NUWC-NPT Technical Report 11,968
5 March 2010

Separation of Sperm Whale Click-Trains for Multipath Rejection and Localization

Paul M. Baggenstoss
Sensors and Sonar Systems Department

NEWPORT

Naval Undersea Warfare Center Division
Newport, Rhode Island

PREFACE

This report was prepared under the DECAF Project ("Density Estimation for Cetaceans from Passive Acoustic Fixed Sensors"), sponsored by the U.S. National Marine Fisheries Service Office of Protected Resources.

The technical reviewer for this report was David J. Moretti (Code 71).

Reviewed and Approved: 5 March 2010

David W. Grande
Head, Sensors and Sonar Systems Department

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (*DD-MM-YYYY*)	2. REPORT TYPE	3. DATES COVERED (*From – To*)
05-03-2010		

4. TITLE AND SUBTITLE

Separation of Sperm Whale Click-Trains for Multipath Rejection and Localization

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Paul M. Baggenstoss

5.d PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

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

Naval Undersea Warfare Center Division
1176 Howell Street
Newport, RI 02841-1708

8. PERFORMING ORGANIZATION REPORT NUMBER

TR 11,968

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

U.S. National Marine Fisheries Service Office

10. SPONSORING/MONITOR'S ACRONYM

11. SPONSORING/MONITORING REPORT NUMBER

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this report, an algorithm is described for separation and grouping of sperm whale clicks received at a single hydrophone. The clicks are separated into click-trains from separate animals and separate propagation paths. The improvement possible in animal localization using multisensor data is then demonstrated. A localization algorithm is described that can handle multiple whales and that provides estimates of the Cramer-Rao lower bound for the position estimates.

15. SUBJECT TERMS

Sperm Whale Vocalization Marine Mammals Acoustic Detection and Localization

Signal Processing Underwater Sound

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Paul M. Baggenstoss
(U)	(U)	(U)	SAR	65	19b. TELEPHONE NUMBER (*Include area code*) (401) 832-8240

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont'd)

LIST OF TABLES

Table **Page**

SEPARATION OF SPERM WHALE CLICK-TRAINS FOR MULTIPATH REJECTION AND LOCALIZATION

1. INTRODUCTION

The localization of sperm whales (Physeter macrocephalus) is mainly accomplished through the measurement of the time-delay between individual clicks received at multiple sensors. Complicating the problem is the existence of multiple propagation paths ("multipath"), which causes false associations and false time-delay measurements. To reduce these false time-delay measurements, the elimination of multipath at individual sensors is proposed. This elimination will be accomplished by first arranging clicks into separate click-trains corresponding to individual whales and propagation paths, and then by classifying entire click-trains as either direct path or reflected path.

Previous work in click-train separation has employed various clues, including spectral and temporal features, inter-click correlation, and multisensor time delay (references 1 – 3). In an earlier publication (reference 4), the author describes a technique for separating the sperm whale clicks received at a single sensor into individual click-trains. The click-trains could be from single whales with multipath or from multiple whales or from both.

Assume that a series of N_c clicks has been received at one sensor. A click similarity metric $L_{i,j}$ is defined here as a measure of similarity between clicks i and j, $1 \le i \le N_c$, $1 \le j \le N_c$, $i \ne j$. It is assumed that $L_{i,j}$ has a higher value if clicks i and j are more similar. The goal is to arrange clicks into related groups or click-trains so as to minimize the sum of $L_{i,j}$ between sequential click-pairs in a group. The hypothesis is that, if the inter-click similarity is measured only between adjacent clicks within a click-train, then the total similarity (i.e., the sum of the inter-click similarity metric over all adjacent click-pairs) is maximized when the clicks are properly grouped. There must also be a penalty for creating a new click-train (otherwise, the metric would be minimized at zero by assigning each individual click to a click-train of one click). Thus, a fixed penalty value is added for each group.

This problem setup is identical to that posed in reference 4—although, in that paper, the goal was the minimization of total error as opposed to the maximization of similarity. The algorithm reported in reference 4 has two shortcomings. The first shortcoming relates to the method that the algorithm uses to search for the maximum total similarity. The algorithm was inspired by the Viterbi algorithm, which is a type of dynamic programming (reference 5). It performed exhaustive searches over limited time windows, moving forward in time, and repeating the exhaustive search at regular time intervals. In environments with three or more whales, the exhaustive searches, even if limited in time, required too much computer time to complete. So, the number of click-pairs that could be considered at one time had to be limited, reducing the algorithm's performance. In the present report, this issue is addressed through the use of a "break-and-reassemble" approach that does not suffer the exponentially increasing computation time of the exhaustive searches, yet has comparable performance.

The second shortcoming of the reference 4 algorithm is that it does not consider click-train correlations. In other words, once the click-trains are formed, the algorithm classifies the click-trains as direct path or reverberation based only on individual click characteristics. A better approach would try to establish whether two click-trains from the same sensor are related by a time delay, and, if they are, then the click-train that comes later in time is likely to be reverberation.

In this report, these issues and related enhancements to the original algorithm are discussed. Further, the algorithm is tested as part of a localization system, and actual localization results are shown using data from the Second International Workshop on Detection, Classification, and Localization of Marine Mammals (Monaco 2005 Workshop) (reference 6), as well as 2007 data from the Atlantic Undersea Test and Evaluation Center (AUTEC) range. Section 2 discusses the process of separating the clicks into some number of click-trains. Section 3 describes the elimination of those click-trains due to multipath. In section 4, the click-trains are used to estimate the position of each whale. In section 5, the application of these techniques to real data is examined.

2. CLICK-TRAIN SEPARATION

2.1 CLICK DETECTION

Clicks are detected on a single sensor within a 12-second data window. A wider time window is not desired because the ultimate goal is position estimation. The animal's position can be regarded as constant in a 12-second window. To detect clicks, a 0.012-second moving average of instantaneous power was used. A time quantization of 0.004 second was achieved by shifting the moving average 0.004 second at each update. The moving average can be normalized by any standard local background estimator that gives an approximately unbiased estimate of local ambient background noise power, such as a median filter. An approximate unbiased background power estimator is needed to prevent changing the effective signal-to-noise ratio (SNR) threshold as a function of click density, since nearby clicks could bias the estimate of local noise power. Before detection, the input data were bandpass filtered to the range of 1800 to 24,000 Hz. Information below 1800 Hz tended to be affected by frequency dispersion effects, so its elimination made it easier to determine click arrival times because the clicks were sharper.

Clicks are detected by first finding a local maximum above the threshold. The detection start time is found by searching backward in time from the peak and finding the point where the instantaneous power falls below one-quarter of the peak value or the point where it stops decreasing. The detection end time is found by searching forward in time from the peak until the power either: (1) exceeds the peak value, or (2) rises by more than 20% in one time step, or (3) reaches a maximum at about 64 msec from event peak location. The detection scheme seeks to declare only one event in the case of two closely separated peaks, such as may occur from the inter-click interval, or two events in the case of two very close clicks. It also allows for slight temporary increases in power on the tail end, such as may occur in reverberation. Since the start and end time of the click is determined "on the fly," the clicks have a variable duration. A short time-series was extracted from the input data stream for each click for later use in click classification and formation of click-trains (separation).

2.2 CLICK PERIODICITY ANALYSIS

A database of manually validated (validation methods are described in sections 5.2.1 through 5.2.3) click-trains totaling some 68,000 clicks was gathered. A histogram of the time delays between adjacent clicks is shown in figure 1. The distribution has a median of 0.752 second. There is a narrow central lobe with outliers. It is possible that many of the outliers are caused by errors in the validation process—for example, by occasional missing clicks. By fitting a two-component Gaussian (kernel) mixture model (GMM) to the distribution, the central lobe and outliers can be separately modeled. The main lobe is fitted with a GMM having a mean of 0.7525 and a standard deviation of 0.1276. The outliers are modeled with a mean of 1.11 and a standard deviation of 0.4648. The main lobe component is a Gaussian distribution estimate of the instantaneous click period for sperm whales.

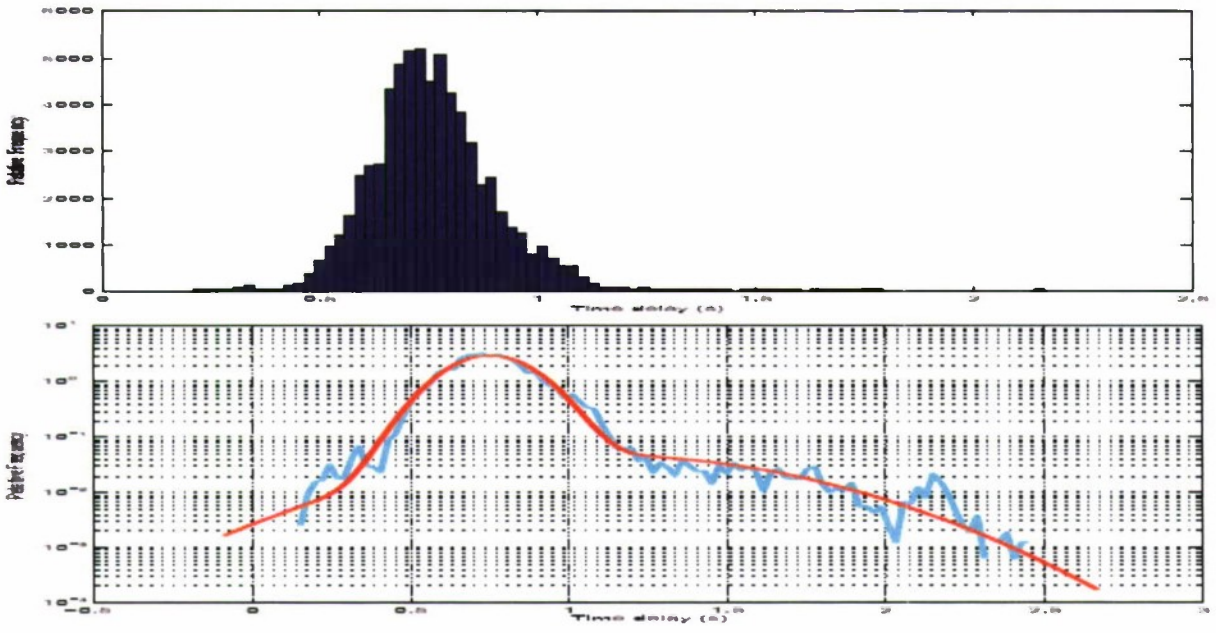

Figure 1. Statistical Analysis of Click-to-Click Time Interval (Period)
(Top: histogram of 68,000 validated click pairs. Bottom: two-component GMM fit. Red is GMM, light-blue is the histogram.)

It is also interesting to analyze the consistency of the click-to-click time delays. A good measure of this consistency is to analyze each available click triple. Given consecutive clicks at times t_i, t_j, t_k, a period consistency measure is

$$c = \log\left(\frac{t_k - t_j}{t_j - t_i}\right). \tag{1}$$

A histogram of c is shown in figure 2. The distribution is centered at 0. As with the time delays, there is a narrow central lobe with outliers. Again, it is possible that many of the outliers are caused by errors in the validation process. Fitting a two-component GMM to the distribution allows the central lobe and outliers to be separately modeled. The central lobe has a standard deviation of 0.0526. The outliers have a standard deviation of 0.1534. The fit is illustrated in figure 2. The main lobe component is a Gaussian distribution estimate of the click period consistency for sperm whales and will be useful later in calculating a measure of goodness-of-fit (see section 2.3.3).

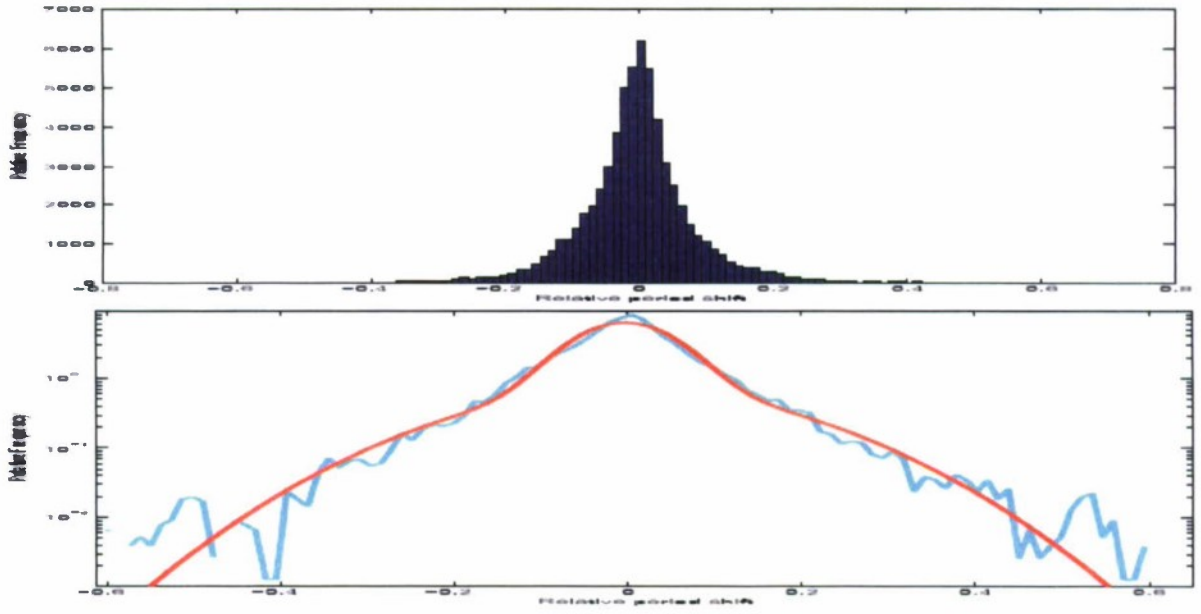

Figure 2. Statistical Analysis of Period Consistency
(Top: histogram of 68,000 validated click pairs. Bottom: two-component GMM fit.)

2.3 CLICK SEPARATION ALGORITHM

2.3.1 Identification of Candidate Click-Pairs

Let there be N_c detected clicks at one sensor within a specified time window of T seconds. The approximately $N_c \times N_c$ potential pairings are examined in order to determine a set of N_p potential click-pairs. Let the time difference be defined as

$$\tau_{i,j} = t_j - t_i .$$

In practice, it is not necessary to examine all (approximately) $N_c \times N_c$ potential pairings because the search can be limited to only click pairings where $t_i < t_j$, and where the time difference falls within a practical range of values:

$$\tau_{\min} < \tau_{i,j} - \tau_{\max} .$$

For sperm whales, typically,

$$\tau_{\min} \cong 0.4 \text{ second}, \quad \tau_{\max} \cong 1.2 \text{ second}.$$

It is also possible to arbitrarily limit the SNR so that the ratio between the two peak SNRs must fall within a prescribed range.

The left and right indexes of the click-pair candidates are given by

$$(i_p, j_p), \quad 1 \le p \le N_p .$$

Thus, candidate click pairing p associates the clicks at time t_{i_p} and time t_{j_p}, and time difference

$$\tau_p = t_{j_p} - t_{i_p} .$$

2.3.2 Calculation of Click Similarity Metric

Let H_1 be the hypothesis that clicks i and j are related (two sequential clicks in a click-train from one whale). Let H_0 be the hypothesis that clicks i and j are unrelated, or are selected at random. There is a subtle difference between H_1 and H_0 because, by chance, random selection could result in two related clicks being selected. To be precise, what is actually meant is the latter definition, *selected at random*. Let $L_{i,j}$ be a statistical measure of how similar clicks i and j are defined:

$$L_{i,j} = \log\left\{\frac{p(H_1|\mathbf{z}_{i,j})}{p(H_0|\mathbf{z}_{i,j})}\right\},$$

where $\mathbf{z}_{i,j}$ are features extracted jointly from the time-series data of clicks i and j. Using Bayes' rule,

$$L_{i,j} = \log\left\{\frac{p(H_1|\mathbf{z}_{i,j})}{p(H_0|\mathbf{z}_{i,j})}\right\},$$

$$L_{i,j} = \log\left\{\frac{p(H_1)p(\mathbf{z}_{i,j}|H_1)}{p(H_0)p(\mathbf{z}_{i,j}|H_0)}\right\}, \tag{2}$$

where $p(H_0)$ and $p(H_1)$ are the *a priori* probabilities of H_0 and H_1:

$$p(H_1) = 1 - p(H_0).$$

Probability $p(\mathrm{H}_1)$ is just the probability that two clicks are related if they are selected at random. A value of 0.03 was used for $p(\mathrm{H}_1)$. The likelihood functions $p(\mathbf{z}_{i,j}|H_0)$ and

$p(\mathbf{z}_{i,j}|H_1)$ (actually, estimates of the true likelihood functions) are represented as a GMM and estimated using standard estimation methods (reference 7). Click-pairs for hypothesis H_1 were collected from the 68,000 validated clicks (resulting in slightly less than 68,000 click-pairs), and their feature information was stored. Likewise, 48,000 randomly paired clicks were collected for H_0.

The eight features include information that compares the clicks, such as spectral comparison and correlation coefficient (the normalized peak of the time-domain correlation of the two time-series), as well as information about individual clicks *i* and *j*. The click-pair features are listed in table 1. In addition to the 8 click-pair features, the difference of 10 single-click features (see table 3) from the left-most click (first in time) was used. Also, the difference of left and right click single-click features was appended. From this total of 28 features, the following 17 features were selected and separated into 3 groups:

[kerr, perr, aerr, K3, D-K3],

[cerr, psdot, derr, K1, K2, K4, TSTD, TM4, D-K1, D-K4, D-TM3],

[dT],

where "D-" stands for feature differencing. Feature selection was accomplished by maximizing classification performance while training on half of the data and testing on the other half. The experiment was then reversed by exchanging the two halves. The results were then combined.

Table 1. List of Click-Pair Features Used for Click Association

Name	Description
kerr	Mean square reflection coefficient deviation
cerr	Coherent correlation measure.
perr	Log power ratio
aerr	Log amplitude (SNR) ratio
psdot	Power spectrum deviation
snrmin	SNR off weakest click
derr	Incoherent correlation measure
dT	Detection time difference

Each of the two groups was independently modeled as a GMM, and the log-likelihood values of the three group likelihood functions were added. This approach reduced the effects of dimensionality on the estimation of the probability density function (PDF) models. Note that the detection time difference (dT) stands alone and is modeled by a fixed Gaussian distribution with a mean of 0.7525 and a standard deviation of 0.1276 (see section 2.2).

A histogram of the likelihood ratio (equations (2)) is shown in figure 3. The classification performance on training data is shown in table 2.

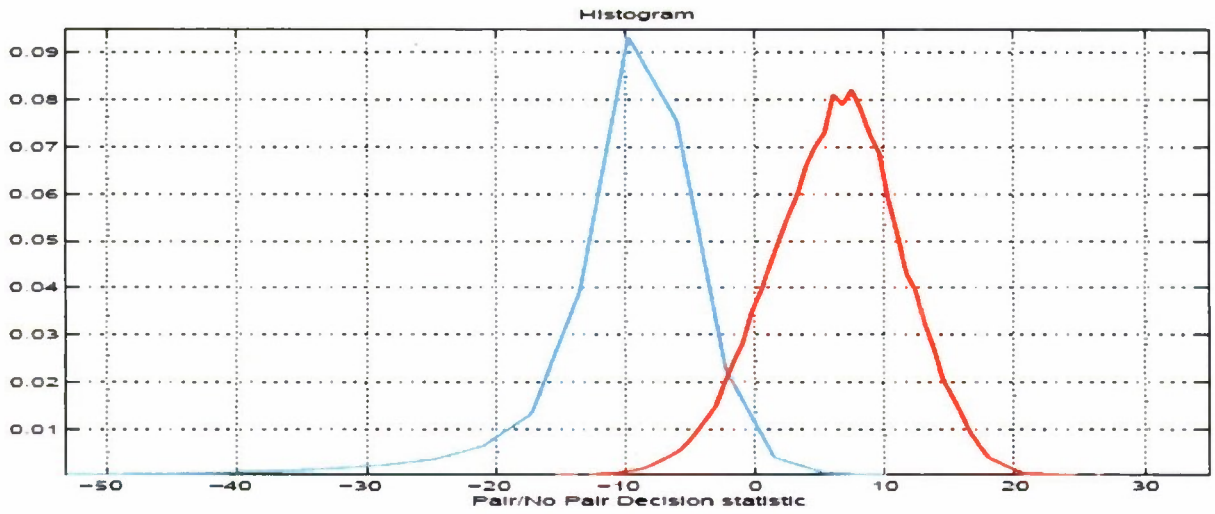

Figure 3. Histogram of Likelihood Ratio (Equation (2)) for Training Data for Class H_0 (Cyan) and H_1 (Red)

Table 2. Confusion Matrix for Single-Click Classification

	Events Classified As:	
	H_0	H_1
H_0 (unrelated)	45,918 (96%)	2082 (4%)
H_1 (related)	3770 (6%)	64,614 (94%)

2.3.3 Optimization Criterion

Solution of the click separation problem is tantamount to finding the best valid subset of the N_p click-pairs. Let S be a valid subset of the N_p candidate click-pairs. To be valid, a given click must appear no more than once as a left-click in a pairing, and no more than once as a right-click in a pairing. Associated with each valid subset is a total similarity value:

$$L_{\text{tot}}(S) = P(S) + H(S) + \sum_{p \in S} L_{i_p, j_p}.$$

The term $P(S)$ is the total lone click penalty. In simple terms, $P(S)$ is the number of clicks not appearing in subset S (unused clicks) multiplied by the specified constant lone penalty value P_{lone}. Typically, P_{lone} is set to a value less than the threshold appropriate for a two-class classifier between H_0 and H_1 (about –2 in figure 3). Increasing P_{lone} makes it more attractive to leave

clicks out of possible pairings, while decreasing P_{lone} makes it more attractive to include clicks in possible pairings. P_{lone} is set manually to maximize qualitative performance.

$H(S)$ is a periodicity measure based on c, the period consistency measure (equation (1)). $H(S)$ increases if the clicks in S exhibit consistent periodicity. This term is calculated from each set of adjacent click-triples. Let (i, j, k) be a click-triple contained in a click-train in the groupings S. Then,

$$h(i,j,k) = -\frac{1}{2}\log(2\pi\sigma_p^2) - \frac{c^2}{2\sigma_p^2},$$

where c is the period consistency measure (equation (1)) and σ_p^2 is the periodicity variance ($\sigma_p^2 = 0.0526^2$) (see section 2.2). The total term $H(S)$ equals the sum over all the triples contained in grouping S:

$$H(S) = \sum_{S \in S} \sum_{i,j,k \in S} h(i,j,k).$$

The click separation approach comes down to the maximization

$$\max_S L_{\text{tot}}(S).$$

2.3.4 Algorithm Initialization

The algorithm is initialized by choosing the null set

$$S = \{\},$$

thus assuming that all candidate pairings are false. Each click is therefore regarded as a click-train of one click. The value of L_{tot} for the null set is simply $N_c\, P_{\text{lone}}$.

2.3.5 Algorithm Update

Let S^* be the current best solution. The algorithm is best described as "break and re-assemble" in that it chooses an arbitrary time t and removes all click pairs in S^* that straddle time t. Thus, it "breaks" the solution at time t. Let this broken subset be called S'. The algorithm then enumerates all the valid subsets of candidate click-pairs that straddle time t and can be added to S to repair the break. Let $\mathcal{R}$ be the set of valid "repair" subsets or, in other words, the set of subsets S such that the union $S' + S$ is a valid subset. In mathematical notation,

$$\max_{S \in \mathcal{R}} L_{\text{tot}}(S'+S).$$

Set $\mathcal{R}$ is very small compared to the total number of valid click-pair subsets, so exhaustive search is very fast. The solution is broken and repaired at arbitrary times t that typically fall between clicks. Time t was selected on a uniform grid by incrementing by τ_{min}. Note that during initialization when $S = \{\ \}$ there is nothing to "break," so the algorithm proceeds by adding the best subsets of click-pairs that straddle time t.

