WHOI-92-04

Characterizing Acoustic Features
of Marine Animal Sounds

by

Kurt M. Fristrup and William A. Watkins

Woods Hole Oceanographic Institution
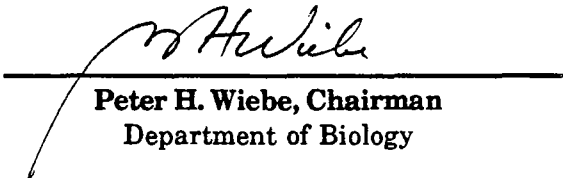Woods Hole, Massachusetts 02543

January 1992

Technical Report

Approved for public release; distribution unlimited.

Approved for Distribution:

Peter H. Wiebe, Chairman
Department of Biology

Abstract

Software tools were designed to characterize the acoustic features of marine animal sounds. These have resulted in a set of calculated measurements that summarize particular aspects of sound sequences. The specificity of these measurements was enhanced by adjusting calculations to compensate for ambient noise. The sound measures included statistics for Aggregate Bandwidth, Intensity, Duration, Amplitude Modulation, Frequency Modulation, Short-term Bandwidth, Center Frequency, and Amplitude Frequency Interaction. The efficacy of noise compensation was tested for each statistic. Then, the sound measures were tested on a subset of 200 sequences of marine animal sounds, including sequences from 20 species: six baleen whales, 13 toothed species, and one seal. The statistics were reviewed for each species and a graphical comparison of all species was generated using principal components analysis. Preliminary results confirm that such sounds can be classified by means of relatively simple statistical algorithms, and we are encouraged to continue toward a system for automatic classification of marine animal sounds.

1 Introduction

Marine animals produce a remarkable variety of sounds (Watkins and Wartzok 1985). A primary goal of the bioacoustic program at the Woods Hole Oceanographic Institution (WHOI) has been to parse this variation into sensible classes of signals. Marine mammal sounds in particular contain distinctive features associated with species (op. cit.), individual identity (Caldwell, Caldwell and Tyack 1990), and certain behaviors . These features have never been examined in a broad context, comparing the sounds of a wide variety of species. Do the differences in these sound features remain distinctive as the scope of comparison widens? With our own ears, we can often distinguish acoustic features that appear to be species-specific, and sometimes features unique to individual animals; can we specify numerical algorithms that objectively recognize these distinctions?

The logistic requirements for addressing these questions have been formidable. To quantify the interspecific and intraspecific variability in marine animal sounds, a large number of sounds must be analyzed for each individual or species to be differentiated. Many biological and environmental attributes potentially explain acoustic variability. Therefore, these numeric results had to be referenced: species, population, group, social context, behavior, activity, individual identity, sex, reproductive situation, age, season, geographic location, water depth, and sound propagation. Thus, a necessary resource for such acoustic distinctions has been a system for integrating the sound sequences with associated biological and environmental data.

The SOUND database system organized for marine animal sounds (Watkins, Fristrup, and Daher 1991) has provided this resource. The databases and associated files contain thousands of digitized sound segments spanning more than seventy species recorded from all the world's oceans. The database describes the time, geographic location, recording conditions, identity of the animal(s) producing the sounds, the behavioral observations

associated with sound production, etc. These SOUND databases represent years of work by several people, and the analyses reported here depend on the availability of such. In turn, these analyses complemented and extended the capabilities of the database. New relational database structures have been implemented to permit flexible and convenient integration of these statistical results with the biological and environmental information about the sounds.

The quantification of time-frequency characters of the animal sounds for these analytic distinctions has had no precedent on this scale. No prior work has dealt with so many species and such a variety of repertoires from individual animals. The WHOI studies of marine animal acoustics, which have continued since William E. Schevill's work in the late 1940's, have provided the heuristic basis for these statistical decisions. We have learned to utilize many different acoustic features to describe and diagnose sounds. As a first step toward the development of an automatic, non-subjective system for separating the different animal sound sequences, we have devised statistical measures to recognize familiar acoustic features.

This report describes the numerical procedures that have been used, and it demonstrates their effectiveness with a trial set of 200 digitized sequences of marine mammal sounds. These preliminary results suggest that the gross acoustic features we analyzed can be useful indicators of species identity, and that with refinement they might provide the basis for finer distinctions.

