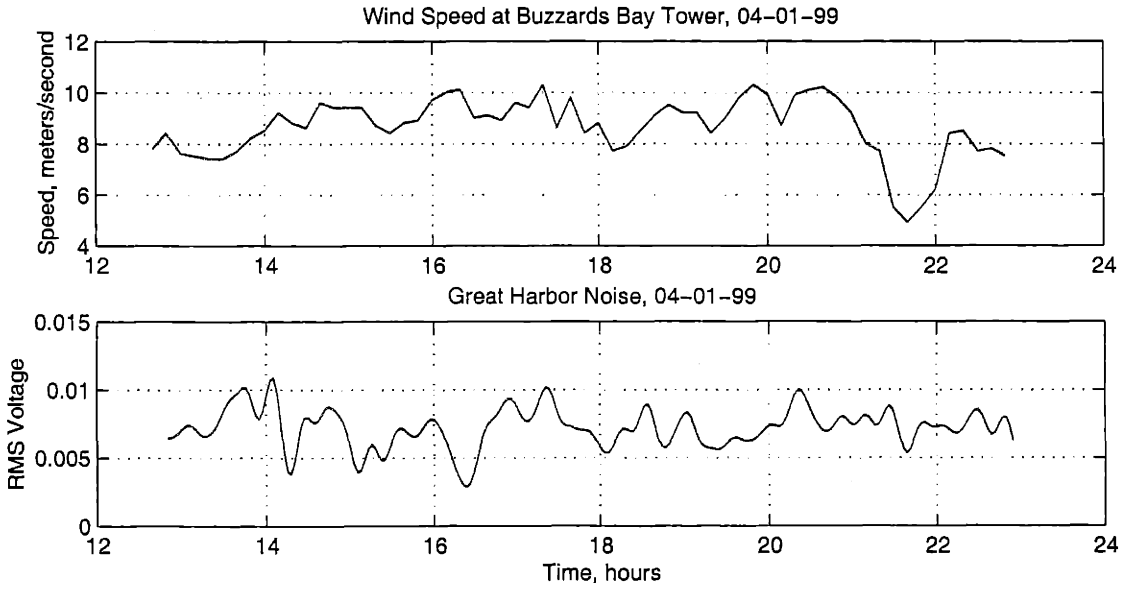

Figure A-16: 04-01-99 Wind Speed and RMS Noise Voltage

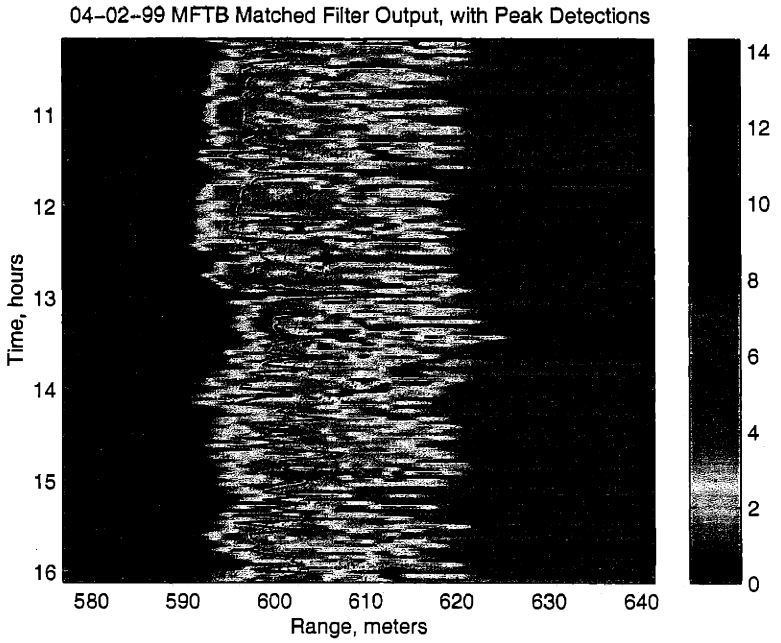

Figure A-17: 04-02-99 Mid Frequency Toneburst

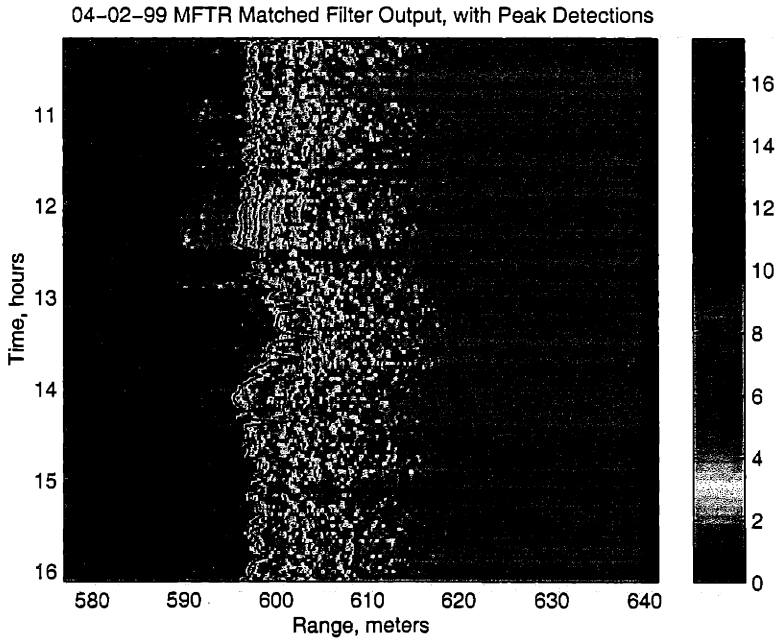

Figure A-18: 04-02-99 Mid Frequency Turyn Code

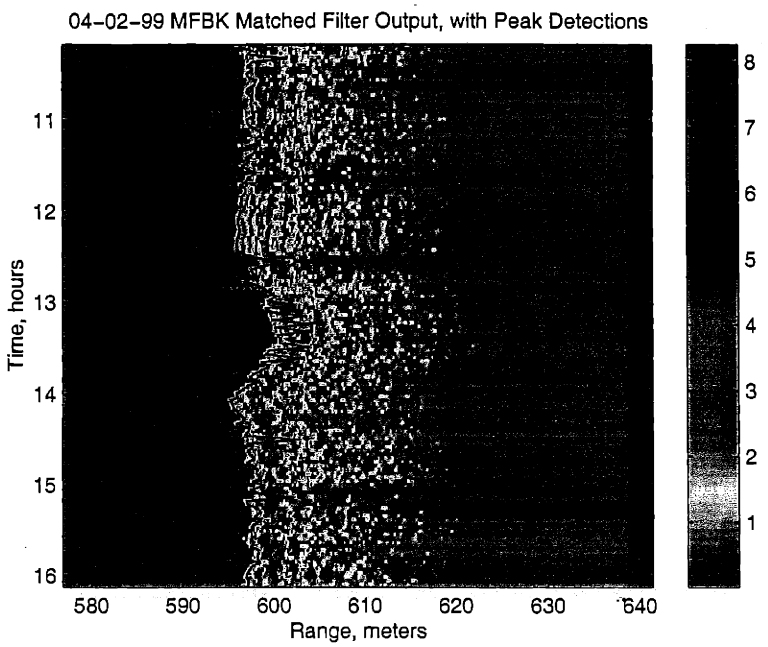

Figure A-19: 04-02-99 Mid Frequency Barker Code

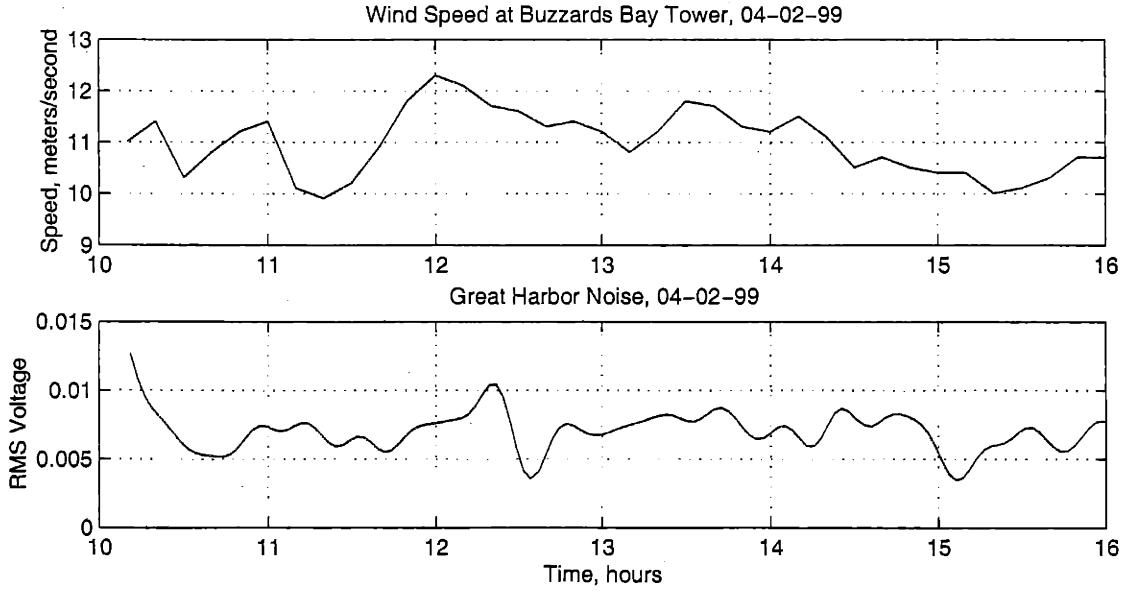