2.3.6 Click-Train Merging

Often the algorithm produces two overlapping click-trains that use alternating clicks from a given whale. To prevent this, click-trains are merged when possible and a test is run for any increase in the total similarity measure. Merging was done by forming one click-train from the union of the clicks in two click-trains. This technique has led to a significant performance improvement.

2.3.7 Algorithm Convergence

The algorithm declares convergence when L_{tot} stops increasing after repeated attempts.

3. MULTIPATH ELIMINATION

3.1 INDIVIDUAL CLICK CLASSIFICATION

Before click-trains can be classified as reverberation, one needs to know how to classify individual clicks. To assist in the rejection of multipath, individual clicks are classified as direct path, surface path, or bottom-bounce reverberation. There is no inherent benefit in distinguishing the surface path from the bottom reverberation path, but the characteristics of the two reflection types are very different. It has been found that it is best to separate them for purposes of PDF estimation. The PDF models of the two classes are typically better than a single PDF model of the joined classes.

A set of single-click features is extracted from the time-series of each click (that is, the input data samples in a time window encompassing the click). These features are listed in table 3. GMMs are used for PDF modeling. Single-click training data were manually collected for 37,000 direct-path clicks; 26,000 surface-path clicks; and 4200 reverberation (bottom-bounce) clicks. Table 4 shows the results of a classification experiment using the single-click training data. Individual clicks were classified using a classical Neyman-Pearson classifier or so-called Bayes classifier for equal class prior probabilities. Thus, the likelihood function is calculated for each of the three hypotheses and the likelihood values are compared. A feature scatter diagram is shown in figure 4 for a typical feature pair..

Table 3. List of Single-Click Features Used for Path Identification

Name	Description
SNR	Log of signal-to-noise ratio
K1	First autoregressive reflection coefficient
K2	Second autoregressive reflection coefficient
K3	Third autoregressive reflection coefficient
K4	Fourth autoregressive reflection coefficient
TSTD	Second moment of signal energy about center of mass
TM3	Third moment of signal energy about center of mass
TM4	Fourth moment of signal energy about center of mass
S1 – 4	Amplitude attack rate
LEN	Click detection length

Table 4. Confusion Matrix for Single-Click Classification

	Events Classified As:			Percent Classified As:		
	Direct	**Surface**	**Reverberation**	**Direct**	**Surface**	**Reverberation**
Direct Path	30,545	6467	877	81	17	2
Surface Path	604	25,396	232	2	97	1
Reverberation Path	52	368	3843	1	9	90

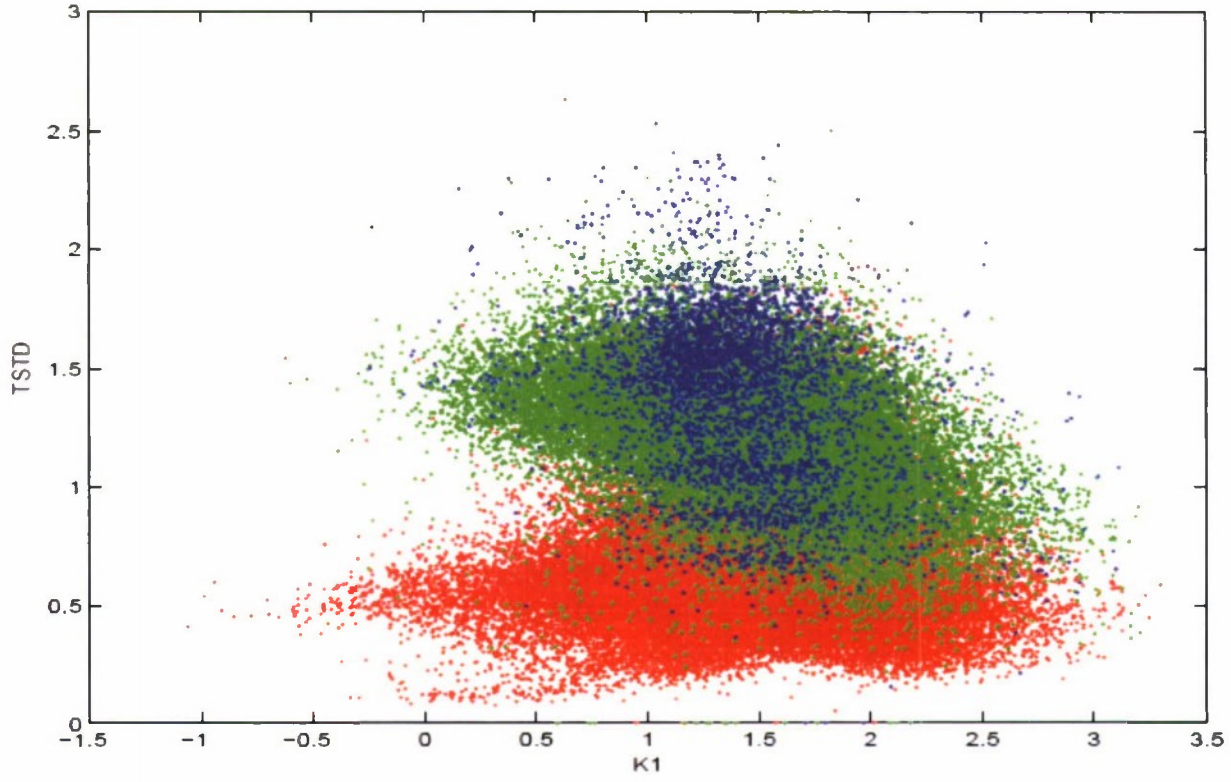

Figure 4. Scatter Diagram for Features TSDT and K1 for Direct Path (Red), Surface Path (Green), and Reverberation Path (Blue) (While the GMM model operates in many dimensions, viewing the data in two dimensions adds insight into the separability of the classes in the feature space.)

3.2 CLICK-TRAIN CORRELATION

To assist in classifying a click-train as direct path or multipath, the click-trains found at a given sensor are cross-correlated. The hypothesis is that two click-trains that are different propagation paths of the same original click-train will have a high correlation coefficient. If they correlate, the later click-train will be rejected as multipath.

Two click-trains are correlated using the frequency-domain approach. The artificial frequency-domain representation of each click-train is created by assuming that the clicks are impulses (Dirac delta functions) of equal amplitude. The sampling rate is chosen to provide a time-domain resolution of 20 msec. The two click-trains are cross-correlated by calculating the inverse fast Fourier transform (FFT) of the product of the frequency-domain representations of the two click-trains. To produce smooth correlograms, the frequency-domain is limited (by Hanning shading centered at zero frequency) to one-quarter of the sampling rate, effectively providing an 80-msec resolution. Finer resolution is not desired as there is error in the click arrival times, as well as error due to animal motion during the time window (normally 12 seconds). Correlation peaks are searched in a range of time depending on the bottom depth, which influences the maximum difference in propagation time between paths (±1.4 seconds was

used here). An example of a correlogram is shown in figure 5. Superimposed on the correlogram is the computed maximum correlation value, which depends on the delay value. The maximum correlation value equals the number of potential overlapping clicks (upper red line). In the figure, the maximum correlation value varies between 2 and 11, depending on the number of clicks that potentially overlap. The correlation value must be compared with this maximum value when determining if the correlation value is adequate. The actual unnormalized correlation at –0.5 second is about 10.5, which is very good compared to the maximum value of 11. Thus, it indicates that the two click-trains are indeed multipath related and the later click-train is multipath. In the figure, the side lobe peaks are due to the 1.1-second period of the click-trains.

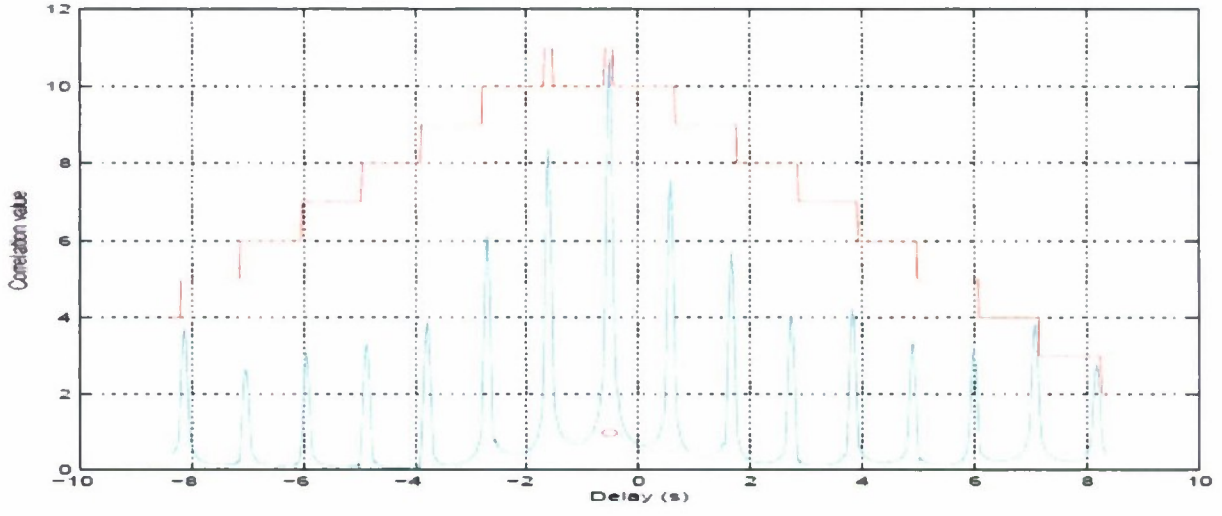

Figure 5. Example of Inter-Click Correlation

(The light blue line can be interpreted as the number of overlapping clicks (Y-axis) when the two click-trains are superimposed at a particular delay (X-axis). The red line is the upper limit of correlation (the number of potentially overlapping clicks). The actual unnormalized correlation at –0.5 second is about 10.5, which is very good compared to the maximum value of 11. The upper limit indicates that there are 11 available time-overlapping clicks.)

3.3 MULTIPATH ELIMINATION RULES

Multipath click-trains are eliminated if the individual clicks within the click-train are overwhelmingly classified as surface or reverberation using single-click classification. This can be determined by calculating the median single-click likelihood value in the click-train. Click-trains with questionable classification results are not eliminated based on single-click classification. Those click-trains that remain are processed by click-train correlation (section 3.2). If the correlation value exceeds a threshold of about 0.90 (as compared with the maximum potential value as described in section 3.2), the later click-train is eliminated. It was determined that at least four clicks need to be involved in the correlation; otherwise, the result is not meaningful.

4. LOCALIZATION

4.1 LOCALIZATION APPROACH

The problem of localizing sperm whales using click-trains received at multiple sensors works primarily by measuring time delays between the clicks received at multiple sensors and comparing them with a propagation model. Previous approaches (references 6, 8 – 10) have estimated time delays using clicks received at multiple sensors without first forming click-trains or removing multipath.

4.2 INTER-SENSOR TIME-DELAY ESTIMATION APPROACH

For localization, the click-train correlation described in section 3.2 is applied to correlate click-trains received at pairs of sensors. Time delays are subjected to a minimum correlation value of 0.9 and a minimum number of clicks (four). This results in a set of time delays from available sensor pairs. Each time delay carries a weight corresponding to the number of clicks involved in the correlation (the unnormalized correlation value in section 3.2).

Besides the ability to reject invalid time delays, correlation can be used to form inter-associated groups. All of the available time-delay measurements can be organized into inter-related groups based on correlation. When initializing the localization solutions, this correlation significantly limits the number of potential locations to test.

4.3 LOCALIZATION MODEL

Assume for the moment that there is only one whale and that a single 12-second window of time is being considered. Thus, the whale is regarded as stationary. (Multiple whales and false time delays will be introduced later.) Let $\mathbf{z}$ be an initial guess of whale position. Let there be M time-delay values available from click-train correlation between pairs of sensors. For a two-dimensional solution, at least two time-delay values from three or more sensors are assumed; and, for a three-dimensional solution, at least three time-delay values from four or more sensors are assumed. Let τ_i, $1 \le i \le M$, be the time delays, and let w_i, $1 \le i \le M$, be the associated weights that represent the approximate number of clicks that are correlated to produce the time-delay estimates.

Let $T(\mathbf{z}, s, r)$ be the model propagation time difference between sensor s and sensor r, assuming that the whale is at location $\mathbf{z}$. Let time-delay measurement τ_i be based on sensors s_i and r_i. Then, assume that

$$\tau_i = T(\mathbf{z}, s_i, r_i) + u_i,$$

where u_i is zero-mean, Gaussian, independent noise with variance σ^2. For the log-likelihood function, a weighted sum of the log-likelihood functions is formed for the independent measurements:

$$Q(\tau_1,...\tau_M;\mathbf{z},\sigma^2) = \log p(\tau_1...\tau_M|\mathbf{z}) = \sum_{i=1}^{M} w_i\left\{-\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}[\tau_i - T(\mathbf{z},s_i,r_i)]^2\right\}, \quad (3)$$

where $p(\ )$ is a PDF or a *likelihood function*. Weighting the terms by w_i is equivalent to replicating the τ_i measurement w_i times. The approach is to maximize $Q(\tau_1,...\tau_M;\mathbf{z},\sigma^2)$ over $\mathbf{z}$. Note that σ^2 can be estimated in parallel as the mean square time-delay error. The maximization of $Q(\tau_1,...\tau_M;\mathbf{z},\sigma^2)$ is more efficient if the first and second derivatives are analytically obtained.

4.3.1 First Derivatives

Let the model time-delay error be

$$u_i = \tau_i - T(\mathbf{z},s_i,r_i).$$

Let

$$\mathbf{z} = [x,y,z].$$

Taking derivatives with respect to component x of vector $\mathbf{z}$,

$$\frac{\partial Q(\tau_1,...\tau_M;\mathbf{z},\sigma^2)}{\partial x} = \sum_{i=1}^{M} w_i\left\{-\frac{u_i}{\sigma^2}(-T^x(\mathbf{z},s_i,r_i))\right\},$$

where $T^x(\mathbf{z},s,r)$ is the partial derivative with respect to the x-component of $\mathbf{z}$. Let $D(\mathbf{z},\mathbf{z}_s)$ be the distance from position $\mathbf{z}$ to sensor s:

$$\begin{aligned}
T^x(\mathbf{z},s,r) &= \frac{\partial T(\mathbf{z},s,r)}{\partial x}, \\
&= \frac{1}{C}\frac{\partial}{\partial x}\{D(\mathbf{z},\mathbf{z}_s) - D(\mathbf{z},\mathbf{z}_r)\}, \\
&= \frac{1}{C}\frac{\partial}{\partial x}\left\{[(x-x_s)^2 + (y-y_s)^2 + (z-z_s)^2]^{1/2} - [(x-x_r)^2 + (y-y_r)^2 + (z-z_r)^2]^{1/2}\right\}, \\
&= \frac{1}{2C}\left\{D^{-1}(\mathbf{z},\mathbf{z}_s)\,2(x-x_s) - D^{-1}(\mathbf{z},\mathbf{z}_r)\,2(x-x_r)\right\}, \\
&= \frac{1}{C}\left\{D^{-1}(\mathbf{z},\mathbf{z}_s)\,(x-x_s) - D^{-1}(\mathbf{z},\mathbf{z}_r)\,(x-x_r)\right\},
\end{aligned}$$

where (x_s,y_s,z_s) is the sensor s position and C is the speed of sound. Partial derivatives with respect to y and z are obtained similarly.

4.3.2 Second Derivatives

Taking second derivatives,

$$\frac{\partial^2 Q(\tau_1,...\tau_M;\mathbf{z},\sigma^2)}{\partial x \partial y} = \sum_{i=1}^{M} w_i \left\{ -\frac{1}{\sigma^2} T^x(\mathbf{z},s_i,r_i) T^y(\mathbf{z},s_i,r_i) - \frac{1}{\sigma^4} u_i T^{xy}(\mathbf{z},s_i,r_i) \right\}.$$

4.3.3 Fisher's Information Matrix

The Fisher's information matrix is a $P \times P$ matrix where P is the total number of parameters (three or four including σ^2, depending on whether depth z is required). Let x and y be two arbitrary position parameters (they can represent any pair in the set $\{x,y,z\}$. Then, the x,y element of the Fisher's information matrix is

$$I_{x,y} = -\varepsilon\left\{ \frac{\partial Q(\tau_1,...\tau_M;\mathbf{z},\sigma^2)}{\partial x \partial y} \right\},$$

where $\varepsilon\{\ \}$ is the statistical expected value (mean value). One needs to take the expected value of the negative of the above second derivatives. Note that $\varepsilon\{u_i\} = 0$. Thus,

$$I_{x,y} = \sum_{i=1}^{M} w_i \left\{ \frac{1}{\sigma^2} T^x(\mathbf{z},s_i,r_i) T^y(\mathbf{z},s_i,r_i) \right\}.$$

Note that in the above expressions x and y are general position parameters that stand for any two components of $\mathbf{z}$.

4.3.4 Solution Error Ellipse

It is important not only to know the solution $\mathbf{z}$ that maximizes equation (3) but also the error variance associated with this solution. This information is important when updating a recursive tracker. It is also critical in establishing, for example, whether the depth component can be trusted. It is very common for the x,y position to be accurate but the depth estimate to be wildly off the mark. This condition typically occurs when there are no nearby sensors. It is also of interest to know along which axis in the x,y plane is the most error expected to lie.

The x,y error ellipse is defined as a region in space (drawn around the position estimate) that can be regarded as the region in the x,y plane where—to some degree of certainty—the true position must be. It is possible to use equation (3) to estimate the solution error ellipse based on the Cramer-Rao lower bound.

The Cramer-Rao lower bound matrix (reference 11) is the inverse of the Fisher's information matrix $\mathbf{C} = \mathbf{I}_{x,y,z}^{-1}$. Here, $\mathbf{C}$ is the 3×3 error covariance matrix for the three-dimensional parameter set $\mathbf{z} = \{x, y, z\}$ for a location solution. What this value represents is the

asymptotic error covariance matrix. As the amount of data goes to infinity ($M \rightarrow \infty$), the error covariance

$$\varepsilon((\mathbf{z}-\mathbf{z}_0)(\mathbf{z}-\mathbf{z}_0)') \rightarrow \mathbf{C},$$

where $\mathbf{z}$ is a solution estimate and $\mathbf{z}_0$ is the true location.

The solution error ellipse for solution $\mathbf{z}$ is the locus of points $\mathbf{z}^*$ where the inner product

$$(\mathbf{z}-\mathbf{z}^*)'\mathbf{I}_{x,y,z}(\mathbf{z}-\mathbf{z}^*) = K,$$

where K is a constant, typically 1. These points can be drawn using eigen-analysis of matrix $\mathbf{I}_{x,y,z}$ A MATLAB code segment to do this is provided below.

```
%
% code segment to plot ellipse at solution z using FIM Ixyz
%
function plot_ellipse(z,Ixyz,min_std_x)

    % Here we calculate C=inv(I) taking into account the possibility of singular matrix Ixyz.
    % min_std_x is the minimum positional standard deviation for a component x,y, or z
    [V,D]=eig(Ixyz);
    e1=V(:,1); e2=V(:,2);
    d=diag(D);
    d=max(d, ([1000000 1000000 1000000]').^(-2));
    d = 1./d;
    d=max(d, (min_std_x*[1 1 1]').^2);
    C = V*diag(d)*V';

    % determine the individual standard deviations of each component
    xyz_std=sqrt(diag(C));

    % now isolate x,y component and draw ellipse
    [V,D]=eig(C(1:2,1:2));
    d=diag(D);

    % plot '+' at center
    plot(z(1),z(2), 'b+','linewidth',2);

    % now plot ellipse
    ang=[0:10:360]';
    u= V(1:2,1:2) * diag(sqrt(d(1:2))) * V(1:2,1:2)' * [cos(pi/180*ang) sin(pi/180*ang)]';
    plot(z(1)+u(1,:),z(2)+u(2,:));
return
```

4.3.5 Localization Maximization Iteration

The parameters of the solution include the time-delay error variance σ^2 and the position vector

$$\mathbf{z} = [xyz]'.$$

Let $\mathbf{z}_k$ be the position vector at the k^{th} iteration. The algorithm update equation is

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \mathbf{I}_{xyz}^{-1}(\mathbf{z}_k)\delta(\mathbf{z}_k),$$

where $\mathbf{I}_{xyz}(\mathbf{z}_k)$ is the Fisher's information matrix and δ is the vector of first derivatives (of $Q(\mathbf{z}, \sigma^2)$ with respect to x, y, and z). Both are computed using previous solution $\mathbf{z}_k$.

4.4 LOCALIZATION INITIALIZATION

To localize, it is necessary to obtain starting points. One possible initialization approach is to determine likely starting points based on intersections of hyperbolas. Each time-delay measurement corresponds to a locus of points forming a hyperbola in a two-dimensional geometric plot (see figure 6). The depth of the whale must be assumed in order to create the hyperbolas. A depth of 500 meters can be used. Figure 6 shows such a set of hyperbolas drawn using the time delays obtained from a single 12-second window from dataset 3. The initial position needs to be accurate enough so that the time-delay errors are small. Using the geometric mean of the sensor positions, for example, would probably not be close enough as an initial guess. For the manual initialization, an operator manually entered a position based on observing the hyperbola plot.

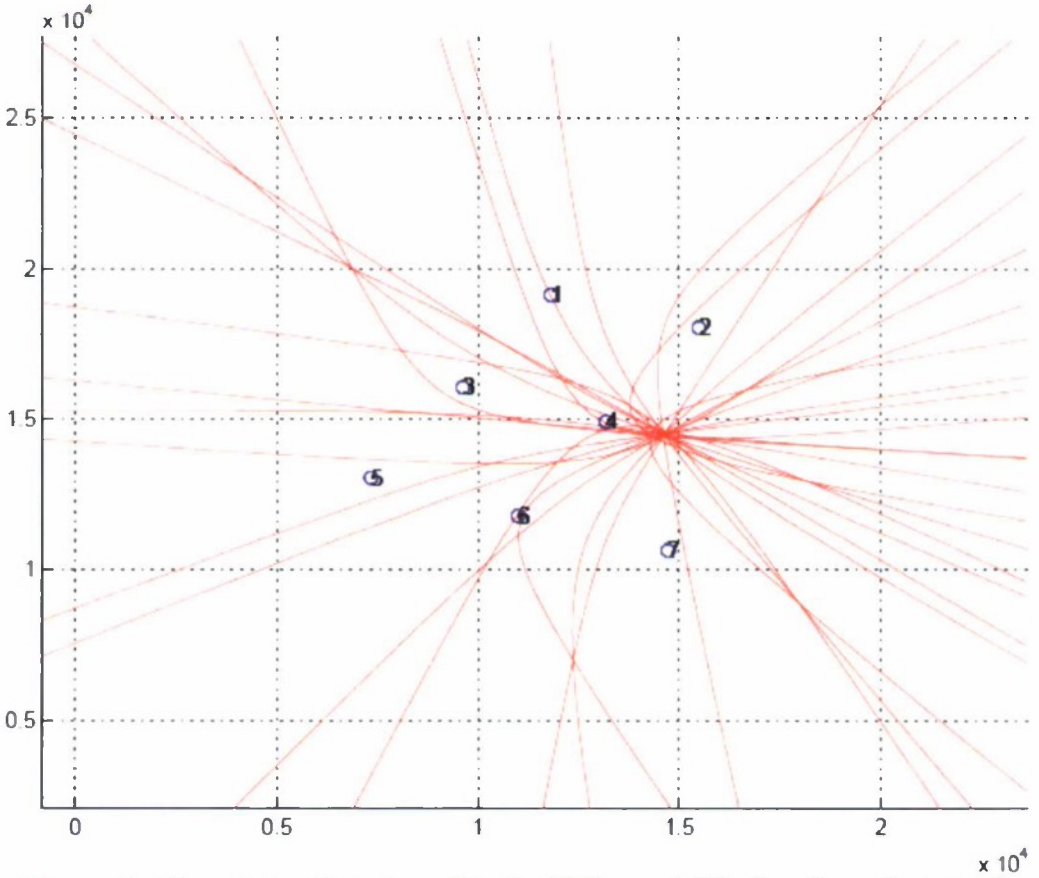

Figure 6. Hyperbola Plots for a Single 12-Second Window from Dataset 1 Using Initial Depth Estimate of 500 Meters

4.5 MULTIPLE WHALES AND FALSE TIME DELAYS

Not all of the inter-sensor time-delay measurements that are produced for a given data time window are generally associated with a given whale position solution. Either they are false time-delays altogether, or they may be associated with other valid solutions.