2 The Statistics

Our statistical estimation techniques were based on our experience with the marine animal sounds in our tape library. We were guided by the following criteria:

- Each statistic was designed to emphasize particular parameters of animal sounds that we recognized as important for distinguishing species.

- Each statistic had to be insensitive to sound artifacts introduced by propagation in the ocean (multipath, fading, frequency-dependent attenuation, etc.).

- Most statistics needed to be relatively insensitive to noise and assumed a minimum of 15 dB signal/noise.

- Most statistics had to be insusceptible to the shape (relative frequency emphases) of the ambient noise power spectra.

- Most statistics needed to be related to obvious features in the time-frequency analysis displays of these sounds (duration, frequency range, etc.) – so we could recognize the effectiveness of the statistics in making the discriminations.

These criteria reflected our interest in discriminating among the animal sounds rather than making selections that were largely controlled by differences in the ambient backgrounds. The choice of criteria did not take into consideration changes to the sounds contributed by the orientation and movements of sound sources. A number of other effects also have not been addressed, including means for dealing with frequency-dependent attenuation. Statistics that obviously would be sensitive to distortion of phase information have been avoided in these analyses.

The basic unit of data used for our feature extraction programs was one FFT (Fast Fourier Transform) block. For most files, this was 256 sample points, but for very short files

3

(low sampling rates) the FFT size was decreased to obtain no fewer than 16 blocks. Adjacent blocks did not overlap. A constant was subtracted from each sample point such that the mean for each block became zero, and the data were then tapered with a Hamming window. These choices eliminated explicit correlation between adjacent blocks and smoothed the resulting power spectra; the cost of this was a reduction in the degrees of freedom for our analyses.

The noise compensation technique begins by estimating the power spectrum of sounds that were present throughout the sound cut. To identify blocks of data containing only noise energy, intensity measures were computed for up to 600 blocks of data distributed evenly through a sound cut. Blocks were sorted by intensity, and the blocks between the fifth and tenth percentiles in level were used to form a noise power spectrum. We eliminated the bottom five percent to avoid using atypically quiet sections (tape dropout, etc.). During subsequent processing of these data, a multiple (currently 6.67x) of this noise spectrum was subtracted from each block's power spectrum (negative values set to zero). All spectral statistics were computed from this reduced power spectrum. To obtain the amplitude estimate for the block, the adjusted spectrum values were summed; this indirect method of computing amplitude, which exploits Parseval's relation (Oppenheim and Schafer 1989, p. 574), prevents loud noise components from dominating the amplitude statistics.

2.1 Abbreviations in Statistical Formulae

- t_i, time in seconds: the interval from the beginning of the sound cut to the beginning of the ith FFT block.

- ss_i, relative intensity in arbitrary units: sum of the adjusted power spectrum values for the ith block.

- $mm_i = min(ss_{i-1}, ss_i)$: the smaller of two adjacent ss values.

4

- F_{50} in Hertz: the frequency that bisects the area under the power spectrum density.

- N_{50} in Hertz: the minimum number of frequency bins required to accumulate fifty percent of the total signal energy.

- F_{75} in Hertz: the highest frequency encountered when calculating N_{50}.

- F_{25} in Hertz: the lowest frequency encountered when calculating N_{50}.

Considerable use of symbols could not be avoided, but wherever possible we have used descriptive terms for ease of interpretation.

Table 1: Summary of Statistics

	Equation	Page
Duration: Results in Table 2		
Total Duration	2	8
Sound Concentration	3	8
Amplitude Modulation: Results in Table 3		
Amplitude Mean	4	9
Amplitude Standard Deviation	5	9
Attack Fraction	6	9
Attack Proportion	7	10
Amplitude Skewness	8	10
Frequency Modulation: Results in Table 4		
Upsweep Mean	9	10
Upsweep Fraction	10	11
Upsweep Proportion	11	11
Time Frequency Correlation	12	11
Time Upsweep Correlation	13	12
Short-term Bandwidth: Results in Table 5		
Short-term Bandwidth Mean	14	12
Short-term Spectral Concentration	15	13
Short-term Spectral Asymmetry	16	13
Aggregate Bandwidth: Results in Table 6		
Total Upper Frequency - F_{75}		5, 13
Total Lower Frequency - F_{25}		5, 13
Total Spectrum Concentration - N_{50}		5, 13
Modal Upper Frequency - F_{75}		5, 13
Modal Lower Frequency - F_{25}		5, 13
Modal Spectrum Concentration - N_{50}		5, 13
Center Frequency: Results in Table 7		
Median Frequency Mean	17	14
Total Spectrum Median Frequency		
Modal Spectrum Median Frequency		
Amplitude Frequency Interaction: Results in table 8		
Amplitude Frequency Correlation	18	14
Amplitude Upsweep Correlation	19	14

Intensity measures were computed for each block of data. These values were used as weights to find the "center" of the sound, defined as the weighted average of the time values (eq. 1). This statistic is only used as a reference point for subsequent calculations of Sound Duration (eq. 2) and Amplitude Skewness (eq. 8).