Figure A-20: 04-02-99 Wind Speed and RMS Noise Voltage

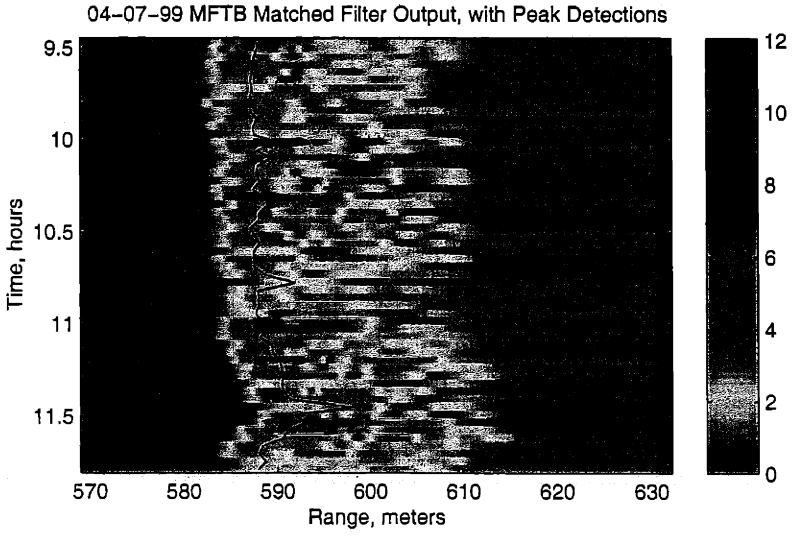

Figure A-21: 04-07-99 Mid Frequency Toneburst

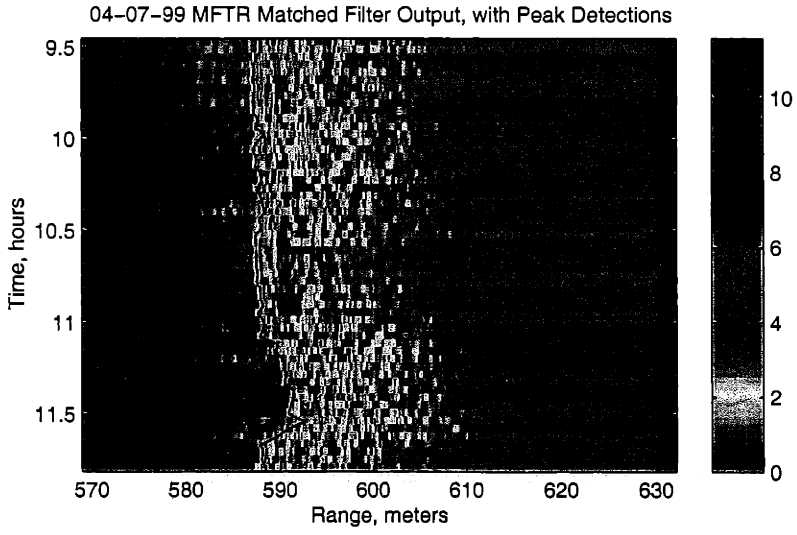

Figure A-22: 04-07-99 Mid Frequency Turyn Code

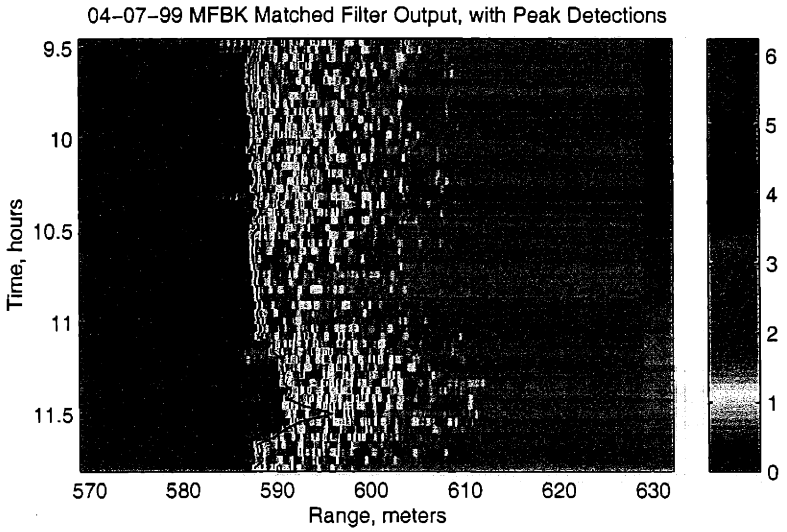

Figure A-23: 04-07-99 Mid Frequency Barker Code

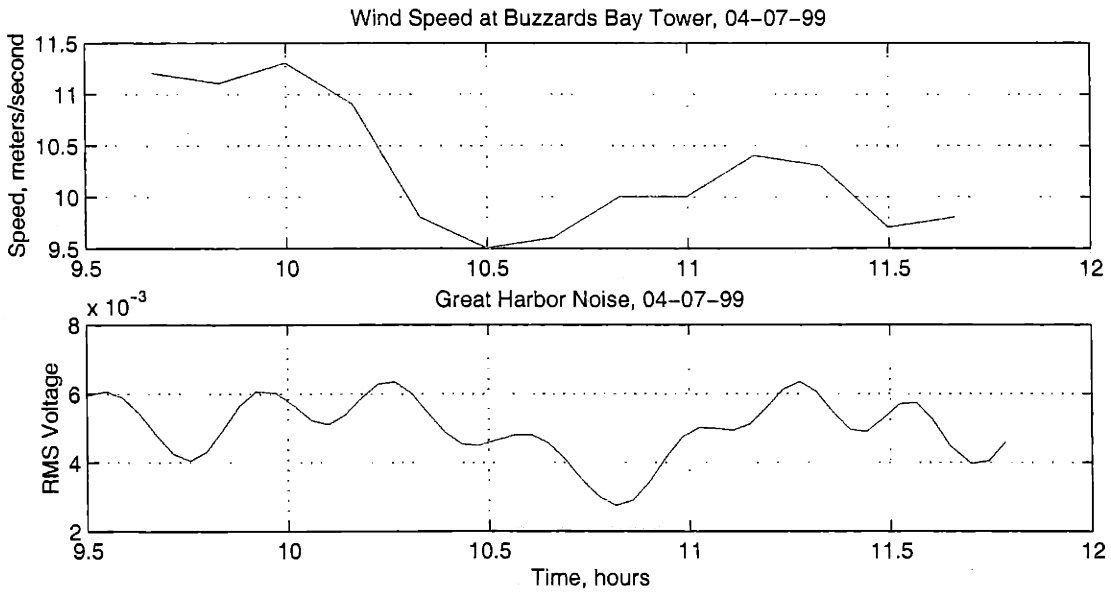

Figure A-24: 04-07-99 Wind Speed and RMS Noise Voltage

Appendix B

Epilogue: The Real Problem with REMUS at LEO-15

The research reported in this thesis was motivated by the problems with REMUS navigation observed at LEO-15 during July, 1998. While it was originally thought that the poor results obtained during these field operations could be the result of either a hardware glitch or the particular acoustic propagation environment, the focus of this work quickly narrowed to the acoustics once the obvious problem of the bandwidth limited transducers was eliminated, as described in Chapter 3. However, more recent work in this area has revealed that while the environment at LEO-15 is a particularly difficult one because of the multipaths supported, the primary source of the problem may in fact have been in the REMUS hardware after all.