To handle the possibility that some time delays are not associated with the given solution, it is necessary to augment the weight w_i in equation (3) with a membership probability calculated for the current position estimate $\mathbf{z}$. Thus, let there be K potential solutions with current localization solutions $\mathbf{z}_k$. The membership probabilities are given by

$$w_{k,i}, \quad 1 \le k \le K, \quad 1 \le i \le M.$$

Thus, a given time delay i is "soft-assigned" to the solutions. Two ways of calculating the membership probabilities $w_{k,i}$ are considered—"open-loop" (single pass) and "closed-loop" (iterative).

4.5.1 Open-Loop Membership Weights

For the open-loop method, a fixed weight that depends only on the time-delay error is assigned. The weight used here is

$$w_{k,i} = n_i \exp\left(-\frac{(\tau_i - T(\mathbf{z}_k, s_i, r_i))^2}{2\sigma_m^2} \right),$$

where σ_m is the fixed membership time-delay standard deviation of 0.1 second and n_i is the number of clicks associated with time-delay measurement i. The standard deviation σ_m acts as a "soft" threshold for time delays to be included in solution k.

4.5.2 Closed-Loop Membership Weights

The open-loop method fails if time delays are close matches to more than one solution. Take, for example, figure 7 where three hyperbolas intersect more than one solution (there are four solutions indicated by the star-like convergence points for like-colored hyperbolas). Because each of these time delays (hyperbolas) is consistent with more than one solution, the open-loop scheme would probably assign each of these time delays more or less equally to two solutions, biasing the localization positions. There is no easy way to know *a priori* which solution the time delays belong to. The color-coding seen in the figure is an automatically generated grouping based on correlation, but it cannot be relied on in every instance to make the correct associations. Thus, at the outset it must be assumed that any time delay can associate with any solution.

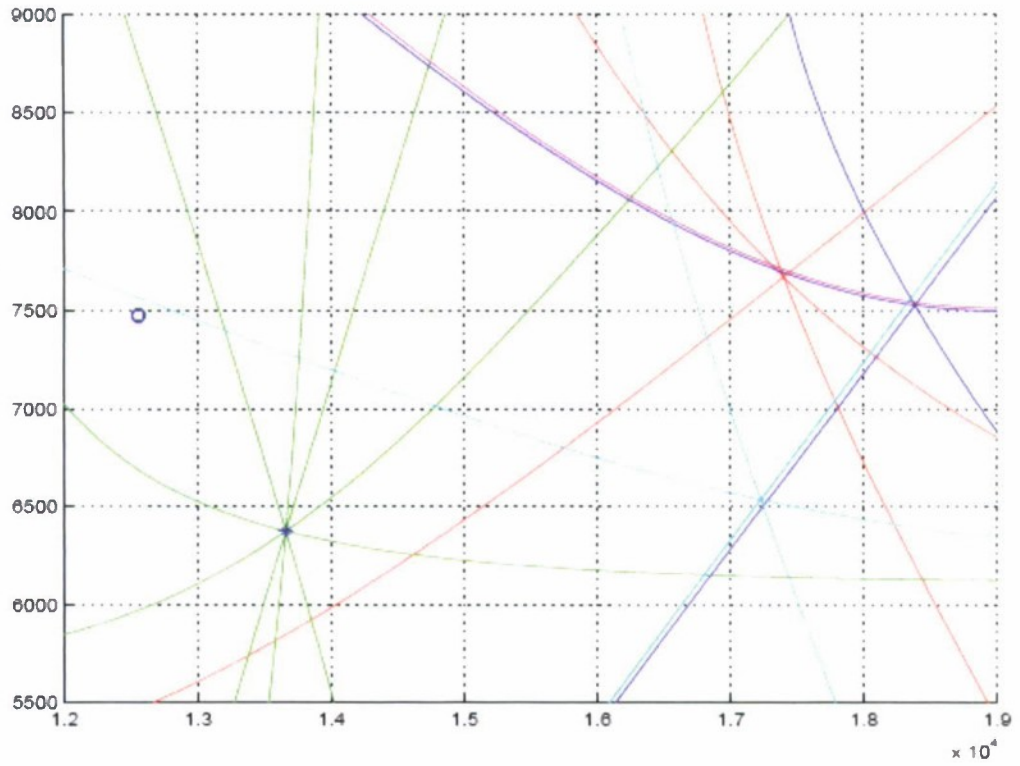

Figure 7. Hyperbolas from Dataset 2 Simultaneously Showing Four Whales (A depth of 760 meters is assumed; this is the solution depth for the left-most whale.)

Closed-loop membership probabilities allow an interactive "negotiation" to take place that attempts to resolve "disputes" happening when a given time delay suits more than one solution. These "disputes" can be resolved if, for a given solution, the time-delay errors for the correctly associated time delays are smaller than the time-delay errors for the falsely associated time delays. Then, the incorrectly associated time delays will have lower weight. This lower weight will reduce the bias, which in turn will drive the solution closer to the correct location, further reducing the time-delay errors for the correctly associated time delays. Eventually, the incorrectly associated time delays will be driven out of the solution altogether.

Let α_k be the weight of solution k and assume that

$$\sum_{k=1}^{K} \alpha_k = 1.$$

Weight α_k may be thought of as a relative measure of the number of time delays that each solution "owns." The process is started by assuming that the solutions are equally likely, setting $\alpha_k = 1/K$. Let σ_k be the solution-dependent time standard deviation. One can initialize σ_k to 0.1 second, for example.

To proceed, the unnormalized weights are computed:

$$\tilde{w}_{k,i} = \alpha_k n_i (2\pi\sigma_k^2)^{-1/2} \exp\left(\frac{(\tau_i - T(\mathbf{z}_k, s_i, r_i))^2}{2\sigma_k^2} \right).$$

The weights are then normalized to sum to 1 over k:

$$w_{k,i} = \frac{\tilde{w}_{k,i}}{\sum_{l=1}^{K} \tilde{w}_{l,i}}.$$

These weights are used in the update of the solutions (section 4.3). In contrast to the open-loop method, these weights grow larger if σ_k^2 grows smaller (the solution is tighter) or if solution k becomes "stronger" (larger α_k). At each iteration, the localization solutions $\mathbf{z}_k$ are updated using the new time-delay membership weights $w_{k,i}$. On the next iteration, the solution weights are recalculated:

$$\alpha_k = \frac{\sum_{i=1}^{M} w_{k,i}}{\sum_{l=1}^{K} \sum_{i=1}^{M} w_{l,i}}.$$

The algorithm is summarized as follows:

1. Initialize α_k, σ_k^2, $1 \le k \le K$, as described.

2. Compute $\tilde{w}_{k,i}$, $1 \le k \le K$, $1 \le \mathrm{i} \le M$, as described.

3. Compute $w_{k,i}$ from $\tilde{w}_{k,i}$ as described.

4. Update α_k as described.

5. Execute one iteration of the localization algorithm (section 4.3.5), noting that σ^2 and w_i are replaced by the solution-dependent σ_k^2 and $w_{k,i}$.

6. Go to step 2.

As the algorithm iterates, weak or redundant solutions are "pruned." A solution is eliminated if (1) α_k falls below a certain threshold, or (2) two solutions become physically closer than about 30 meters (the weaker solution is removed), or (3) the membership functions of two solutions become highly correlated, indicating that they represent the same set of time delays.

5. APPLICATION TO REAL DATA

5.1 DATASETS

Five localization datasets are considered. Details of these datasets are given in table 5.

Table 5. Dataset Details

Dataset	Length (minutes)	No. of Sensors	No. of Whales	Description	Identi-fication	Nickname
1	25	5	1	Monaco 2005 #2	(set 1-5)	"monaco"
2	20	6	4	Monaco 2005 #1	(set 6-9)	"monaco"
3	50	7	1	AUTEC BRS Drive 2007 Dive 1	(set 10)	"new"
4	50	8	1	AUTEC BRS Drive 2007 Dive 2	(set 11)	"new"
5	50	7	1	AUTEC BRS Drive 2007 Dive 3	(set 12)	"new"
6	>3 hours	3	2^{+}	AUTEC BRS Drive 2007 Set 3-Song 1	(set 13)	"latest"

5.2 CLICK SEPARATION

5.2.1 First Case Study

An example of the click-separation algorithm output for a 12-second input time window is shown in figure 8 for an "easy" problem consisting of a single sperm whale. Data are from dataset 1. The algorithm has organized the clicks into five groups (click-trains). A detailed analysis of each click-train is provided in figures 9 through 11. The clicks associated with the direct path are shown in figure 9, the clicks associated with the surface path are shown in figure 10, and the clicks associated with a bottom path are shown in figure 11. In the top panel of each of these figures, the click amplitude is plotted as a function of time. In the middle panel, a concatenation of the short-time Fourier transform (STFT) or "spectrogram" of each click is shown. In the bottom panel, inter-click correlograms are shown. For each consecutive click-pair in the sequence, the correlogram amplitude is plotted as intensity in a vertical strip. Each vertical strip is centered at the maximum amplitude, which results in the horizontal line in the center.

Several clues establish the identity of the direct-path click-train. First is the consistent arrival time difference between the direct and the reflected paths. This clue is difficult to see by eye in figure 8. A naive algorithm can be easily fooled by being one period ahead or behind. To reliably establish time precedence, it is necessary to use correlation (see section 3.2). The second clue is the relative SNR of the direct path seen in figure 8. The direct path is often but not always the strongest path. The third clue is evidence of the inter-pulse interval (which is normally explained as a reverberation occurring within the animal's skull). This evidence appears in the correlograms (bottom panel in figure 9). A consistent 4-msec peak (both at +4 and –4 msec) can be seen. This evidence also appears in the spectrogram (middle panel of figure 9).

A closeup is shown in figure 12. Again, the 4-msec inter-click interval can be faintly seen. The inter-click interval varies, possibly with aspect angle, and tends to be lost in reflected paths due to multipath distortion.

Further evidence that this click-train is a direct path is shown in table 6. This table lists classification statistics for the 14 clicks in the direct-path click-train in figure 8. Columns are SNR and probability (in percent) for direct path, surface path, and reverberation path. These last three columns are the result of the likelihood comparison described in section 3.1. Of the 14 single-click classifications, 9 are direct path. Many, however, are classified as surface path. These statistics demonstrate the difficulty of using single-click information to classify clicks, especially direct path versus surface path. Additional information will be obtained from correlation of click-trains (see section 3.2). Despite the fact that several of the clicks are classified as surface, one can still establish them as direct path.

Consider the same statistics generated from the surface-path clicks (table 7). For these, the surface-path probability is 100 for all clicks but one. This is a common occurrence and is probably due to the very unique character of the surface path seen in this dataset.

Table 6. Single-Click Statistics for Direct-Path Clicks
[1 4 7 10 14 17 21 25 28 31 34 37 40 43]

SNR	P (Direct)	P (Reverberation)	P (Surface)
186	100	0	0
204	97	0	3
204	100	0	0
189	100	0	0
167	97	0	3
180	93	0	7
156	99	0	1
128	99	0	1
108	87	0	13
93	0	0	100
83	8	0	92
86	1	0	99
65	40	0	60
61	16	0	84

Table 7. Single-Click Statistics for Surface-Path Clicks [3 6 9 12 15 19 23 27 30 33 36 39 42].

SNR	P (Direct)	P (Reverberation)	P (Surface)
111	0	0	100
146	0	0	100
112	0	0	100
126	0	0	100
158	0	0	100
131	0	0	100
82	0	0	100
96	0	0	100
94	0	0	100
112	0	0	100
65	1	16	83
130	0	0	100
141	0	0	100

Figures 10 – 12 show distinctly different correlograms and spectrograms. The correlograms are more random-looking and, especially for the reverberation path, the amplitude envelope is far more diffuse. For the reverberation path, the single-click classification statistics in table 8 are consistently for reverberation except for one click. It can be seen that a classification of the entire click-train can be accomplished reliably by averaging or in some way combining the single-click classifications. This must be augmented by information from correlating the click-trains.

Table 8. Single-Click Statistics for Reverberation-Path Clicks [18 22 26 29 32 35 38 41 44].

SNR	P (Direct)	P (Reverberation)	P (Surface)
5	13	85	2
4	0	100	0
5	1	95	4
4	2	94	4
5	0	100	0
3	45	55	0
4	0	100	0
4	96	4	0
3	0	100	0

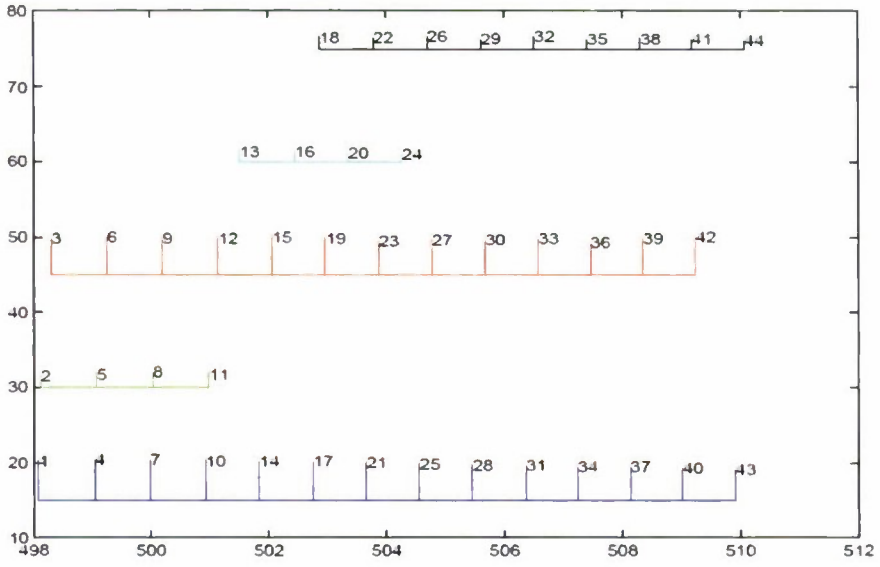

Figure 8. Illustration of a Click Grouping for a 12-Second Time Window After Convergence (First Case Study)

(Click SNR is plotted as a function of time. Click-trains have been rendered in separate colors and vertically separated for clarity. Clicks [1 4 7 10 14 17 21 25 28 31 34 37 40 43] are the direct path, clicks [3 6 9 12 15 19 23 27 30 are a reverberation path, probably bottom-reflected. Clicks [13 16 20 ...] are possibly a second bottom-reflected path.)

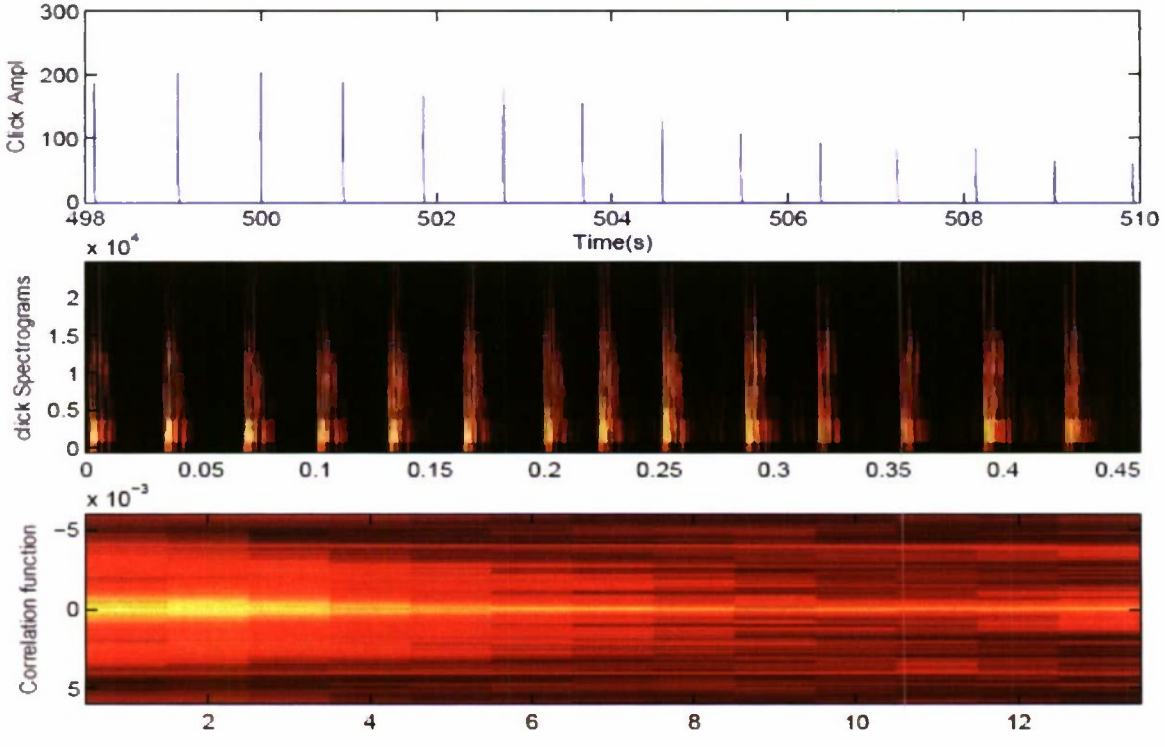

Figure 9. Analysis of Direct-Path Click-Train [1 4 7 10 14 17 21 25 28 31 34 37 40 43]

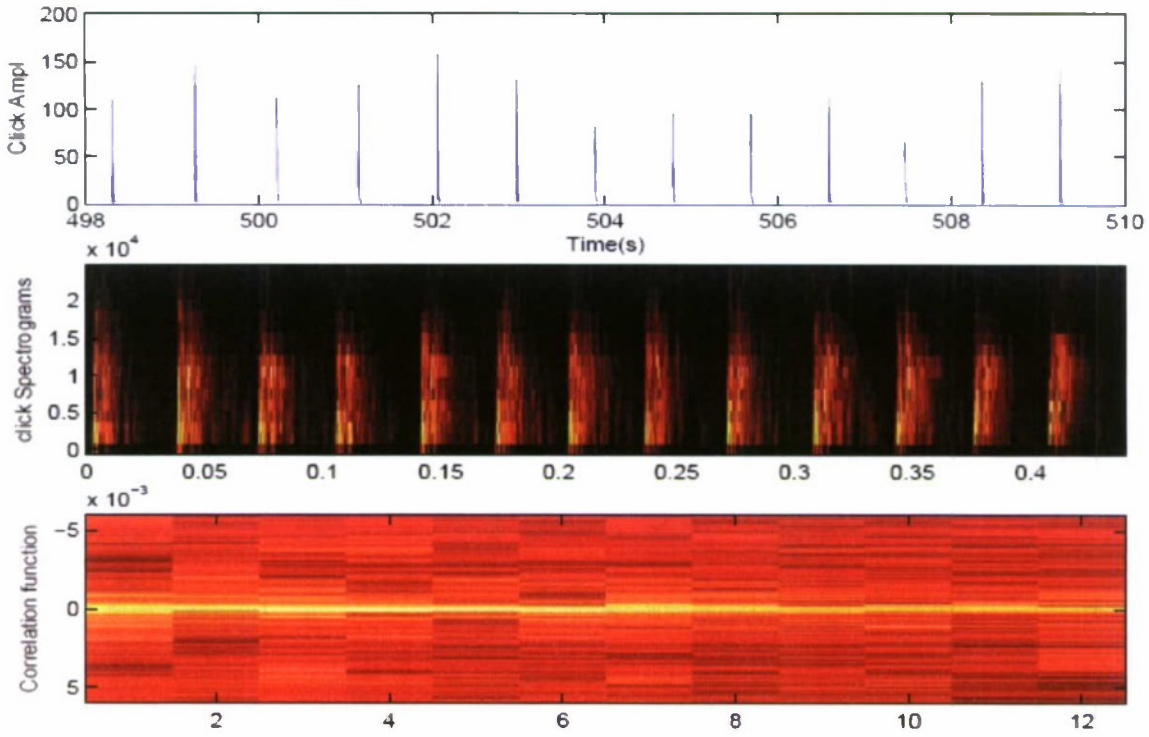

Figure 10. Analysis of Click-Train [3 6 9 12 15 19 23 27 30 33 36 39 42]

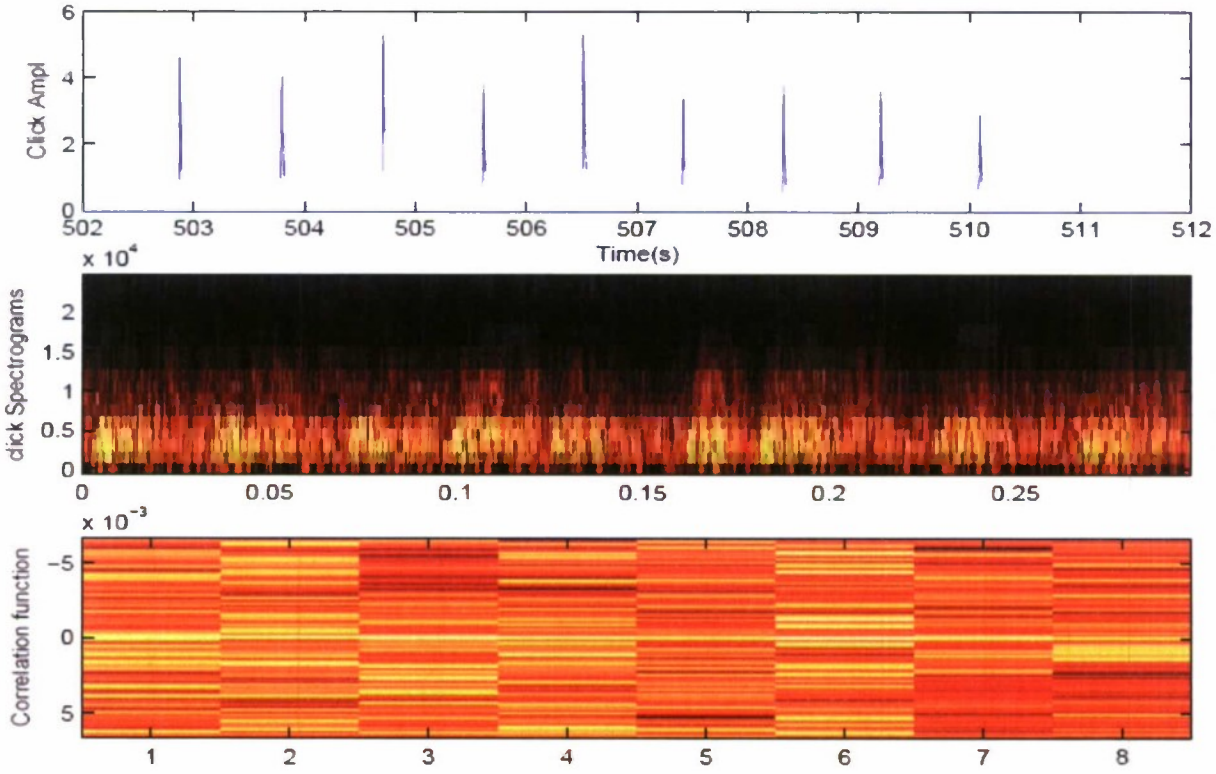

Figure 11. Analysis of Click-Train [18 22 26 29 32 35 38 41 44]

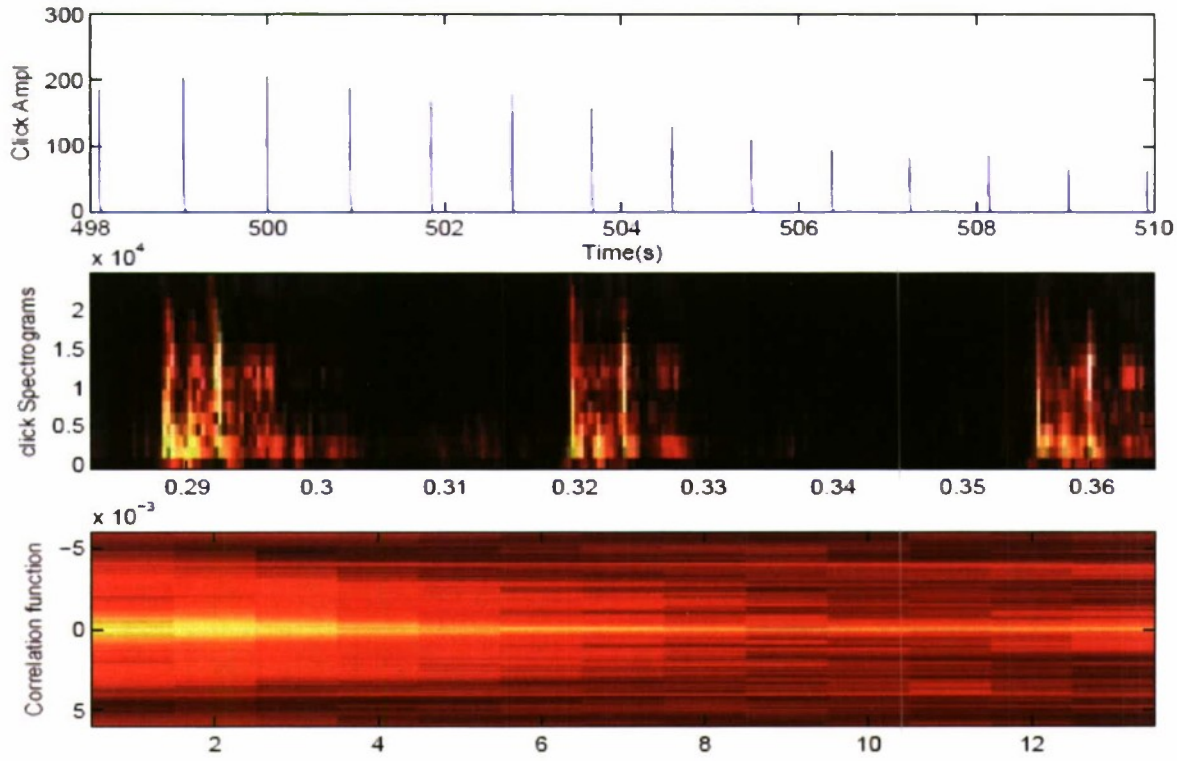

Figure 12. Closeup Spectrograms for Three Clicks Showing a 4-msec Inter-Click Interval

5.2.2 Second Case Study

A second example of the click-separation algorithm output for a 12-second input time window is shown in figure 13 for a problem consisting of two sperm whales. The algorithm has organized the clicks into six main groups (click-trains). In this example, the clicks are drawn in a color depending on the result of the single-click classifier (section 3.1). The baseline of the click-train is painted in a color depending on the classification of the entire click-train based on summing the single-click probabilities of the component clicks. Data are from dataset 6.