- Signal Center: weighted mean of t, ss as weights.

$$\bar{t} = \frac{\sum_{i=0}^{N} ss_i t_i}{\sum_{i=0}^{N} ss_i} \tag{1}$$

2.2 Sound Duration

The Sound Duration was computed by Equation 2. It yielded a gross estimate of total duration, including any intervals of silence between sound elements.

- Sound Duration: weighted standard deviation of t, ss as weights.

$$4\sigma_t = 4\sqrt{\frac{\sum_{i=0}^{N} ss_i t_i^2}{\sum_{i=0}^{N} ss_i} - \bar{t}^2} \tag{2}$$

The Sound Concentration was computed by Equation 3. It yielded an estimate of duration that would result if the sound were "packed" such that all silent sections were removed. It responds only to the relative amplitudes of different blocks, and it is insensitive to their ordering in the sound.

- Sound Concentration: equivalent statistical bandwidth of the amplitude values.

$$\frac{\left(\sum_{i=0}^{N} ss_i \right)^2}{\sum_{i=0}^{N} ss_i^2} \tag{3}$$

The ratio of these two duration estimates can be used to measure duty cycle.

2.3 Amplitude Modulation

A reference value was computed for the average level of the sound: the Amplitude Mean was computed by Equation 4.

- Amplitude Mean: average ss value.

$$\overline{ss} = \frac{\sum_{i=0}^{N} ss_i}{N} \tag{4}$$

The Amplitude Standard Deviation (Eq. 5) can be used to measure the average magnitude of amplitude modulation. However, Amplitude Mean and Amplitude Standard Deviation are scaled arbitrarily by idiosyncrasies of the digitizing process. To form a useful diagnostic, they must be used together to form a scale-independent measure like the coefficient of variation (Eq. 5/ Eq. 4).

- Amplitude Standard Deviation: standard deviation of ss values.

$$\sigma_{ss} = \sqrt{\frac{\sum_{i=0}^{N} ss_i^2}{N} - \left(\frac{\sum_{i=0}^{N} ss_i}{N}\right)^2} \tag{5}$$

The occurrence and magnitude of sections of sound with increasing and decreasing amplitudes was measured by assessing the fraction of blocks in which a subsequent block had a larger (or lower) amplitude than the current block. This Attack Fraction was computed by Equation 6.

- Attack Fraction: the fraction of blocks in which subsequent block has a larger ss value than the current block.

$$\frac{\sum_{ss_{i+1} > ss_i} 1}{N} \tag{6}$$

9

The proportion of average changes in amplitude to the sum of the average increases and decreases in amplitude values was calculated and called the Attack Proportion. These statistics (Eq. 6 and Eq. 7) were similar in function to the "attack" and "decay" terms used to refer to the initial and terminal amplitude modulation of individual notes in music. The attack proportion was computed by Equation 7.

- Attack Proportion: the proportion of average increase in ss values relative to the sum of average increases and decreases in ss values.

$$\frac{\sum_{\Delta ss_i > 0} \Delta ss_i}{\sum_{\Delta ss_i > 0} \Delta ss_i - \sum_{\Delta ss_i < 0} \Delta ss_i} \tag{7}$$

Gross asymmetry in the amplitude modulation relative to the "center" of the sound was weighted by a function of time and amplitude. This Amplitude Skewness was computed by Equation 8.

- Amplitude Skewness: weighted skewness of t, ss as the weights.

$$\frac{\frac{\sum_{i=0}^{N} ss_i t_i^3}{\sum_{i=0}^{N} ss_i} - 3\bar{t}\frac{\sum_{i=0}^{N} ss_i t_i^2}{\sum_{i=0}^{N} ss_i} + 2\bar{t}^3}{\sigma_t^3} \tag{8}$$

2.4 Frequency Modulation

The frequency modulation of a sound was expressed by differences between median frequencies of adjacent power spectrum estimates. The average upsweep trends (downsweep is a negative upsweep) for entire sound cuts were calculated by comparing weighted averages of the change in frequency and amplitude values, giving the Upsweep Mean using Equation 9.