During additional testing of the REMUS navigation systems in June 1999, it was discovered that too much gain was being taken in the initial preamplifier stages, resulting in clipping of moderately strong signals. This gain was subsequently reduced, and the performance immediately improved. The reason for the improvement was, however, initially misunderstood. Intuitively, one would expect that clipping would equalize the matched filter output of all arrivals, making it impossible to distinguish between them and resulting in arrival hopping. This is not, however, the case. As figure B-1 shows, even when high gain is taken and the results are then clipped at a low level, the relative heights of the matched filter output peaks are preserved. Given this, the range estimation should not have suffered from the clipping to the extent observed. However, clipping does result in the creation of spectral copies at multiples of the third harmonic of the original carrier. Without filtering,

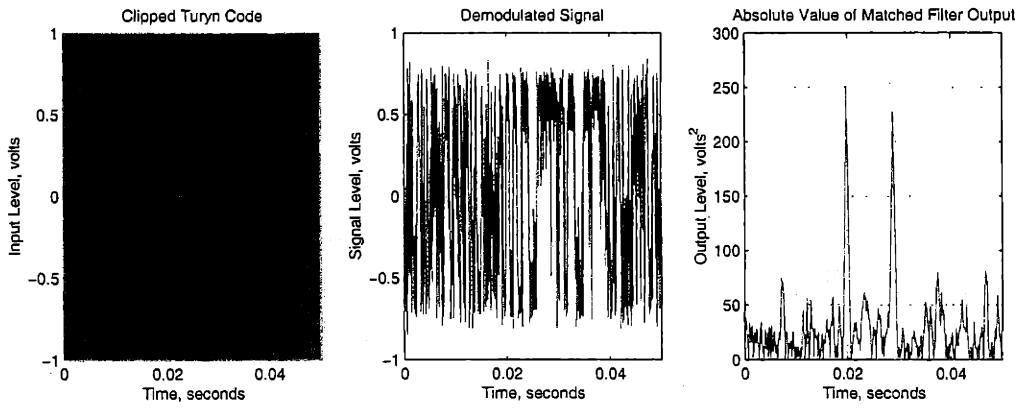

Figure B-1: Clipped Turyn Code, for the Same Conditions as Shown in Chapter 2

aliases of these under sampled spectra will be present in the input signal. It was the presence of this aliased interference, along with the effects of the waveguide, which were most likely the cause of the errors observed at LEO-15 in 1998.

Bibliography

[1] T.C. Austin. The application of spread spectrum signaling techniques to underwater acoustic navigation. In *Proceedings, IEEE Autonomous Underwater Vehicle Conference '94*, Cambridge, MA, 1994.

[2] R.H. Barker. Group synchronizing of binary digital systems. In W. Jackson, editor, *Communication Theory*, pages 273–287. Academic Press, New York, 1953.

[3] P. Beckmann and A. Spizzichino. *Scattering of Electromagnetic Waves from Rough Surfaces.* The Macmillan Company, New York, 1963.

[4] M. Deffenbaugh. *Optimal Ocean Acoustic Tomography and Navigation with Moving Sources.* PhD dissertation, MIT/WHOI Joint Program in Oceanography and Applied Ocean Science and Engineering, Applied Ocean Physics and Engineering, April 1997.

[5] G.V. Frisk. *Ocean and Seabed Acoustics.* P T R Prentice-Hall, Englewood Cliffs, NJ, 1994.

[6] A.M. Kerdock, R. Mayer, and D. Bass. Longest binary pulse compression codes with given peak sidelobe levels. *IEEE Transaction on Information Theory*, February 1986.

[7] D.B. Klfoyle and A.B. Baggeroer. Acoustic telemetry. To appear in IEEE Journal of Oceanic Engineering, 2000, 1999.

[8] V.O. Knudsen, R.S. Alford, and J.W. Emling. Underwater ambient noise. *Journal of Marine Research*, 7:410, 1948.

[9] L.N. Liebermann. Refelction of underwater sound from the sea surface. *Journal of the Acoustical Society of America*, 20:498, 1948.