To validate the results of the display, the operator has several possibilities. A detailed analysis of the direct-path click-train [3 8 14 ...] is provided in figure 14. A detailed analysis of the surface-path click-train [4 9 15 ...] is provided in figure 15. A detailed analysis of the direct-path click-train [30 35 41 ...] is provided in figure 16. The characteristic clues can be seen for each type of click-train, as described in section 5.2.1.

To further validate the classifications, click-train correlation is used (section 3.2). Click-train correlation results for click-trains [2 7 11 ...] and [4 9 15 ...] are shown in figure 17. The

lower graph in this figure shows an overlay of the click times for the click-trains. When the circles coincide with the dots at the estimated delay, a correct solution is indicated. This supports the classification of these chains as direct and surface. The characteristic of a surface path is a delay of less than about 0.5 second.

Correlation results for click-trains [2 7 11 ...] and [30 35 41 ...] are shown in figure 18. The results support the classification of these chains as direct path and bottom reverberation. The characteristic of a reverberation path is a delay of greater than about 0.5 second.

Correlation results for click-trains [2 7 11 ...] and [5 10 16 ...] are shown in figure 19. The results support the classification of these chains as unrelated.

Correlation results for click-trains [3 8 14 ...] and [5 10 16 ...] are shown in figure 20. The results support the classification of these chains as direct and surface.

Correlation results for click-trains [3 8 14 ...] and [29 37 43 ...] are shown in figure 21. The results support the classification of these chains as direct and bottom reverberation.

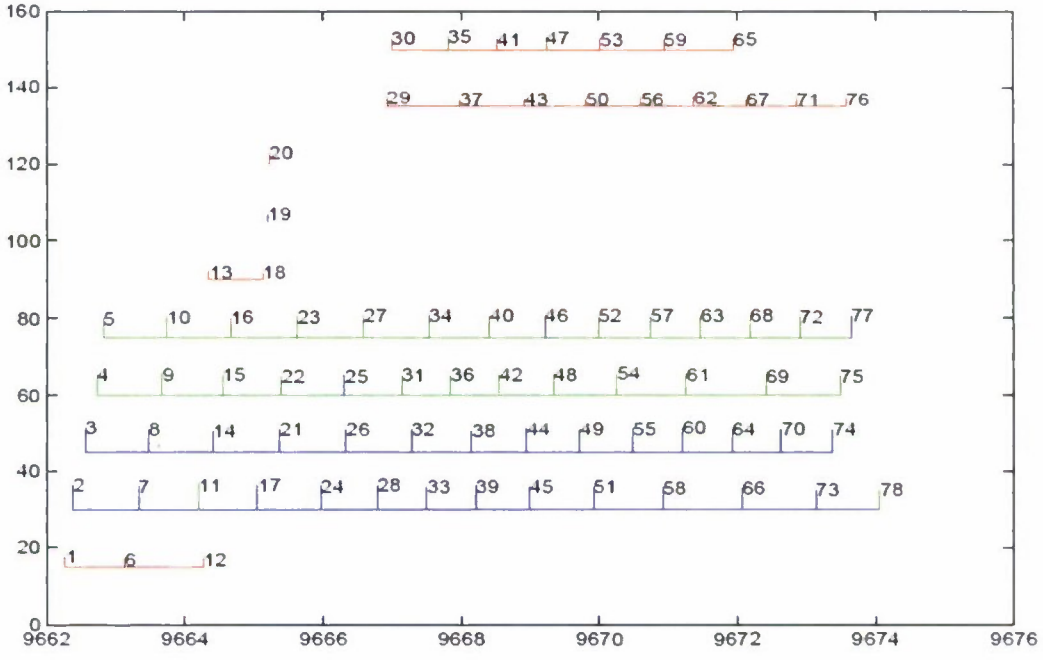

Figure 13. Illustration of a Click Grouping for a 12-Second Time Window After Convergence (Second Case Study)

(Click SNR is plotted as a function of time. Click-trains are rendered in separate colors depending on click-type classification (blue = direct path, green = surface path, red = bottom reverberation) and vertically separated for clarity. Clicks [2 7 11 ...] and [3 8 14 ...] are direct path; clicks [4 9 15 ...] and [5 10 16 ...] are surface; and [29 37 43 ...] and [30 35 41 ...] are bottom reverb. File: "Set3-Song1-082307-0545L nvt2.31".)

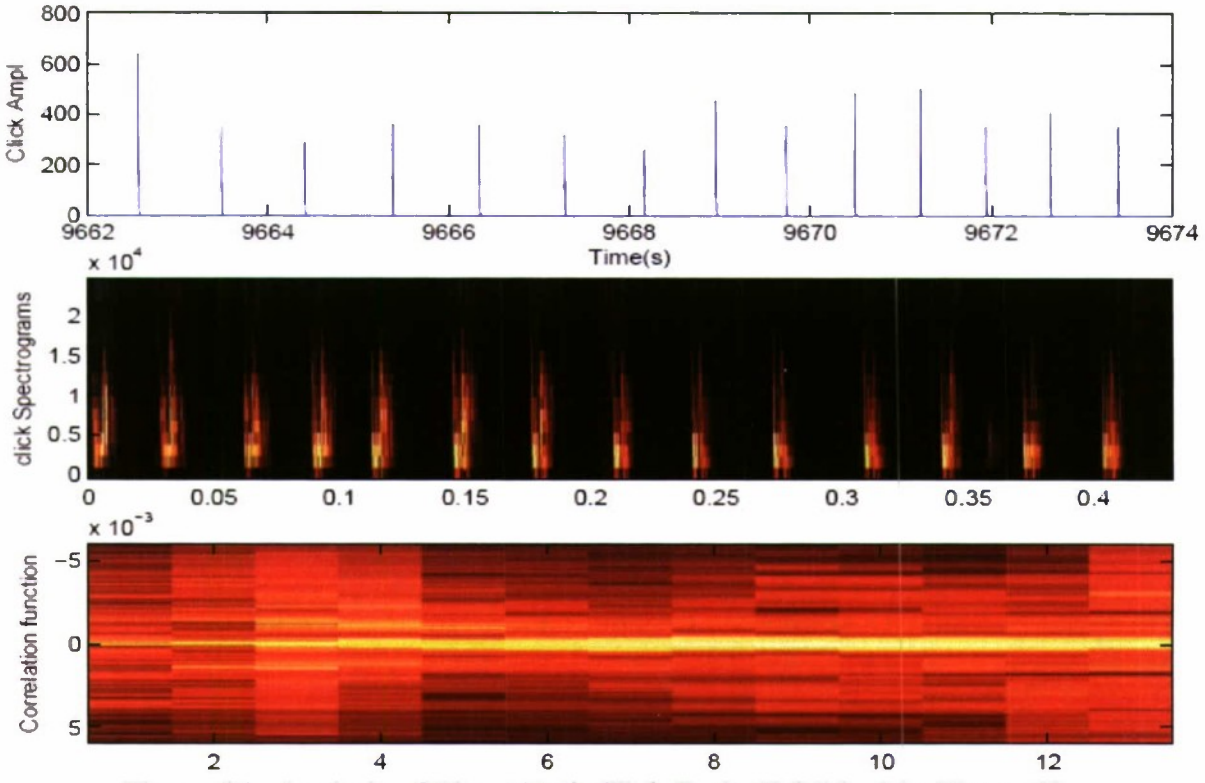

Figure 14. Analysis of Direct-Path Click-Train [3 8 14 ...] in Figure 13

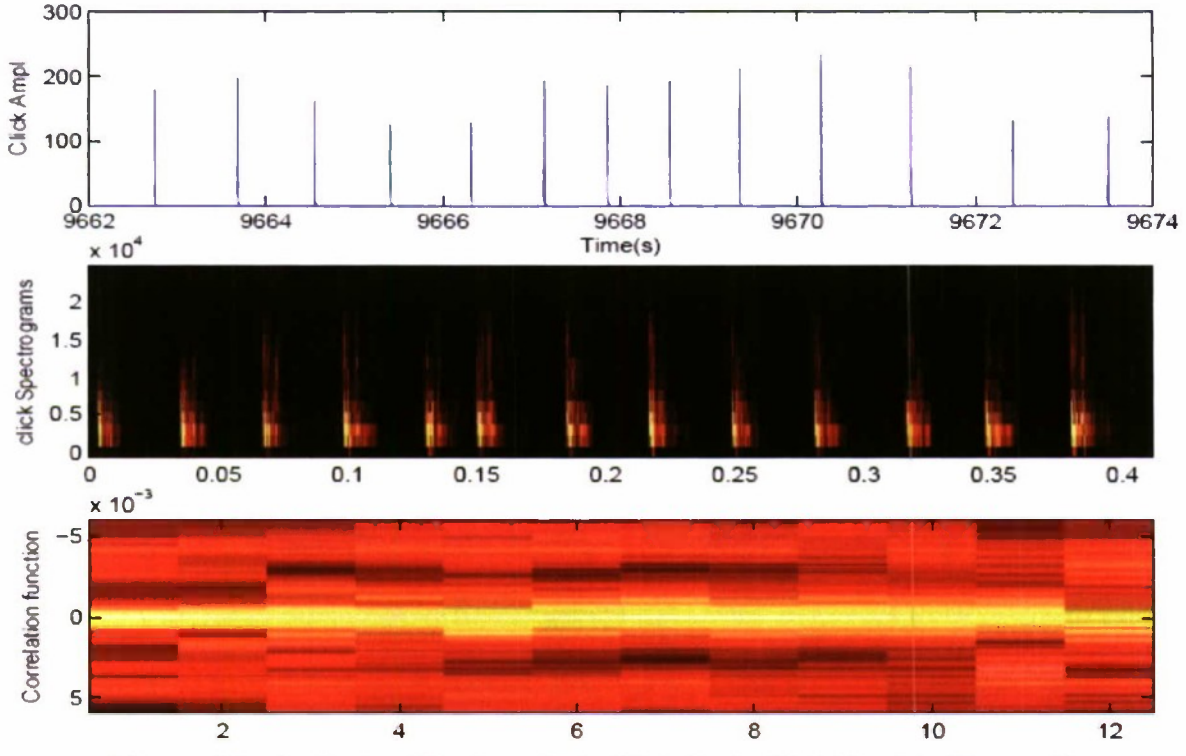

Figure 15. Analysis of Surface-Path Click-Train [4 9 15 ...] in Figure 13

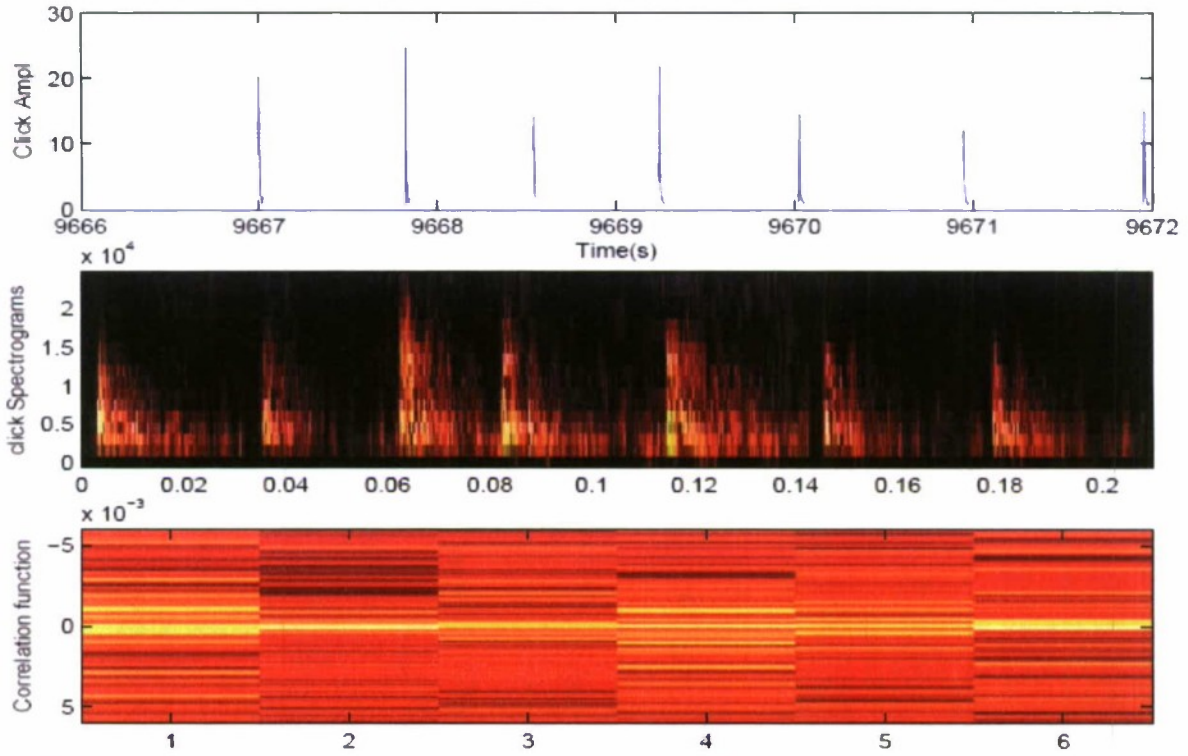

Figure 16. Analysis of Bottom Reverberation Click-Train [30 35 41 ...] in Figure 13

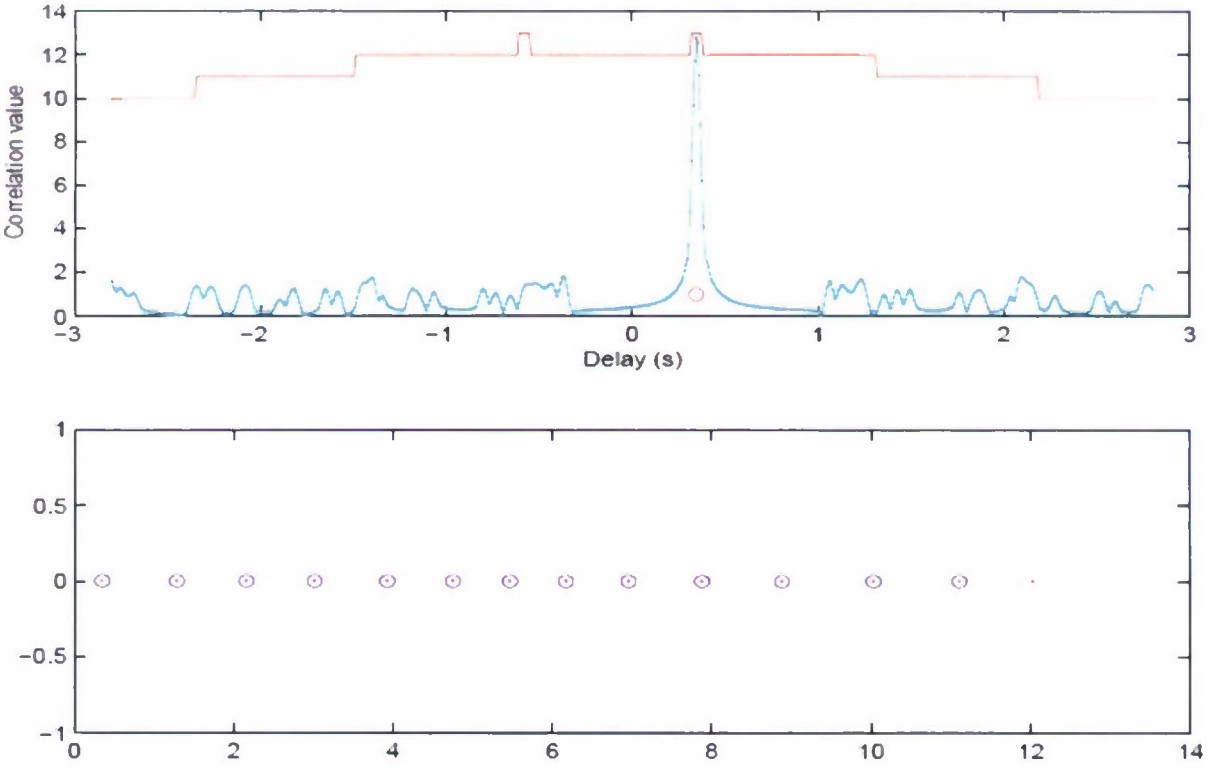

Figure 17. Inter-Click-Train Correlation Results for Click-Trains [2 7 11 ...] and [4 9 15 ...] in Figure 13

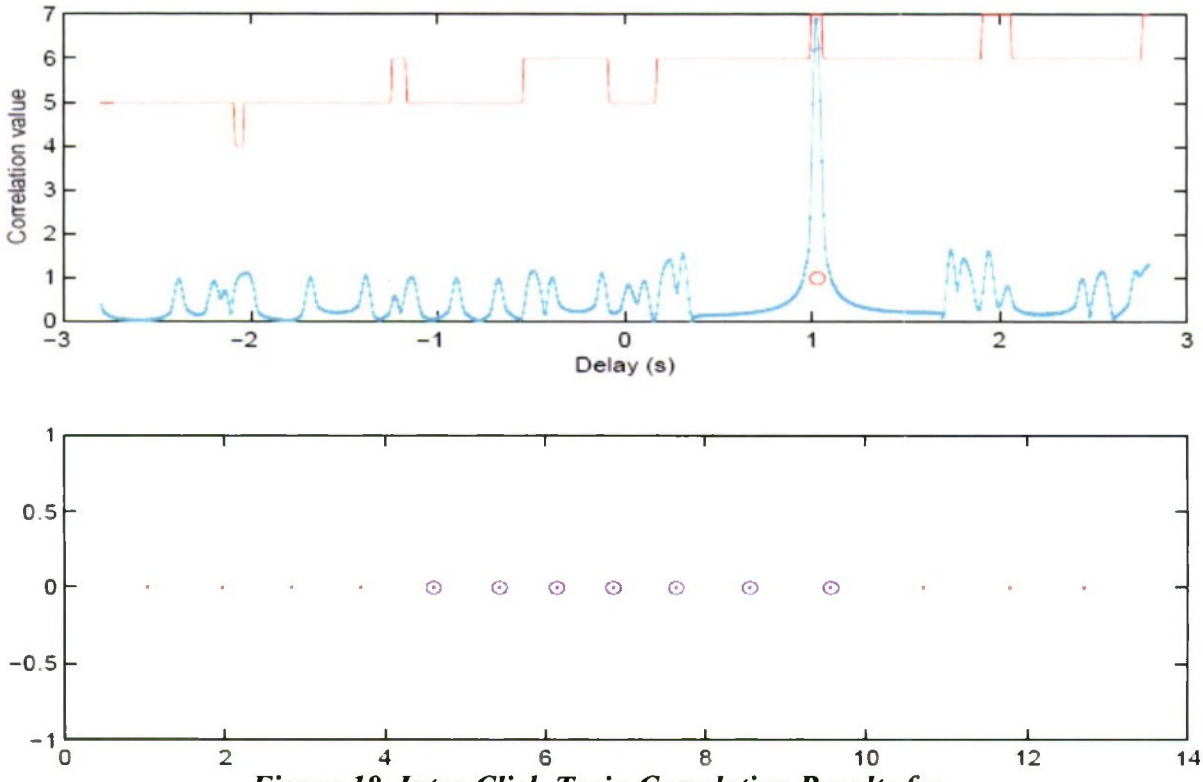

Figure 18. Inter-Click-Train Correlation Results for Click-Trains [2 7 11 ...] and [30 35 41 ...] in Figure 13

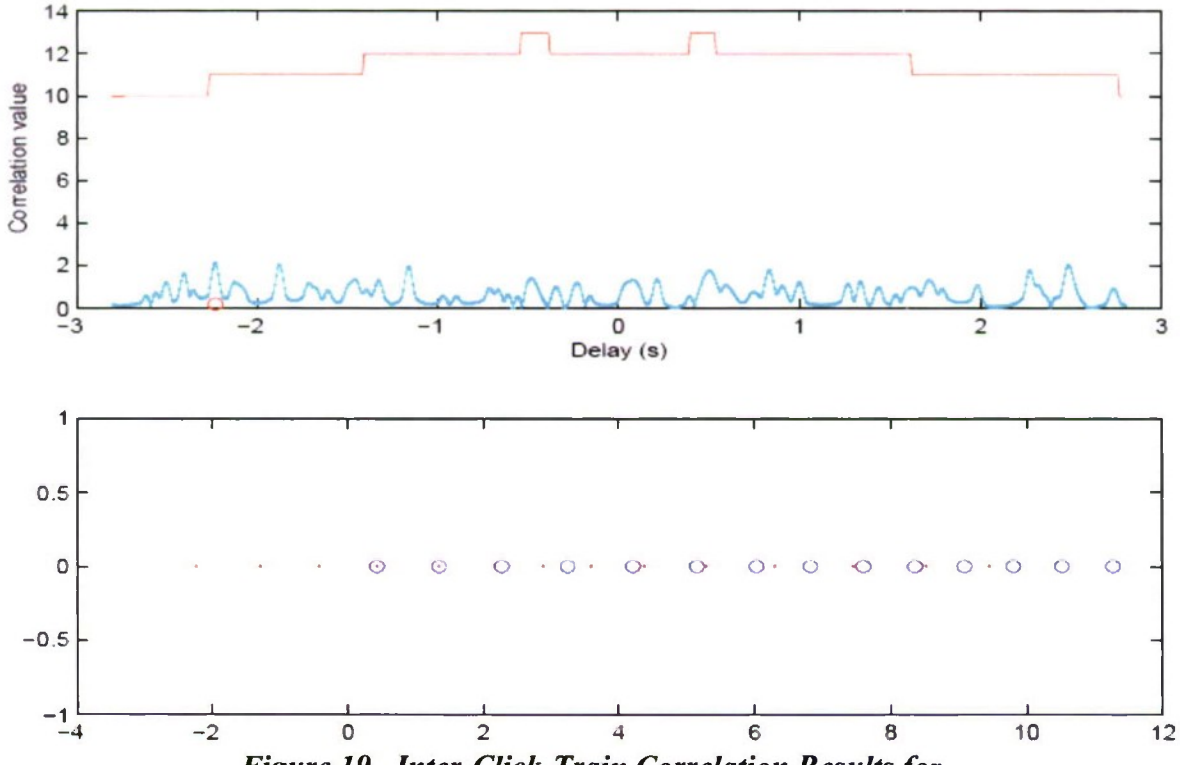

Figure 19. Inter-Click-Train Correlation Results for Click-Trains [2 7 11 ...] and [5 10 16 ...] in Figure 13