- Upsweep Mean: weighted average of the change in F_{50} values, mm as weights.

$$\frac{\sum_{i=1}^{N} mm_i \Delta F_{50i}}{\sum_{i=1}^{N} mm_i} \tag{9}$$

10

Estimates of the relative occurrence of frequency modulation were calculated by determining the fraction of total energy that coincided with increases in frequency. This Upsweep Fraction was computed by Equation 10.

- Upsweep Fraction: fraction of summed mm values that coincide with increases in F_{50}.

$$\frac{\sum_{\Delta F_{50i}>0} mm_i}{\sum_{i=1}^{N} mm_i} \tag{10}$$

The relative magnitudes of frequency upsweeps and downsweeps were calculated as the Upsweep Proportion, using Equation 11.

- Upsweep Proportion: proportion of average weighted increase in F_{50} to the sum of the weighted average increases and decreases in F_{50}.

$$\frac{\dfrac{\sum_{\Delta F_{50i}>0} mm_i \Delta F_{50i}}{\sum_{\Delta F_{50i}>0} mm_i}}{\dfrac{\sum_{\Delta F_{50i}>0} mm_i \Delta F_{50i}}{\sum_{\Delta F_{50i}>0} mm_i} - \dfrac{\sum_{\Delta F_{50i}<0} mm_i \Delta F_{50i}}{\sum_{\Delta F_{50i}<0} mm_i}} \tag{11}$$

The magnitudes of linear relationships between time and median frequency were estimated by calculating correlation coefficients. This correlation used intensity values as weights to focus on portions of the sound cuts with loud signals. This was computed as a Time Frequency Correlation by Equation 12.

- Time Frequency Correlation: weighted correlation between F_{50} and t, ss as weights.

$$\frac{\dfrac{\sum_{i=0}^{N} ss_i t_i F_{50i}}{\sum_{i=0}^{N} ss_i} - \overline{tF_{50}}}{\sigma_t \sigma_F} \tag{12}$$

where

$$\sigma_F = \sqrt{\frac{\sum_{i=0}^{N} ss_i F_{50i}^2}{\sum_{i=0}^{N} ss_i} - \overline{F_{50}}^2}$$

The magnitudes of linear relationships between time and frequency upsweep were also estimated by a weighted correlation coefficient. This Time Upsweep Correlation was computed by Equation 13.

- Time Upsweep Correlation: weighted correlation between ΔF_{50} and t, mm as weights.

$$\frac{\frac{\sum_{i=1}^{N} mm_i t_i \Delta F_{50i}}{\sum_{i=1}^{N} mm_i} - \hat{t}\hat{F_{50}}}{\tau_t \tau_F} \qquad (13)$$

Where

$$\hat{t} = \frac{\sum_{i=1}^{N} mm_i t_i}{\sum_{i=1}^{N} mm_i}, \quad \hat{F_{50}} = \frac{\sum_{i=1}^{N} mm_i \Delta F_{50i}}{\sum_{i=1}^{N} mm_i}$$

$$\tau_t = \sqrt{\frac{\sum_{i=1}^{N} mm_i t_i^2}{\sum_{i=1}^{N} mm_i} - \hat{t}^2}, \quad \tau_F = \sqrt{\frac{\sum_{i=1}^{N} mm_i \Delta F_{50i}^2}{\sum_{i=1}^{N} mm_i} - \hat{F_{50}}^2}$$

2.5 Short-term Bandwidth

Two measures of bandwidth were computed for each block. The gross spread in power spectral values (frequency) was calculated to give the Short-term Bandwidth Mean by Equation 14.

- Short-term Bandwidth Mean: weighted average of $F_{75} - F_{25}$, ss as weights.

$$\frac{\sum_{i=0}^{N} ss_i (F_{75i} - F_{25i})}{\sum_{i=0}^{N} ss_i} \qquad (14)$$

The effective number of frequency bins in each block containing significant power levels (ignoring gaps in the sound) also was computed to give the Short-Term Spectral Concentration by Equation 15.