[10] L.N. Liebermann. Reflection of sound from coastal bottoms. *Journal of the Acoustical Society of America*, 20:305, 1948.

[11] H. Medwin. Speed of sound in water for realistic parameters. *Journal of the Acoustical Society of America*, 58:1318, 1975.

[12] C.L. Piggott. Ambient sea noise at low frequencies in shallow water of the scotian shelf. *Journal of the Acoustical Society of America*, 36:2152, 1965.

[13] W.H. Thorp. Analytic description of the low frequency attenuation coefficient. *Journal of the Acoustical Society of America*, 42:270, 1967.

[14] G.L. Turin. An introduction to matched filters. *IRE Transactions on Information Theory*, pages 311–329, 1960.

[15] R. Turyn. Sequences with small correlation. In H.B. Mann, editor, *Error Correcting Codes*, pages 195–228. Academic Press, New York, 1968.

[16] R.J. Urick. *Principles of Underwater Sound.* Peninsula Publishing, Los Altos, CA, 1983.

[17] R.J. Urick and H.L. Saxton. Surface refelction of short supersonic pulses in the ocean. *Journal of the Acoustical Society of America*, 19:8, 1947.

[18] C.J. von Alt, B. Allen, T. Austin, and R. Stokey. Remote environmental measuring units. In *Proceedings, IEEE Autonomous Underwater Vehicle Conference '94*, Cambridge, MA, 1994.

[19] C.J. von Alt and J.F. Grassle. Leo-15: An unmanned long term environmental observatory. In *Proceedings, Oceans '94*, Newport, RI, 1994.

[20] G.M. Wenz. Acoustic ambient noise in the ocean: Spectra and sources. *Journal of the Acoustical Society of America*, 34:1936, 1962.

[21] A.D. Whalen. *Detection of Signals in Noise.* Academic Press, Inc., Orlando, FL, 1971.

[22] M.J. White. *Eldridge Tide and Pilot Book, 1999.* M.J. White, Boston, 1998.

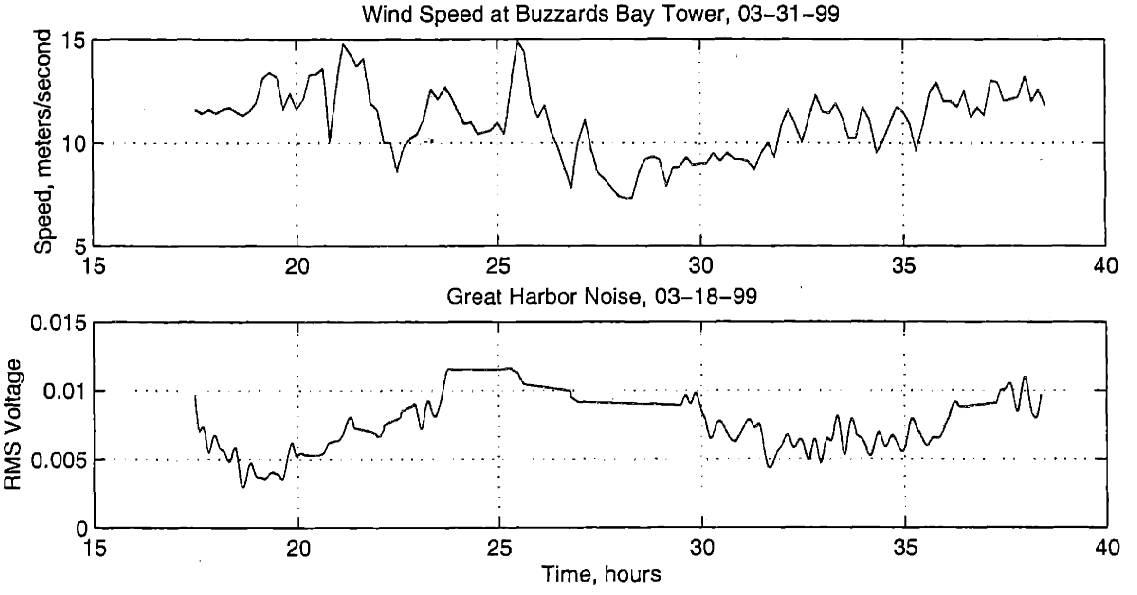

Figure 4-4: 03-18-99 Wind Speed and RMS Noise Voltage