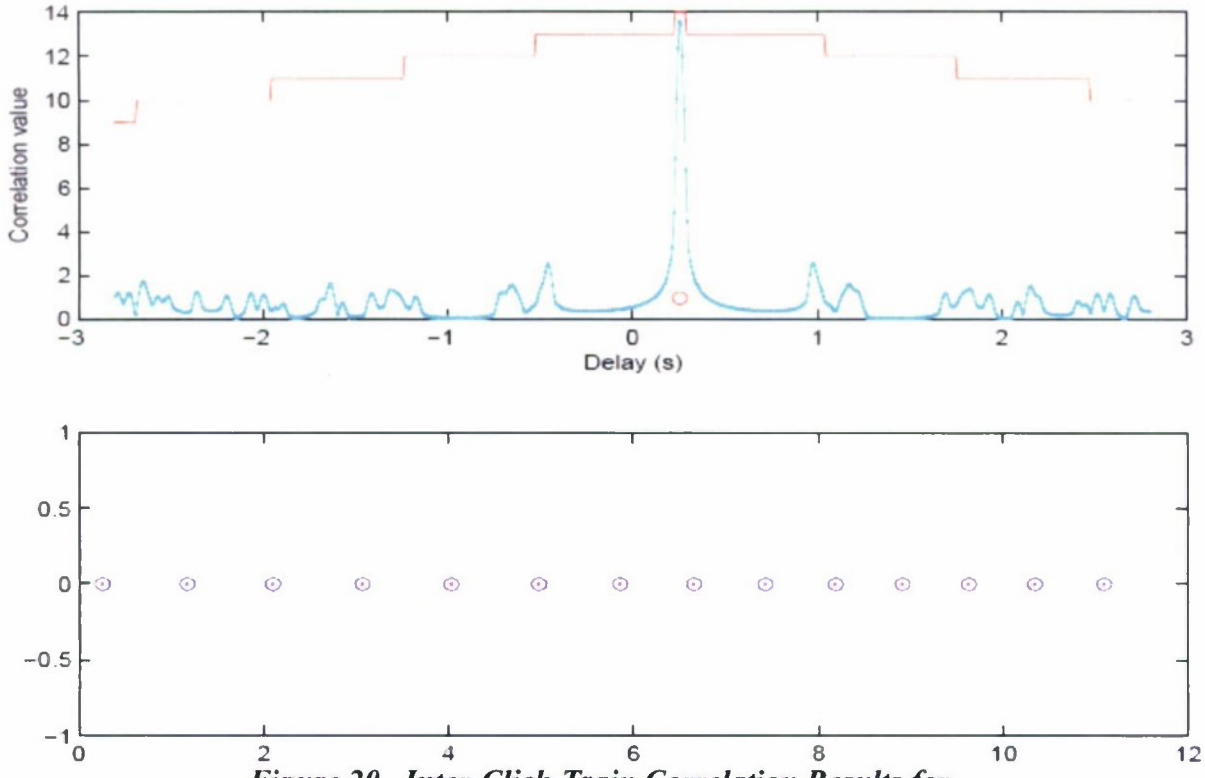

Figure 20. Inter-Click-Train Correlation Results for Click-Trains [3 8 14 ...] and [5 10 16 ...] in Figure 13

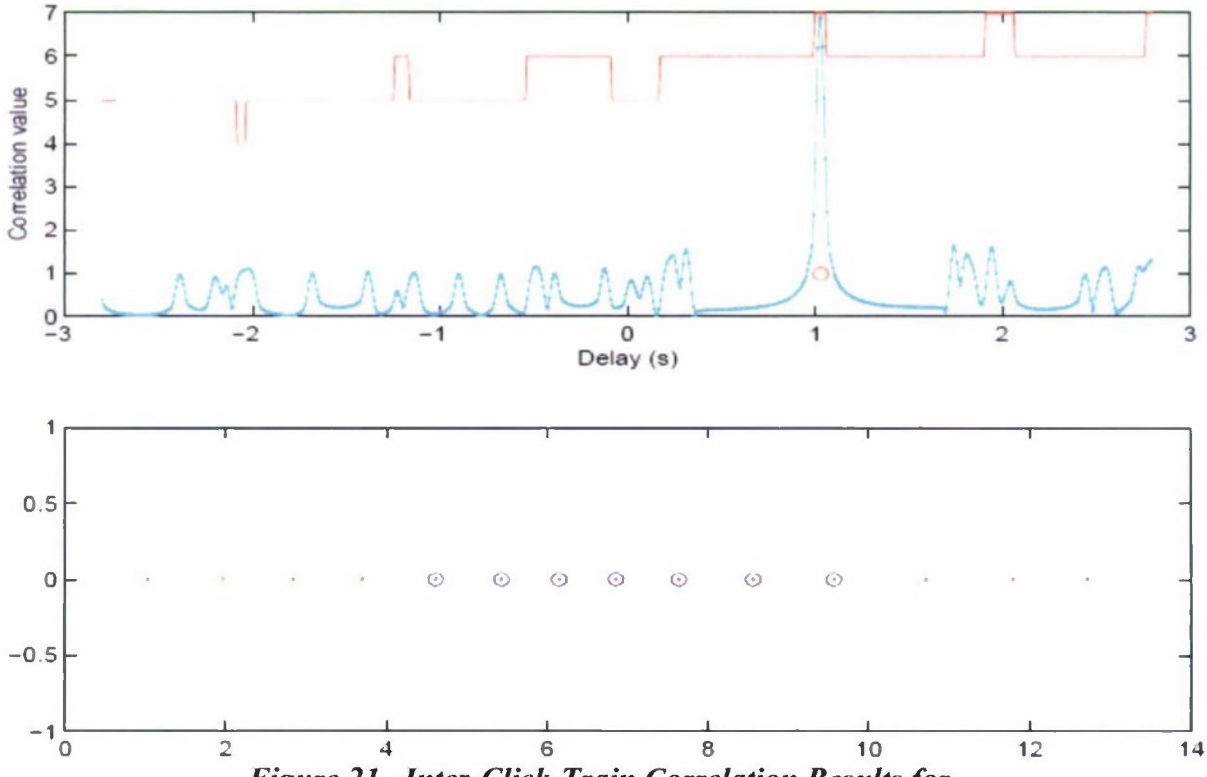

Figure 21. Inter-Click-Train Correlation Results for Click-Trains [3 8 14 ...] and [29 37 43 ...] in Figure 13

5.2.3 Third Case Study

The third case study, also from dataset 6, is peculiar in that there are two reverberation paths (figure 22). Click-train correlation results for click-trains [14 18 22 ...] and [2 8 11 ...] are shown in figure 23. The results support the classification of these chains as direct and surface.

Correlation results for click-trains [2 8 11 ...] and [10 13 17 ...] are shown in figure 24. The results show surface to first bottom reverberation delay.

Correlation results for click-trains [10 13 17 ...] and [3 9 12 ...] are shown in figure 25. The results show first bottom to second bottom reverberation delay.

Correlation results for click-trains [14 18 22 ...] and [3 9 12 ...] are shown in figure 26. The results show the full direct to second bottom reverberation delay.

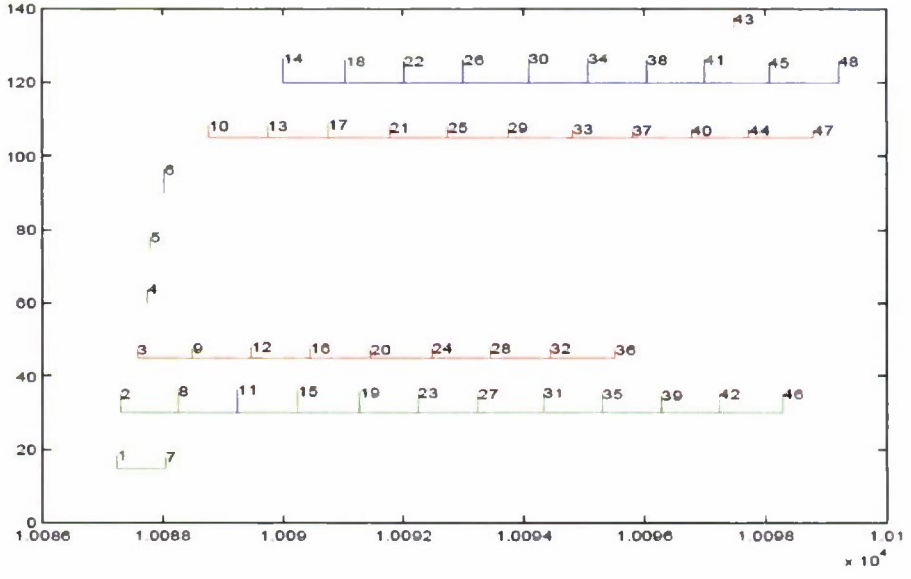

Figure 22. Illustration of a Click Grouping for a 12-Second Time Window After Convergence (Third Case Study)
(Click SNR is plotted as a function of time. Click-trains have been rendered in separate colors depending on the click-type classification (blue = direct path, green = surface path, red = bottom reverberation) and are vertically separated for clarity. File: "Set3-Song1-082307-0545L nvt2.32".)

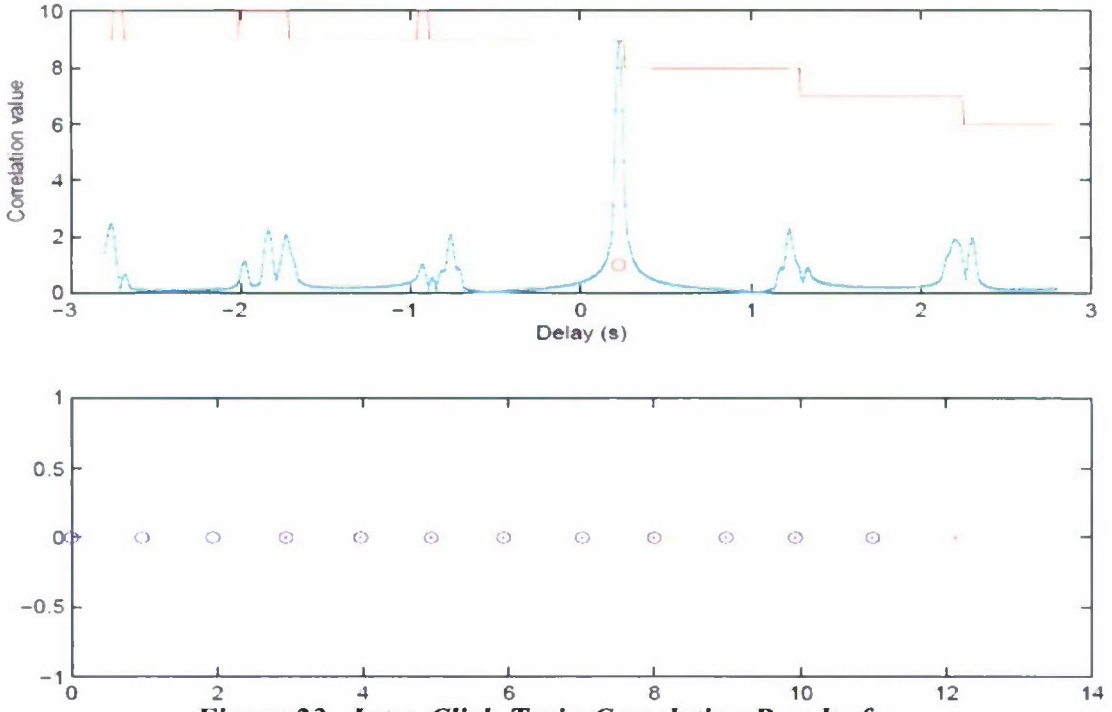

Figure 23. Inter-Click-Train Correlation Results for Click-Trains [14 18 22 ...] and [2 8 11 ...] in Figure 22

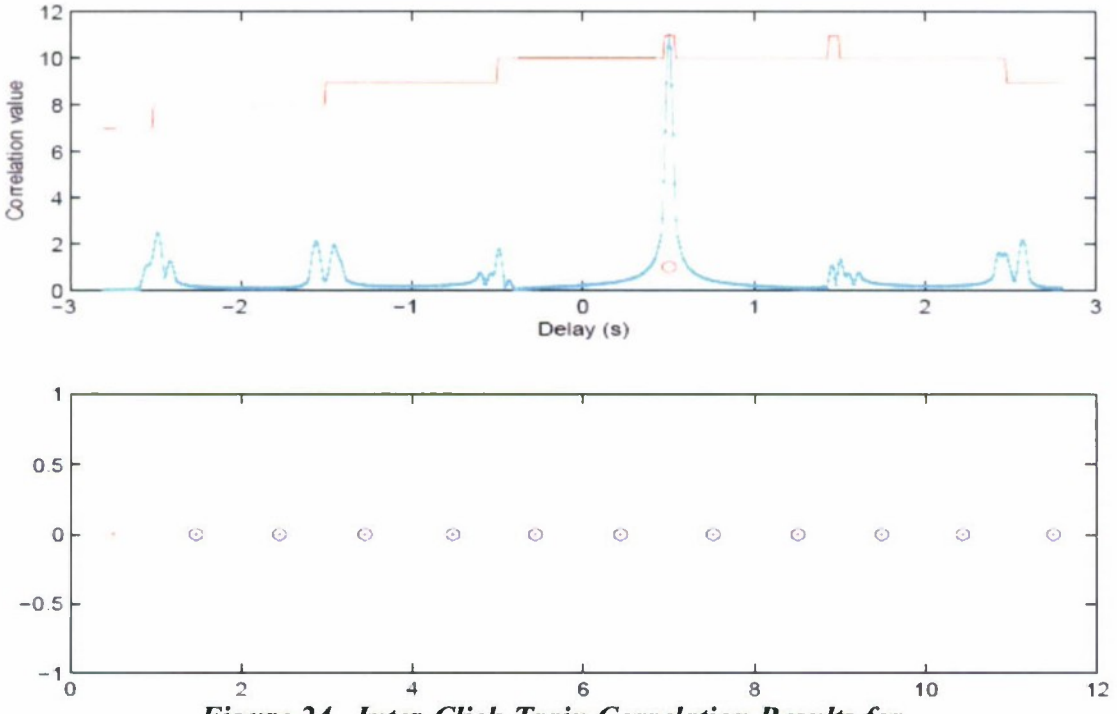

Figure 24. Inter-Click-Train Correlation Results for Click-Trains [2 8 11 ...] and [10 13 17 ...] in Figure 22

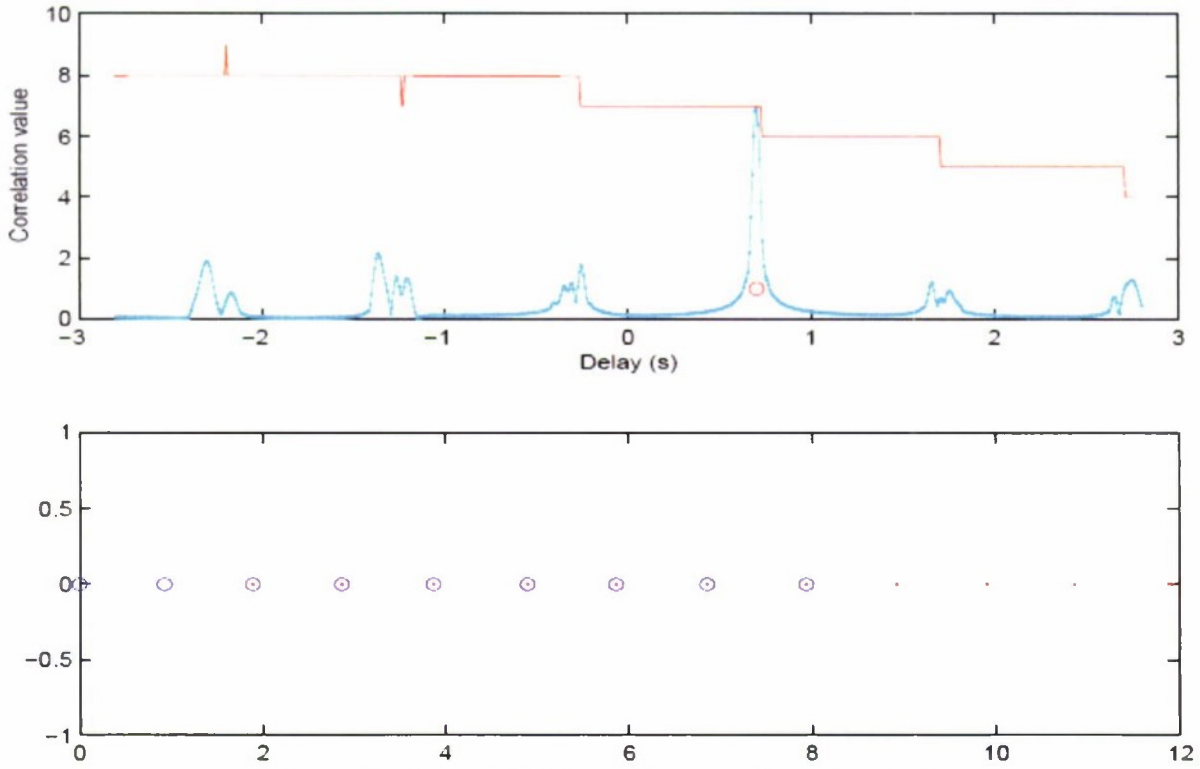

Figure 25. Inter-Click-Train Correlation Results for Click-Trains [10 13 17 ...] and [3 9 12 ...] in Figure 22

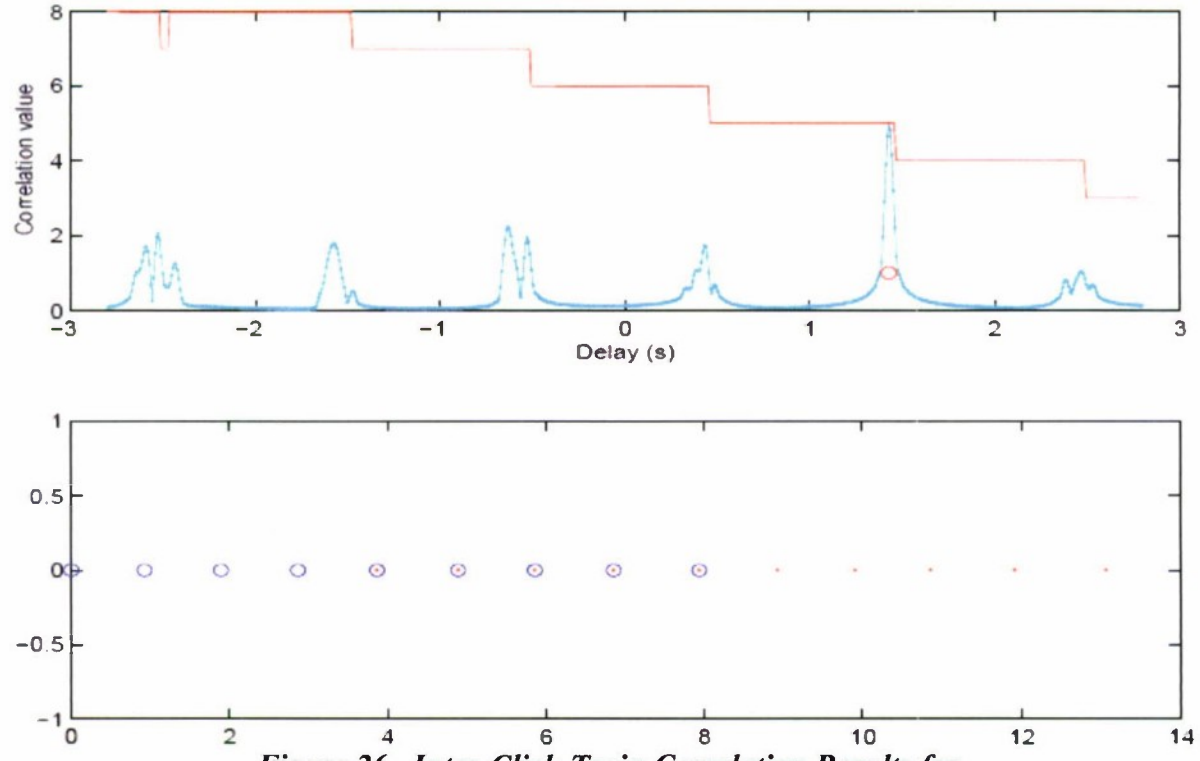

Figure 26. Inter-Click-Train Correlation Results for Click-Trains [14 18 22 ...] and [3 9 12 ...] in Figure 22

5.3 TIME-DELAY ESTIMATES

There is no better way to assess the performance of multipath rejection than to view the resultant time-delay estimates from each pair of sensors. Time-delay estimates are plotted as a function of time for datasets 1 through 6 in figures 27 through 32. A separate graph is plotted for each hydrophone pair. In figure 27 (dataset 1), a consistent time-delay trace is seen for each hydrophone pair. No false time delays are noted for the entire dataset; however, note that any hydrophone pair that includes hydrophone 2 has a 0.9-second time-delay shift at the end of the dataset, probably due to a pause in the recording of hydrophone 2. In figure 28 (dataset 2), the presence of four whales is evidenced by the existence of up to four consistent traces in many of the hydrophone pairs. In figure 29 (dataset 3), the presence of a single whale is seen. A speckling of false time delays is present. Most of these have been traced to the failure to reject surface path. In figure 30 (dataset 4), the presence of a single whale is seen. In figure 31 (dataset 5), the presence of a single whale is seen. In figure 32 (dataset 6), the presence of two loud whales and possibly more is seen.