- Short-term Spectral Concentration: weighted average of N_{50}, ss as weights.

$$\frac{\sum_{i=0}^{N} ss_i N_{50i}}{\sum_{i=0}^{N} ss_i} \qquad (15)$$

Then, the relative emphasis of sideband energy on either side of the dominant frequency was estimated by calculating the Short-term Spectral Asymmetry using Equation 16.

- Short-term Spectral Asymmetry: weighted average spectral asymmetry, with ss as weights.

$$\frac{\sum_{i=0}^{N} ss_i \frac{F_{50i} - F_{25i}}{F_{75i} - F_{25i}}}{\sum_{i=0}^{N} ss_i} \qquad (16)$$

2.6 Aggregate Bandwidth

Two aggregate power spectra were computed for each sound cut: the Total Spectrum was the average of all FFT power spectra for the sound, and the Modal Spectrum accumulated only the power spectrum magnitudes for the frequency bin with the largest value in each FFT block. Both aggregate spectra were processed to extract three statistics related to bandwidth. The Upper (F_{75}) and Lower (F_{25}) Frequencies estimated the bounds of the aggregate spectra. These could be used to compute a bandwidth spanning any gaps in the spectral density. The Spectrum Concentration (N_{50}) provided an estimate of a "packed" or gap-free bandwidth.

2.7 Center Frequency

Three statistics were used to estimate the aggregate center frequency, or "average" frequency, of the entire sound cut. They were the F_{50} values computed for the Total and Modal Spectra, and a weighted average of the instantaneous F_{50} values computed by Equation 17. These statistic produced very similar values in most instances.

- Median Frequency Mean: weighted average of F_{50}, ss as weights.

$$\overline{F_{50}} = \frac{\sum_{i=0}^{N} ss_i F_{50i}}{\sum_{i=0}^{N} ss_i} \qquad (17)$$

2.8 Amplitude-Frequency Interaction

The strength of a linear trend between amplitude and center frequency was computed as the Amplitude-Frequency Correlation by Equation 18.

- Amplitude Frequency Correlation: correlation between ss and F_{50}.

$$\frac{\sum_{i=0}^{N} ss_i F_{50i} - \overline{ss} \sum_{i=0}^{N} F_{50i}}{\sigma_{ss}\sqrt{N \sum_{i=0}^{N} F_{50i}^2 - \left(\sum_{i=0}^{N} F_{50i}\right)^2}} \qquad (18)$$

The magnitude of a linear trend between amplitude and frequency modulation was calculated as the Amplitude Upsweep Correlation, using Equation 19.

- Amplitude Upsweep Correlation: correlation between mm and ΔF_{50}.

$$\frac{\sum_{i=1}^{N} mm_i \Delta F_{50i} - \overline{mm} \sum_{i=1}^{N} \Delta F_{50i}}{\sigma_{mm}\sqrt{N \sum_{i=1}^{N} \Delta F_{50i}^2 - \left(\sum_{i=1}^{N} \Delta F_{50i}\right)^2}} \qquad (19)$$

where

$$\overline{mm} = \frac{\sum_{i=1}^{N} mm_i}{N}, \quad \sigma_{mm} = \sqrt{\frac{\sum_{i=1}^{N} mm_i^2}{N} - \overline{mm}^2}$$

14

3 Noise Performance

To explore the effectiveness of our system of noise compensation, test sounds were copied and contaminated with increasing levels of noise. These test files were processed to generate plots of each statistic relative to signal/noise ratios. The differences between adjacent points along each curve reflected estimation error, because adjacent points represented different noise sequences with similar signal/noise ratios. Any overall trend in these plots reflected imperfect noise compensation.

Against Gaussian white noise, these noise tests were successful. However, the density of this type of noise was favorable for other noise compensation algorithms as well, and we were reminded that noise backgrounds in the ocean were rarely "white" (equal energy at all frequencies). Therefore, noise was generated synthetically to resemble the ambient background for many of our recordings at sea using MATLAB software (The Math Works, Inc.). Parametric spectral estimation procedures extracted parameters for a sixth order autoregressive (AR) process from an ambient noise sample. Noise sequences were generated by filtering white noise with a finite impulse response (FIR) filter constructed from the AR estimates.