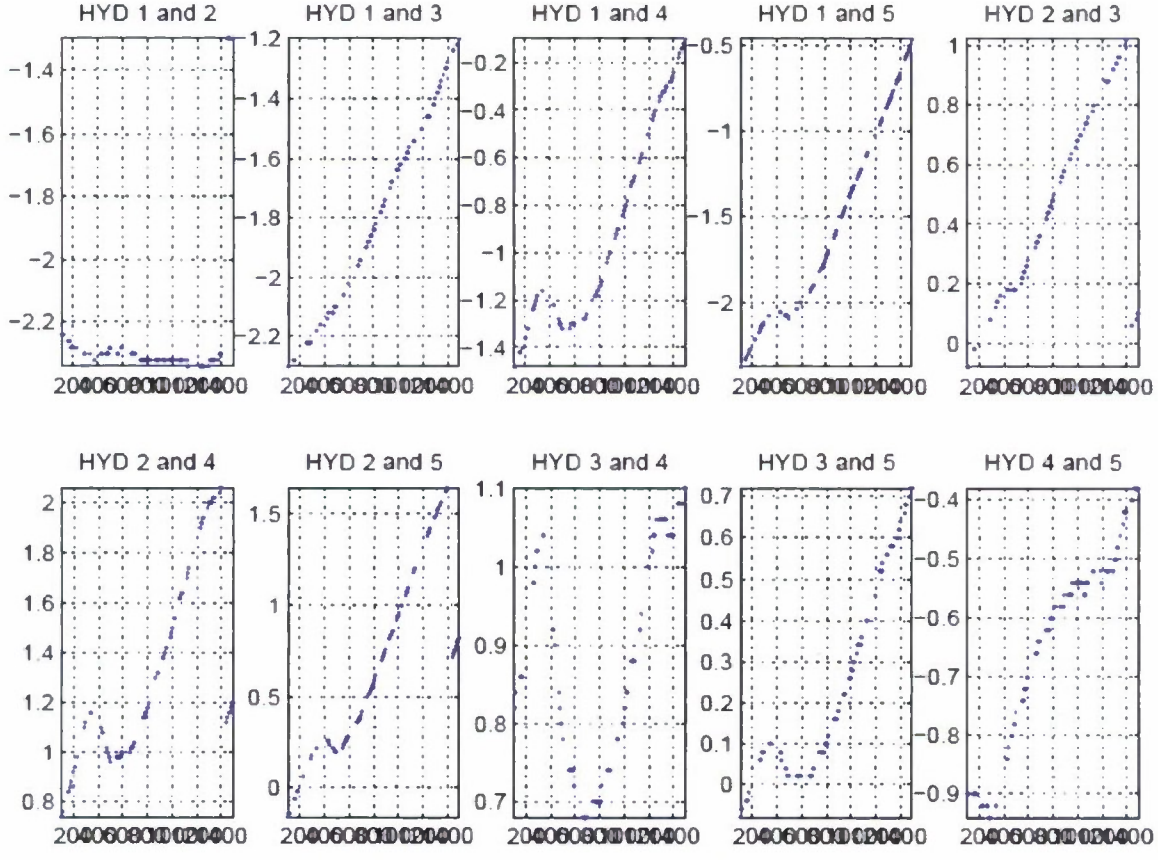

Figure 27. Plots of Inter-Sensor Time-Delay for Dataset 1
(Note the step-function at the end of the run in each sensor pair that includes sensor 2. This is apparently a time-shift due to a glitch in sensor 2.)

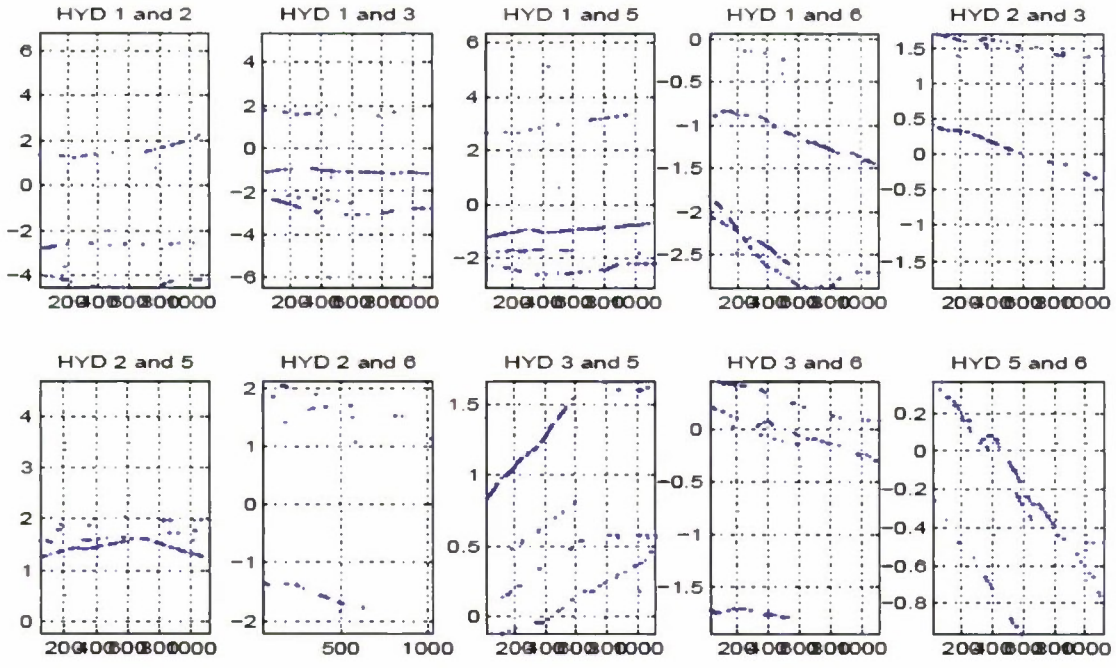

Figure 28. Plots of Inter-Sensor Time Delay for Dataset 2

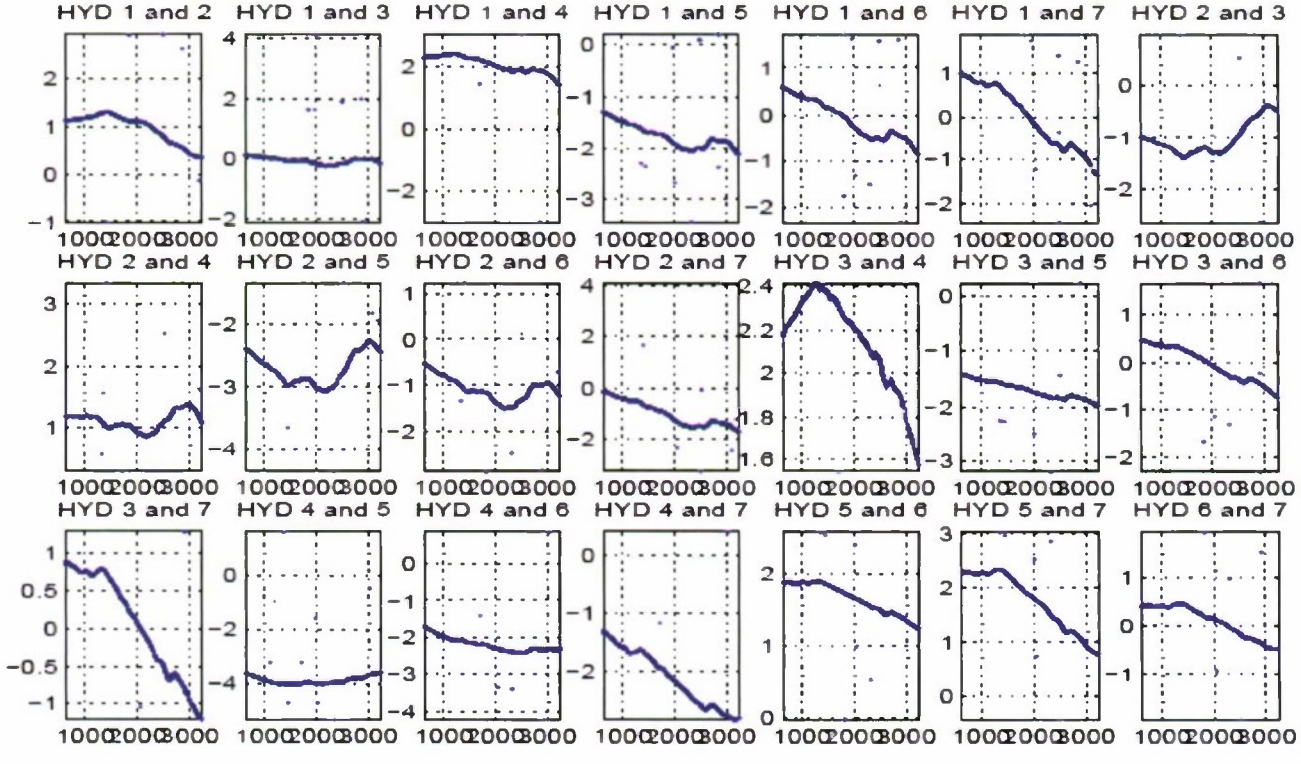

Figure 29. Plots of Inter-Sensor Time Delay for Dataset 3

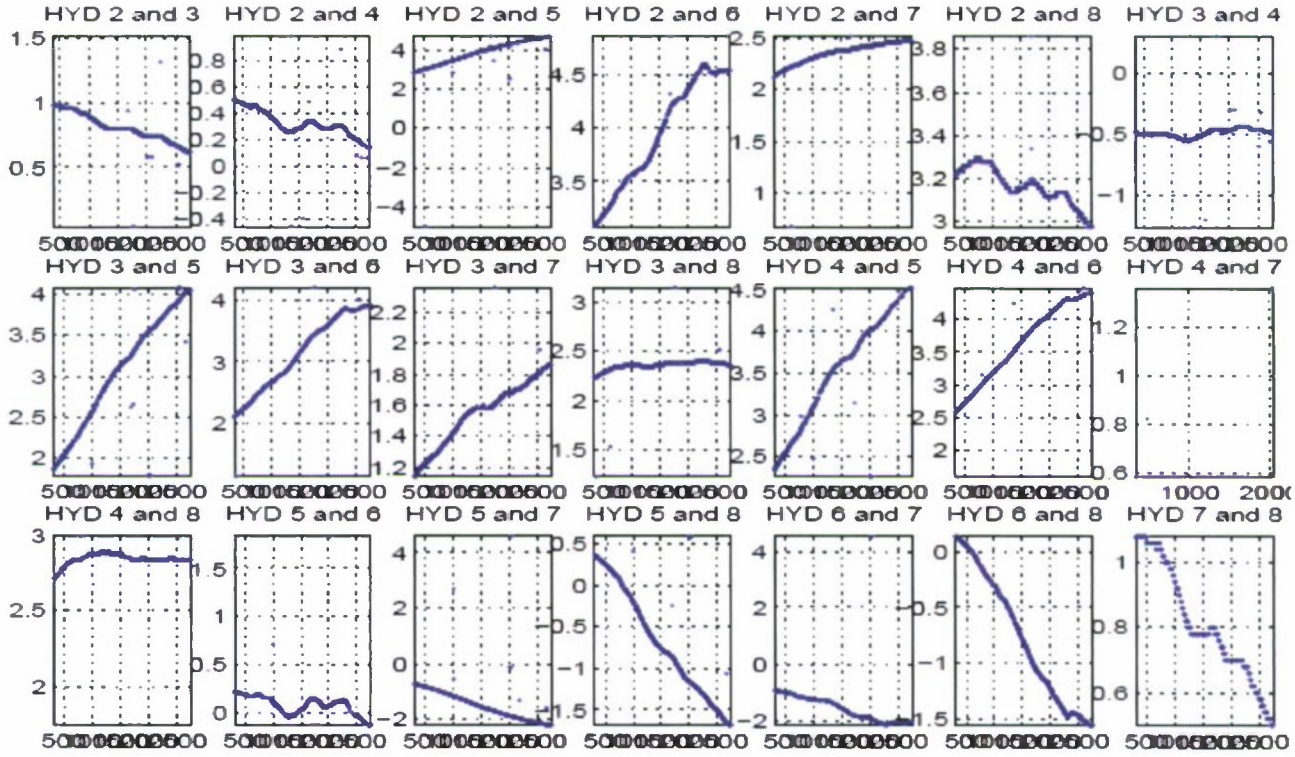

Figure 30. Plots of Inter-Sensor Time Delay for Dataset 4

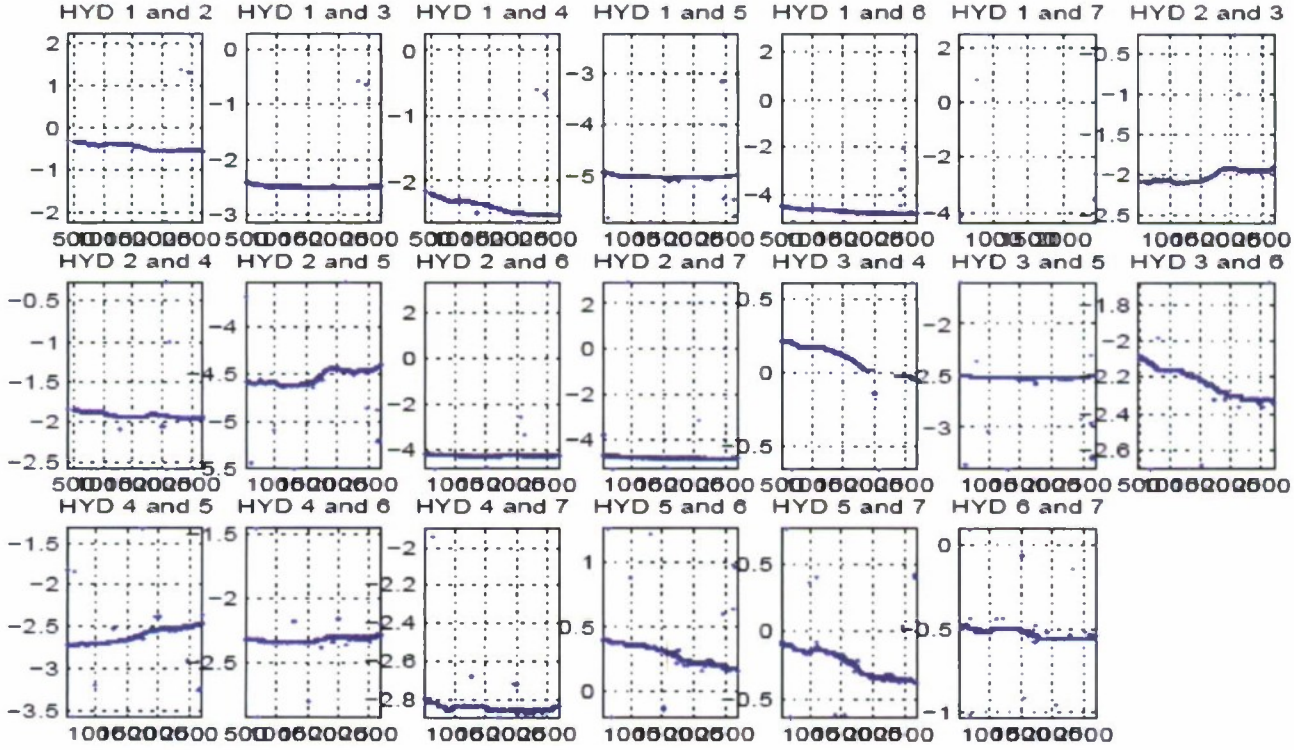

Figure 31. Plots of Inter-Sensor Time Delay for Dataset 5

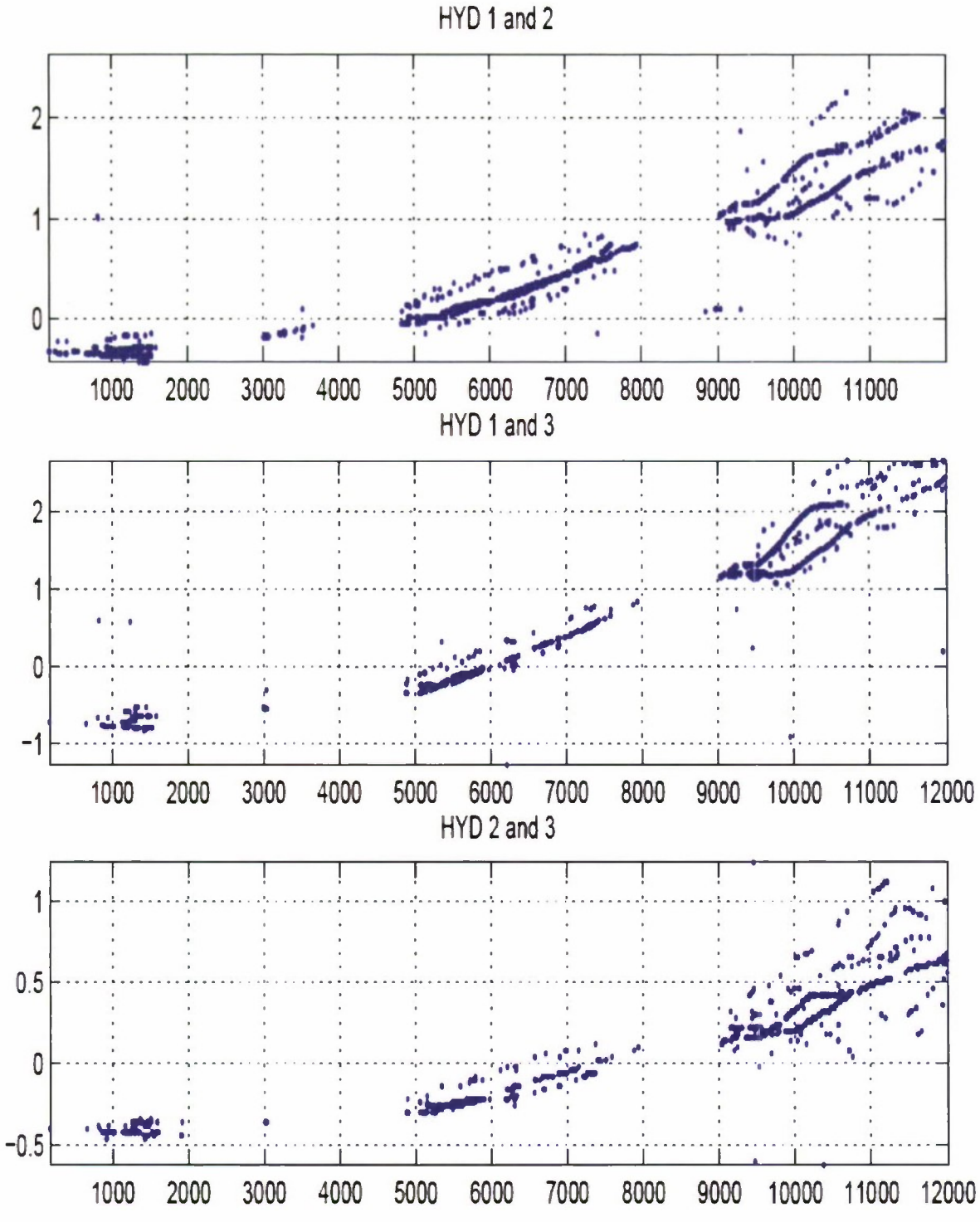

Figure 32. Plots of Inter-Sensor Time Delay for Dataset 6

5.4 LOCALIZATION RESULTS (CASE STUDIES)

An operator can manually initiate a localization solution at the hyperbola crossing points (section 4.4). Using a graphical user interface, solutions were initialized by clicking at the convergence of the hyperbolas in figure 6 using a starting depth of 500 meters, and the algorithm was allowed to converge. The solution converged at a depth of 780 meters, with a final Crammer Rao bound that had a standard deviation of 4.3, 3.4, and 19.0 meters in x, y, and z. The hyperbolas redrawn for the 780-meter depth are shown in figure 33. Note that a much tighter convergence is seen. The time-delay error values are also shown at the bottom of the figure. There is one error value for each time delay (and each hyperbola). Note that one error value is associated with a false time delay (the hyperbola that does not meet at the convergence point). Error absolute values of 0.01 second and below are typical for valid time delays.

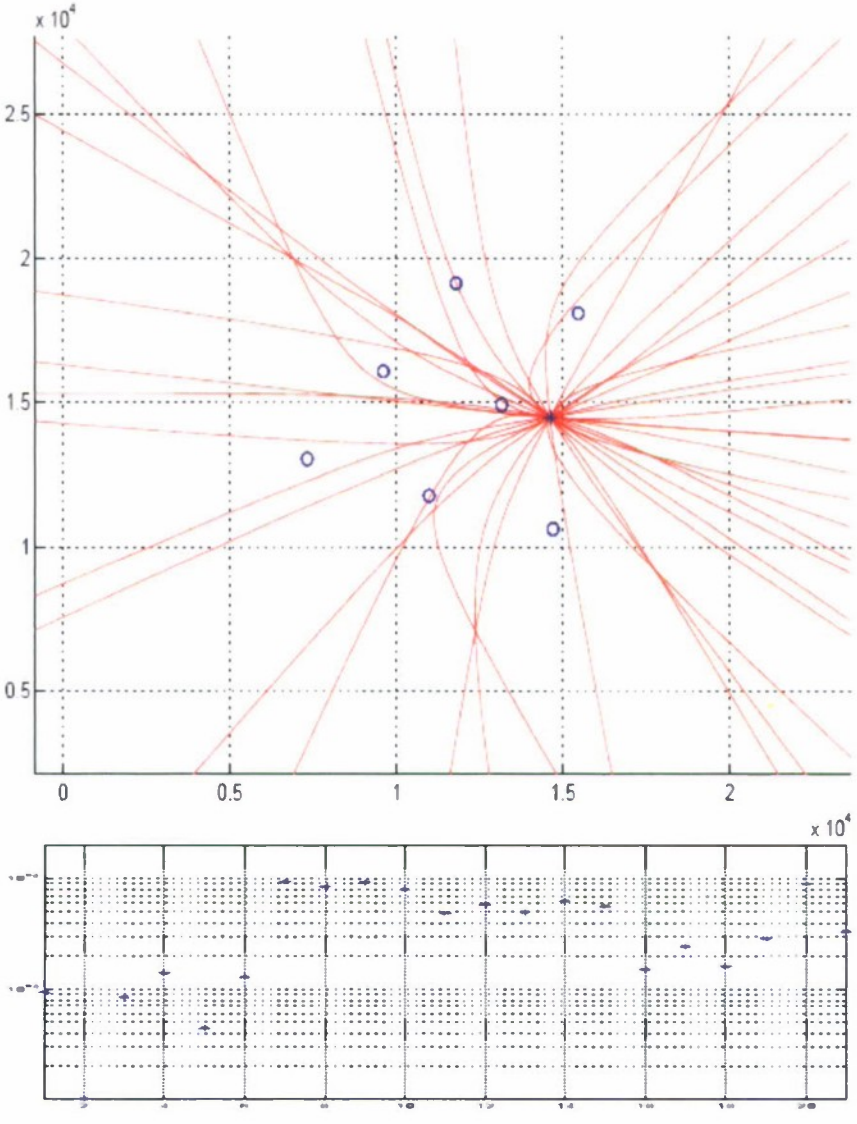

Figure 33. Top: Hyperbolas from Dataset 1 Showing One Whale; Bottom: Time-Delay Error Values After Convergence

5.5 SINGLE WHALES

Only one whale was present in datasets 1, 3, 4, and 5. Localization error ellipses for dataset 1 and a corresponding depth profile are shown in figure 34.

Localization error ellipses for data set 3 and corresponding depth profiles are shown in figure 35 and 36. In the depth profiles, the depth component of the error bound is drawn as a cyan envelope. Note that outlier solutions are usually associated with a widening of the error bounds. These occur typically because there is an insufficient number of sensors to provide good position accuracy or there is poor sensor spatial distribution. This insufficiency is caused by the unavailability of any nearby sensors that provide adequate depth observability. As with the error ellipses in the two-dimensional plane, this provides a reliable "warning" that solutions may have significant error.

Localization error ellipses and corresponding depth profiles for dataset 4 are shown in figures 37 and 38.

Localization error ellipses for dataset 5 and a corresponding depth profile are shown in figure 39, and figure 40 shows hyperbola plots for dataset 5.

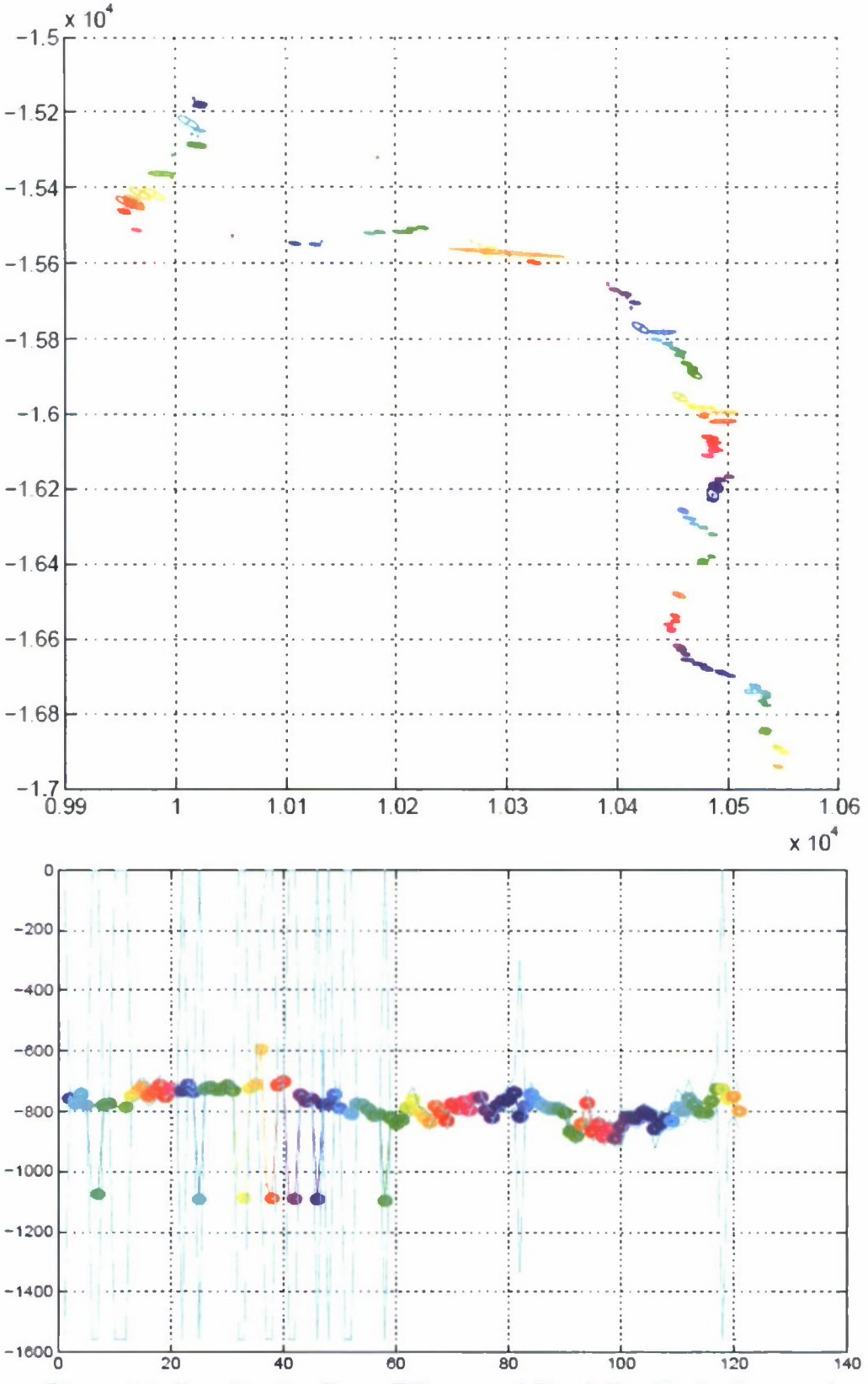

Figure 34. Localization Error Ellipses and Depth Profile for Dataset 1
(Y-axis: depth; X-axis: time update number. The cyan "envelope" is the depth CR bound.)

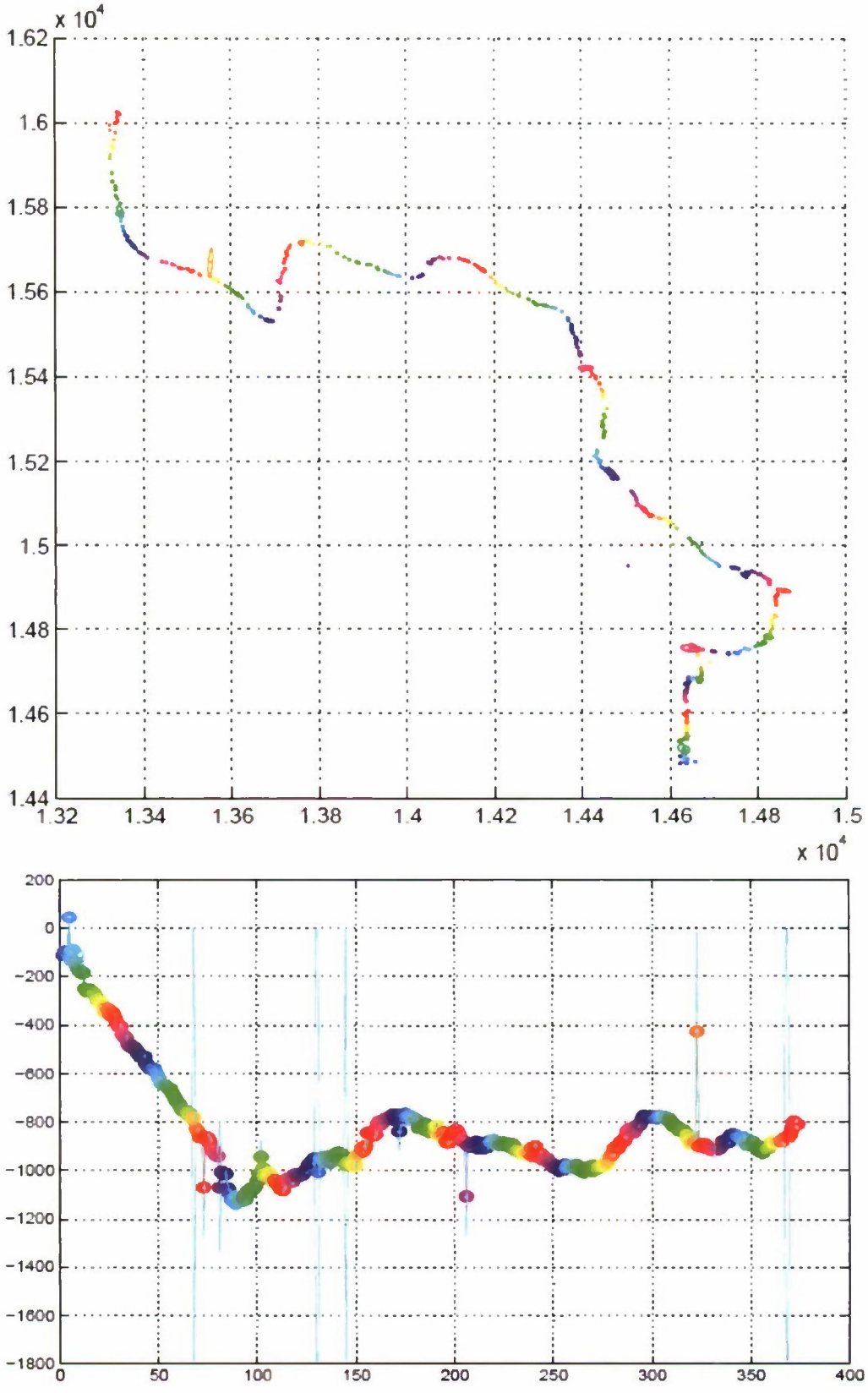

Figure 35. Localization Error Ellipses and Depth Profile for Dataset 3

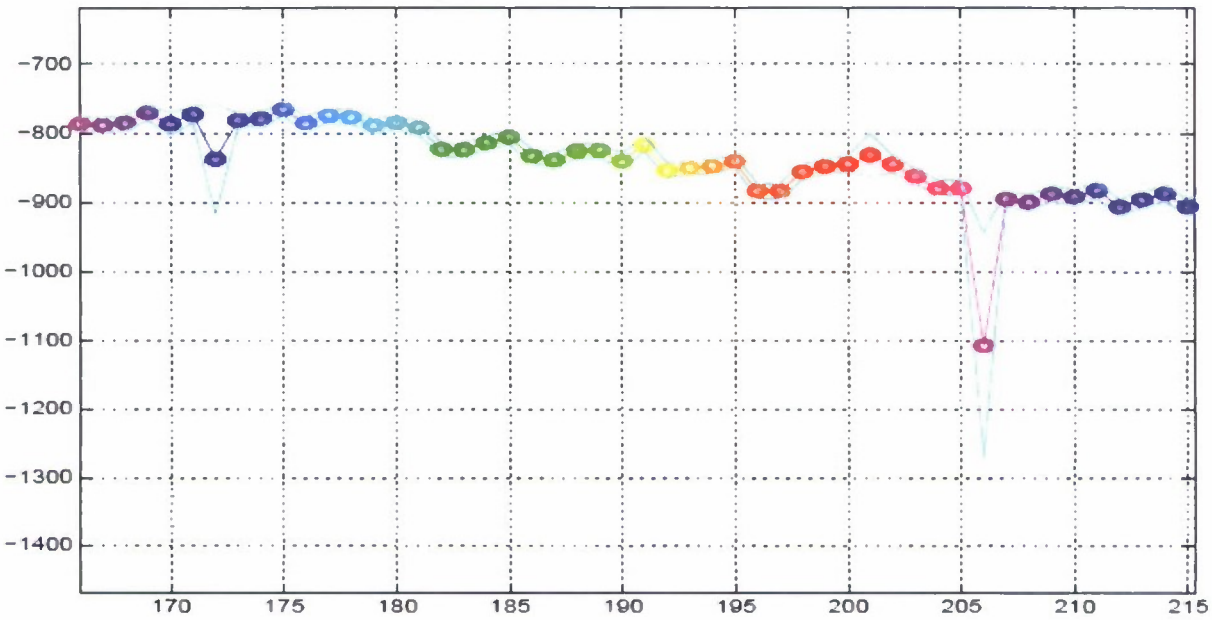

Figure 36. Blowup of Depth Profile for Dataset 3 Corresponding to Figure 35

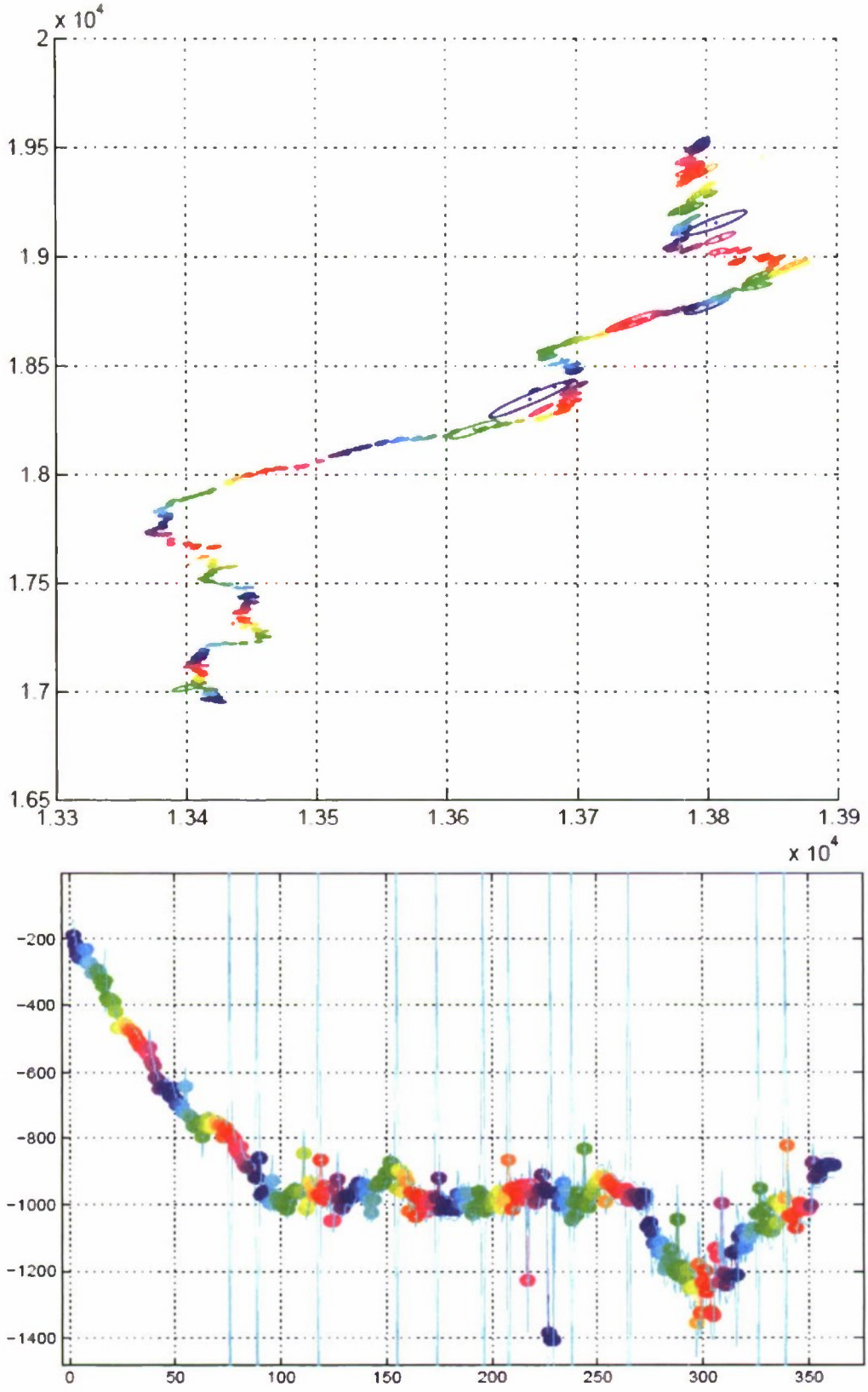

Figure 37. Localization Error Ellipses for Dataset 4

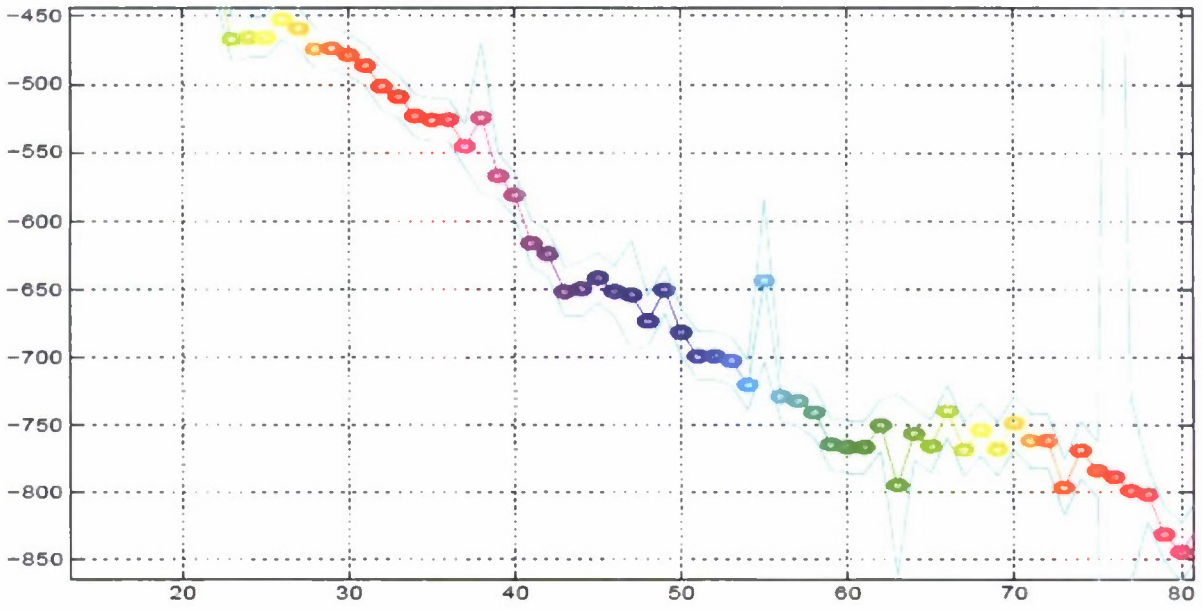

Figure 38. Blowup of Depth Profile for Dataset 4 Corresponding to Figure 37

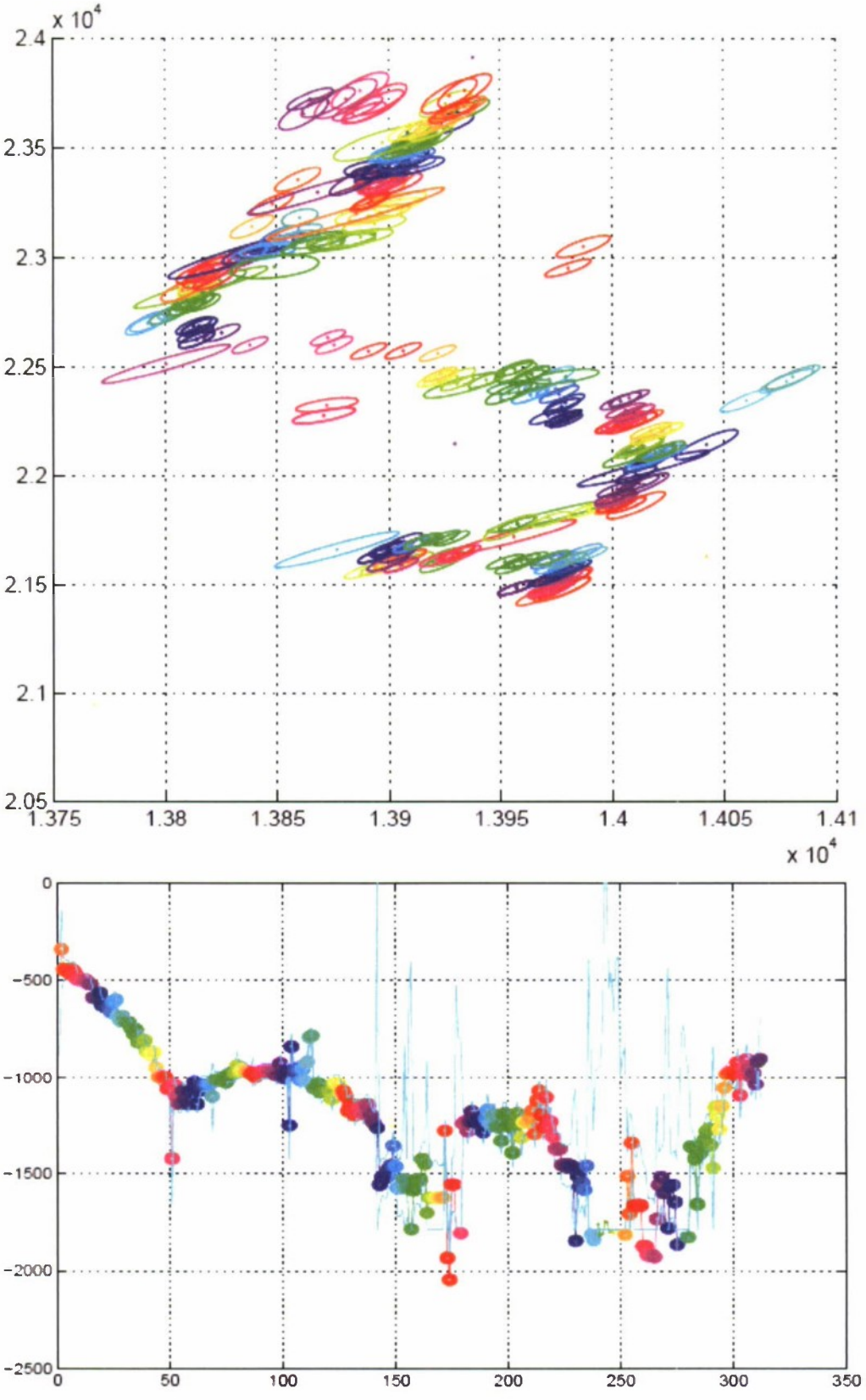

Figure 39. Localization Error Ellipses and Depth Profile for Dataset 5

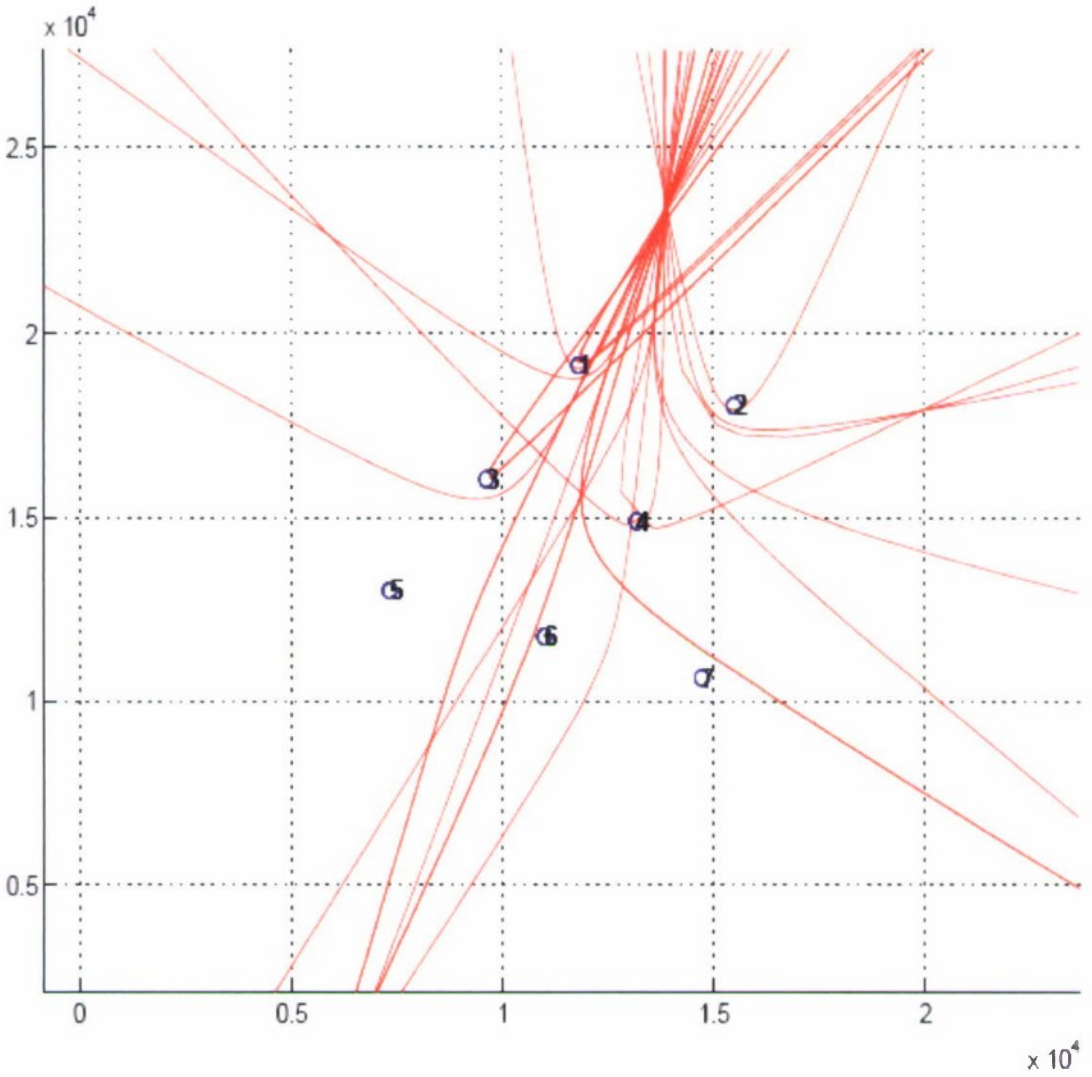

Figure 40. Hyperbola Plots for Dataset 5

5.6 MULTIPLE WHALES

As many as four whales were present in dataset 2. The locations of manually initialized location solutions throughout the data record are shown in figure 41. An example of a hyperbola plot was shown earlier in figure 7, with clear evidence of four simultaneously vocalizing whales. The time-delays have been automatically grouped into inter-related sets, as described in section 4.2. The hyperbolas have been color-coded according to group membership. The error ellipses for each of the four whales are shown in figure 42. The depth profiles are shown in figure 43. Depth was difficult to estimate. Contributing factors include lower SNR for the clicks and the presence of multiple whales, which may complicate the click-train separation process.

A problem that can occur with multiple whales is that time-delay estimates can overlap. For instance in figure 7, the whale positioned farthest to the right (dark blue) has two hyperbolas that pass near solutions belonging to other whales. When solving for the positions of all of these whales using the open-loop membership function, positional errors will be introduced because those time delays belonging to other whales will not be rejected by the membership function (section 4.5). This case was processed using the closed-loop algorithm. Solutions were placed manually by clicking at the convergence point of each solution in figure 7, which was drawn using an average depth of 500 meters. At each of the four convergence points, 10 initial solutions were added in depth increments of 160 meters spanning the water column. As the algorithm iterated, solutions were removed according to the pruning rules (section 4.5.2). In 10 attempts, each time clicking at random locations near the four convergence points, the algorithm correctly reported four surviving solutions with the same final locations. The results are shown in figure 44. Blowups are shown in figures and 45 through 48, where the tight convergence provides good separation from the nearby hyperbolas.