The sound statistics were tested against noise and plotted in Figures 1 through 27 to provide an indication of performance for each statistic. In these figures, the horizontal axis represented signal/noise ratios, proceeding from low to relatively high values. Note that most of our marine mammal sound sequences exceeded 15 dB signal/noise. The vertical axes represented the estimated numeric level for this statistic. The variance may be seen in the relative amplitude fluctuations on the vertical axes, and the trend for this statistic may be seen in the relative changes in the progression from lower to higher signal/noise along horizontal axes.

The performance of each of the statistics relative to noise was summarized in the table

below, with the plots subjectively classified as follows.

Table of Noise Performance for the Statistics

	Low Variance	High Variance
small trend	Signal Center (Fig. 1) Median Frequency Mean (Fig. 25) Upsweep Mean (Fig. 9) Time Frequency Correlation (Fig. 12) Time Upsweep Correlation (Fig. 13) Upsweep Fraction (Fig. 10) Total Spectrum Median (Fig. 23) Modal Lower Frequency (Fig. 21) Modal Spectrum Median (Fig. 24) Modal Upper Frequency (Fig. 20)	Attack Fraction (Fig. 6) Attack Proportion (Fig. 7) Upsweep Proportion (Fig. 11) Short-term Spectral Asymmmetry (Fig. 16) Amplitude Skewness (Fig. 8)
large trend	Sound Duration (Fig. 2) Sound Concentration (Fig. 3 Amplitude Mean (Fig. 4) Amplitude Standard Deviation (Fig. 5) Short-term Bandwidth Mean (Fig. 14) Short-term Spectral Concentration (Fig. 15) Total Upper Frequency (Fig. 17) Total Lower Frequency (Fig. 18) Total Spectrum Concentration (Fig. 19) Modal Spectrum Concentration (Fig. 22)	Amplitude Frequency Correlation (Fig. 26) Amplitude Upsweep Correlation (Fig. 27)

Generally, the statistics performed well in noise. Higher order statistics (standard deviations, correlations) were less consistent, and the least useful were those statistics measuring frequency-amplitude relations and gross asymmetry in the sound waveform envelope.

16

4 Preliminary Analysis of Marine Animal Sounds

A subset of approximately 200 sequences of marine mammal sounds were used to test these statistics. Sounds were selected without attention to any particular acoustic features. They included sequences from 20 species: six baleen whales, 13 toothed whales and dolphins, and one seal. A wide variety of sound types was included in this subset; we also included sounds from pairs of species that were difficult to distinguish aurally.

4.1 Statistical Interdependence

The redundancy in these statistics, for this data set, was examined by a stepwise multiple regression procedure. This analysis treated each sound equally, ignoring the identity of the sequence. At each stage, the algorithm identified the statistic that had the highest linear correlation with other statistics for these data. This statistic was removed, and the analysis repeated. When the correlation coefficients and scatter plots indicated relatively poor fits, the analysis was terminated. We anticipated some redundancy in our statistics; we intended to test alternative sound measures. A more conclusive analysis of redundancy and performance awaits analysis of larger data sets.

For the comparisons described here, the multiple regression functions explained more than 80% of the variance for the first nine statistical estimators. The statistics that were successively eliminated were (with percent explained variance): Amplitude Standard Deviation (0.979, fig. 28), Median Frequency Mean (0.974, fig. 29), Modal Spectrum Median (0.974, fig. 31), Total Upper Frequency (0.944, fig. 31), Total Spectrum Concentration (0.900, fig. 32), Modal Lower Frequency (0.896, fig. 33), Modal Upper Frequency (.0.885, fig. 34), Short-term Spectral Concentration (0.878, fig. 15), and Total Spectrum Median (0.839, fig. 36). In these figures, the horizontal axes represented the predicted value, the vertical axes represented the observed value, and the dark line represented the regression

17

line. Eight of the nine variables appear to scale with center frequency, and we suspect that re-expressing these in relation to center frequency would remove much of this redundancy. Amplitude Standard Deviation and other amplitude variables sould be re-expressed relative to Amplitude Mean for similar reasons. It was not resolved whether simple division by these scaling factors would be appropriate.

4.2 Acoustical Analyses and Biological Information

The ability to select and analyze acoustic measurements based on related biological or environmental observations was crucial for these data. This could have been done by segregating data files for different species, activities, locations, etc. and independently processing each batch. However, it would have been cumbersome and difficult to manage such sorting and data segregation for each new query, especially as the selections became more complicated. A more powerful technique was to link the numerical analyses directly to the text databases. All sound cuts were processed in one batch, and these extensive computations proceeded automatically, unattended. Interactive exploration of relationships among statistics and biological or environmental factors followed, with all of the flexibility and convenience of database queries and reports.