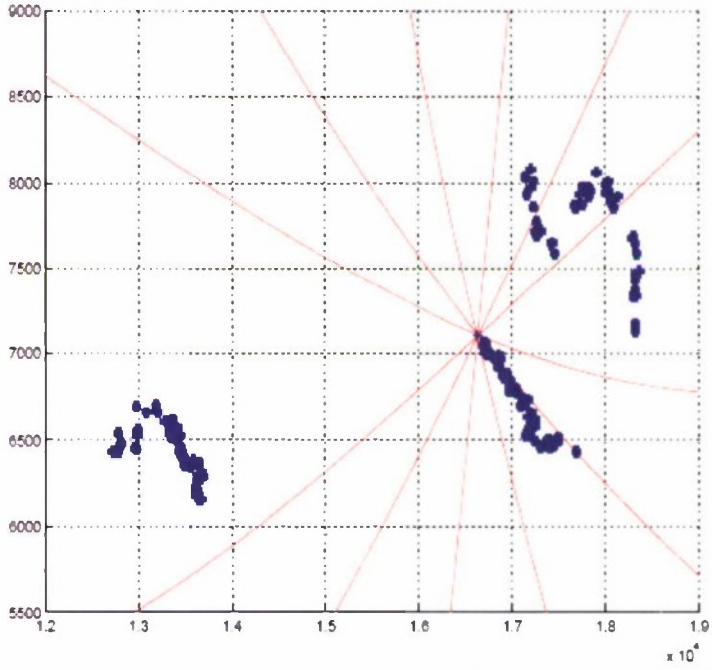

Figure 41. History of Manual Solution Initialization Points Over 20 Minutes (Blue Dots) (The converging hyperbolas are associated with the solution of the current frame.)

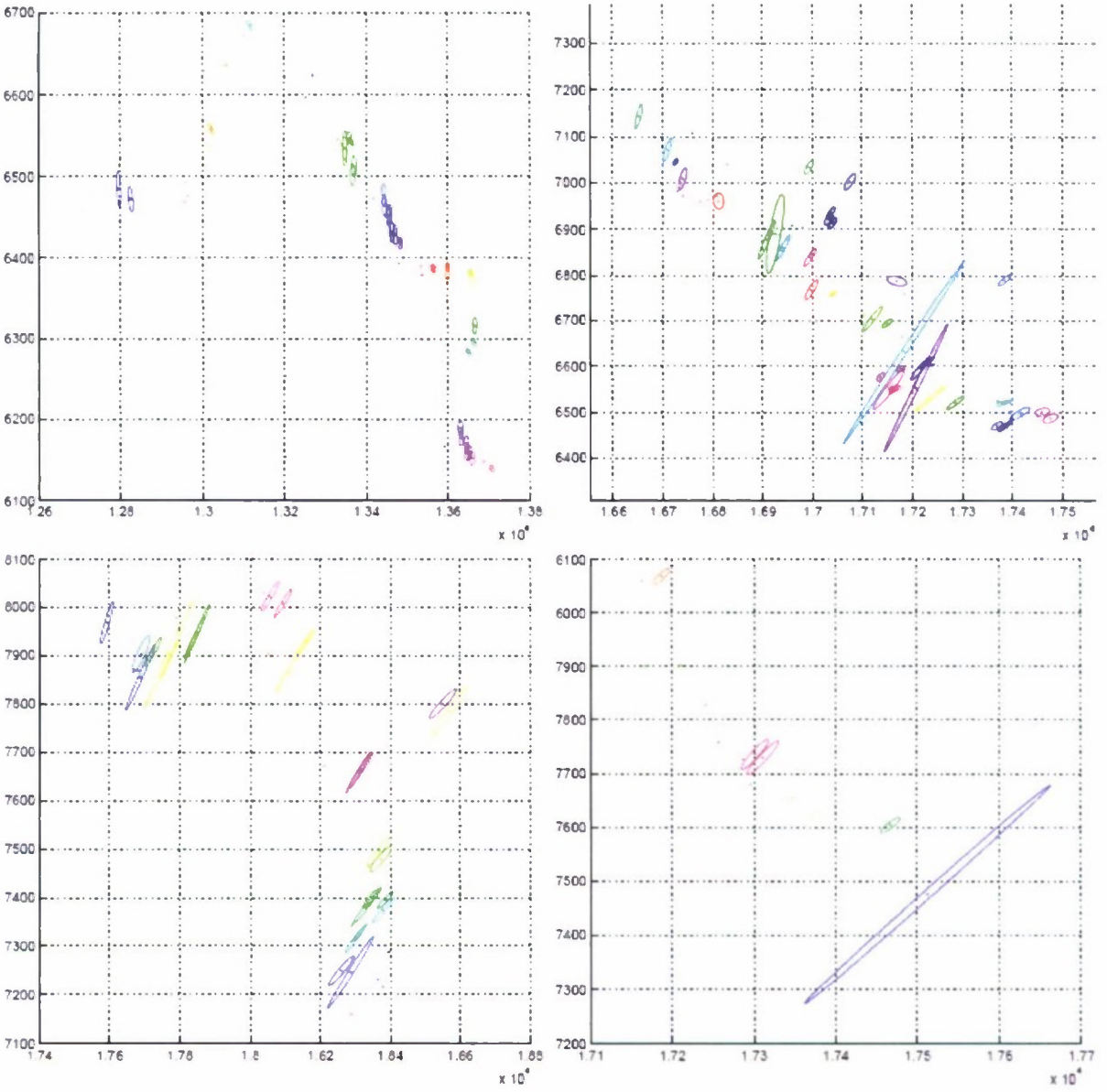

Figure 42. Localization Error Ellipses for Dataset 2, Whales 1-4

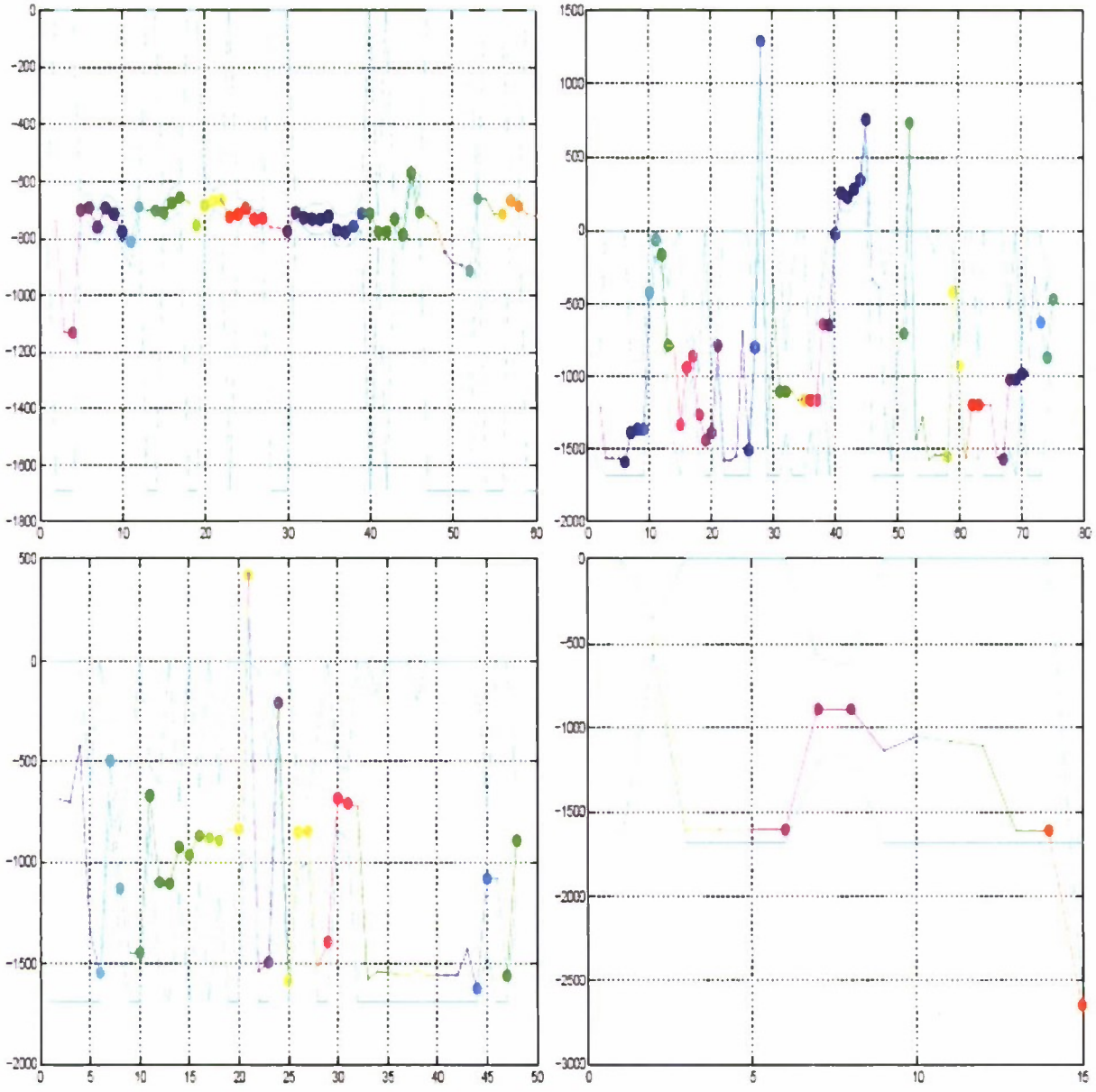

Figure 43. Depth Profiles for Dataset 2, Whales 1-4

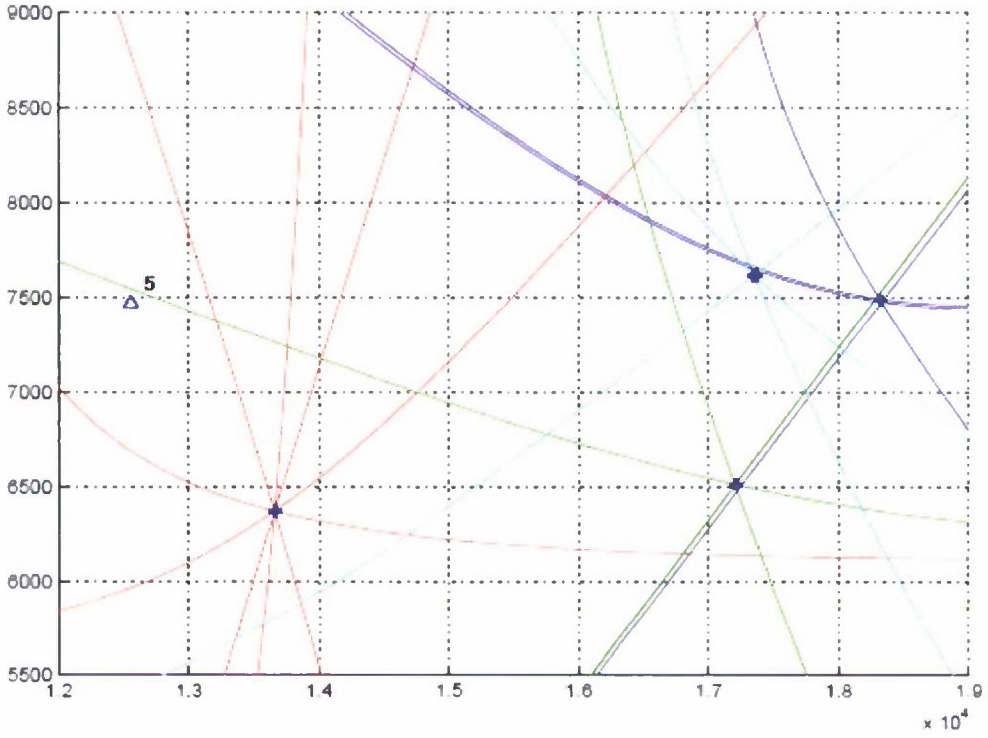

Figure 44. Example of Figure 7 After Convergence of the Closed-Loop Algorithm (The hyperbolas are color coded according to membership to the four remaining solutions. The final solution depth was used when drawing the hyperbolas of each solution. A cross (+) is drawn at the x-y location of each solution.)

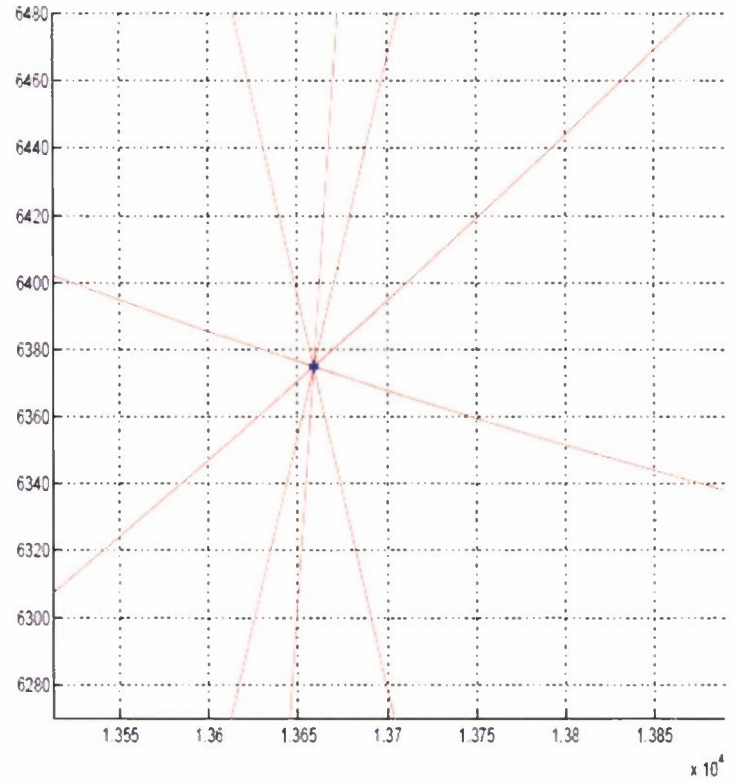

Figure 45. Expanded View of Solution 1 in Figure 44

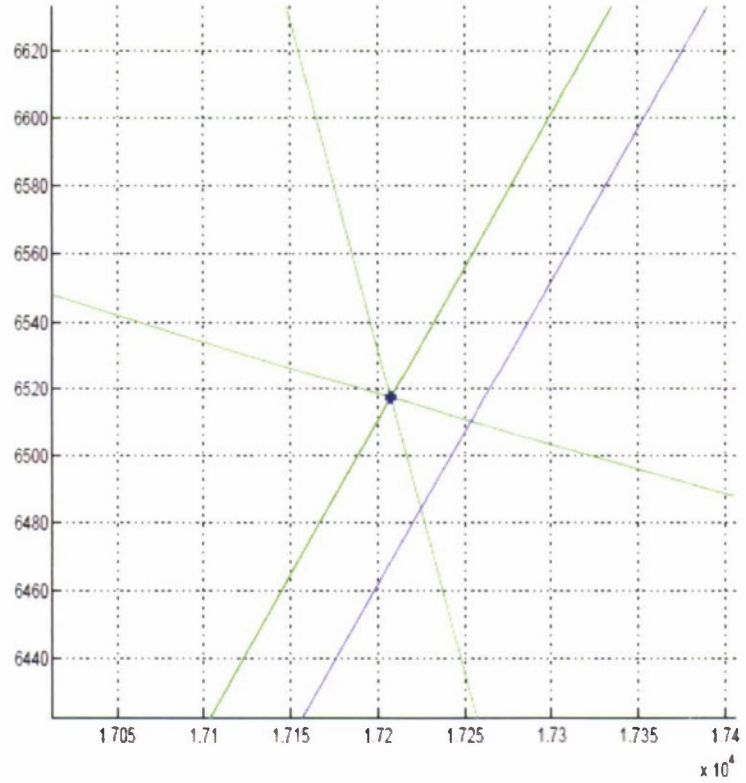

Figure 46. Expanded View of Solution 2 in Figure 44

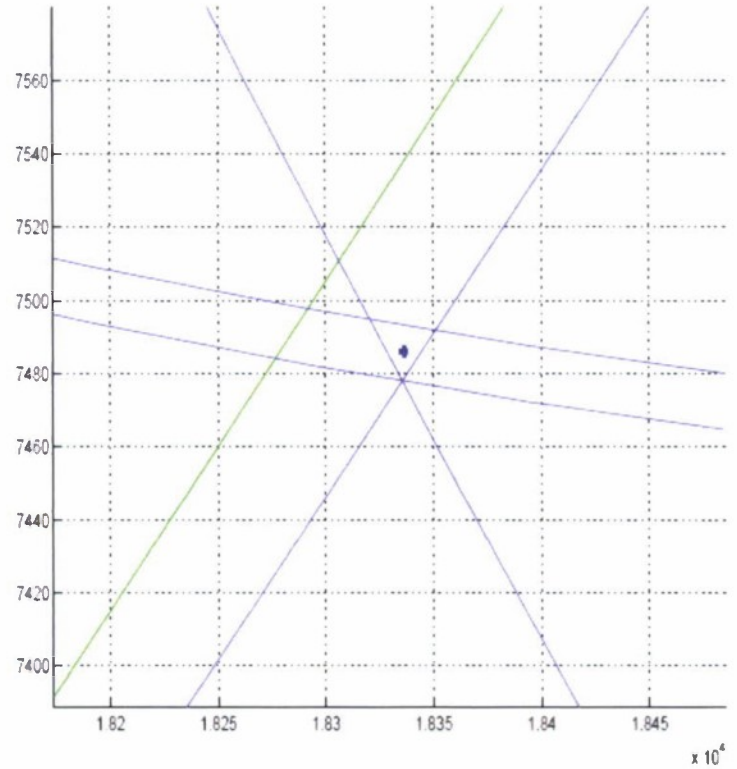

Figure 47. Expanded View of Solution 3 in Figure 44

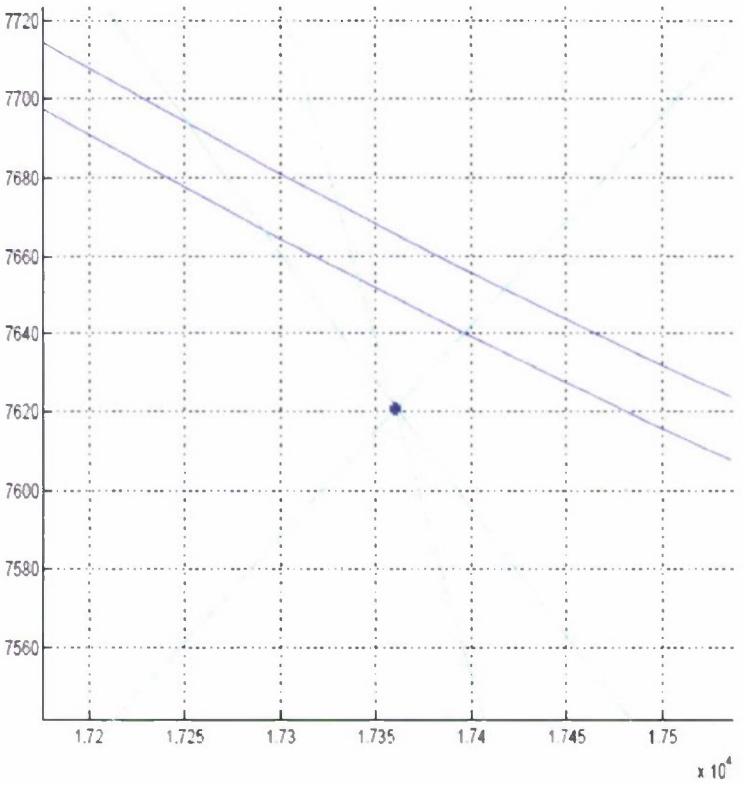

Figure 48. Expanded View of Solution 4 in Figure 44

6. SUMMARY AND CONCLUSIONS

This technical report has presented results in three main areas. First, a sperm whale click separation algorithm that groups clicks into click-trains and solves the computational problem of a previous approach. The algorithm has been demonstrated on benign and demanding real-world examples. How an operator might examine and validate the algorithm's results has been discussed in detail.

Second, two ways that single-sensor information can be used to eliminate multipath have been discussed: (1) features extracted from individual clicks to classify them as direct path or reverberation, and (2) cross-correlation of click-trains formed by the separation algorithm to determine if one click-train is a time-delayed copy of the other.

Third, two versions of a localization algorithm have been demonstrated. Both versions handle multiple simultaneous whales and attempt to resolve "disputes" in which more than one localization solution competes for a given time-delay estimate. One version is "open loop" and assigns time delays to solutions based on a fixed weight that is based on time-delay error. The other version is iterative, or "closed loop," because it recursively updates the weights based on the error variance of each solution. Real results are shown where time delays are correctly assigned to four whales simultaneously vocalizing.

REFERENCES

1. M. van der Schaar, E. Delory, A. Catal, and M. Andr, "Neural Network-Based Sperm Whale Click Classification," *Journal of Marine Biology Association U.K*, vol. 87, 2007, pp. 35–38.

2. T. Ura, R. Bahl, M. Sakata, J. Kojima, T. Fukuchi, J. Ura, K. Mori, T. Nakatani, Y. Nose, and H. Sugimatsu, "Development of AUV-Based System for Acoustic Tracking of Diving Sperm Whales," *Proceedings of OCEANS 2004*, Kobe, Japan, November 2004, pp. 2302–2307.

3. R. Bahl and T. Ura, "Automatic Real-Time Segregation and Classification of Multiple Vocalizing Sperm Whales," *Seisan-Kenkyu Bimonthly Journal of IIS*, University of Tokyo, vol. 55, May 2003, pp. 61–64.

4. P. M. Baggenstoss, "Joint Localization and Separation of Sperm Whale Clicks," *Canadian Journal of Acoustics*, March 2008, pp. 125–131.

5. R. E. Bellman, *Adaptive Control Processes*, Princeton University Press, Princeton, NJ, 1961.

6. R. Morrissey, J. Ward, N. DiMarzio, S. Jarvis, and D. Moretti, "Passive Acoustic Detection and Localization of Sperm Whales (Physeter Macrocephalus) in the Tongue of the Ocean," *Applied Acoustics*, vol. 67, no. 11–12, 2006, pp. 1091–1105.

7. D. M. Titterington, A. F. M. Smith, and U. E. Makov, *Statistical Analysis of Finite Mixture Distributions*, John Wiley & Sons, New York, 1985.

8. S. Jarvis and D. Moretti, "Passive Detection and Localization of Transient Signals from Marine Mammals Using Widely Spaced Bottom Mounted Hydrophones in Open Ocean Environments," *International Workshop on the Application of Passive Acoustics in Fisheries*, Naval Undersea Warfare Center Division, Newport, RI, 2002.

9. P. White, T. Leighton, D. Finfer, C. Powles, and O. Baumann, "Localisation of Sperm Whales Using Bottom-Mounted Sensors," *Applied Acoustics*, vol. 67, 2006, pp. 1074–1090.

10. C. O. Tiemann and M. B. Porter, "Automated Model-Based Localization of Sperm Whale Clicks," *OCEANS 2003 Proceedings*, vol. 2, September 2004, pp. 821– 827.

11. S. Kay, *Modern Spectral Estimation: Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1988.

INITIAL DISTRIBUTION LIST

Addressee	No. of Copies
Office of Naval Research (ONR-321—John Tague, ONR-321US—M Vaccaro, ONR-322—James Eckman)	3
Space and Naval Warfare Systems Center—Pacific (Attn: Stephen Martin)	1
Woods Hole Oceanographic Institution (Attn: Peter Tyack)*	1
University of Massachusetts (Attn: John Buck)*	1
Oregon State University (Attn: David Mellinger)*	1
University of St. Andrews (Attn: Len Thomas, Tiago Marques, David Borchers, Catriona Harris, Danielle Harris)*	5
Science Applications International Corporation (Attn: James Yosco)*	1
Defense Technical Information Center	2

*PDF file distribution via email.