The SOUND text databases for the recordings and the digital sound sequences (Watkins, Fristrup, and Daher 1991) could have accommodated new numeric data from the statistical analyses, but with INMAGIC software this required restructuring the entire database each time the number of numeric fields changed. This was not feasible: the analyses required many iterations and modifications. Therefore, PARADOX software (supports relational database models, with visual, query-by-example interface) was used to provide more flexible linkage between biological and acoustic information. The text information from the SOUND databases remained unmodified as a single table, and additional tables were created to handle the numeric results. Fields were used in these numeric tables that identified

18

the related SOUND text records. Then, subsets of the statistical results were obtained by selecting particular fields in SOUND and reporting the linked numerical information. Note that these queries could be reversed to select pertinent biological or environmental information based on acoustic criteria.

4.3 Preliminary Species Summaries

We generated summaries of the numerical results for each of the species in the trial data set (see Tables 2 through 8). For each species, the mean value of the statistic was listed with the maximum as a superscript and the minimum as a subscript. The number of sounds analyzed (Count) for each species was indicated in Table 1, but not repeated in the other tables. Tables were divided in three sections: baleen whales, toothed species, one seal and a transient sequence (hammer simulating clicks of *Physeter catodon*).

Two aspects of the summaries of sound duration (Table 2) merit comment. Species that are represented by more than ten sounds showed a dramatic variation between upper and lower bounds on both statistics. This may have reflected different selection criteria for the sound cuts. These sound cuts may have been a mix of isolated sounds, long sound sequences with intervals of silence, or continuous choruses from many individuals. All of these are valid data, but we need to differentiate among these classes of recordings in future analyses. Also, the sound "duty cycle" could be calculated by dividing Sound Concentration by Sound Duration. Table 2 indicates that baleen whales could largely be distinguished from toothed species by comparison of sound duty cycles.

Table 2

Species	Count	Sound Duration s.	Sound Concentration s.
B. mysticetus	5	$1.62\,^{2.50}_{.63}$	$.403\,^{.782}_{.149}$
C. marginata	3	$.86\,^{.93}_{.82}$	$.297\,^{.400}_{.217}$
E. glacialis	5	$.83\,^{1.27}_{.39}$	$.265\,^{.639}_{.066}$
E. australis	16	$.82\,^{2.39}_{.18}$	$.443\,^{1.319}_{.126}$
B. acutorostrata	5	$.51\,^{.70}_{.27}$	$.255\,^{.403}_{.149}$
B. physalus	15	$1.07\,^{4.80}_{0.00}$	$.332\,^{.829}_{.100}$
P. catodon	28	$4.15\,^{69.97}_{.32}$	$.057\,^{.172}_{.008}$
D. leucas	1	$1.17\,^{1.17}_{1.17}$	$.517\,^{.517}_{.517}$
S. longirostris	25	$1.39\,^{2.63}_{.10}$	$.089\,^{.256}_{.006}$
S. long. + P. cat.	4	$1.29\,^{1.88}_{.76}$	$.079\,^{.127}_{.023}$
S. bredanensis	7	$1.72\,^{2.41}_{.98}$	$.230\,^{.384}_{.095}$
C. commersonii	2	$1.46\,^{2.31}_{.61}$	$.102\,^{.142}_{.061}$
D. delphis	13	$1.68\,^{4.97}_{.66}$	$.182\,^{1.147}_{.020}$
G. griseus	23	$4.06\,^{8.32}_{.17}$	$.414\,^{1.295}_{.013}$
G. macrorhynchus	11	$1.68\,^{4.47}_{.62}$	$.403\,^{.846}_{.086}$
G. melaena	11	$1.37\,^{8.10}_{.14}$	$.253\,^{1.584}_{.039}$
O. orca	7	$.98\,^{1.83}_{.56}$	$.393\,^{.921}_{.040}$
P. crassidens	6	$5.94\,^{9.79}_{3.23}$	$.470\,^{1.098}_{.077}$
P. phocoena	8	$1.54\,^{3.48}_{.47}$	$.066\,^{.125}_{.030}$
I. geoffrensis	2	$1.31\,^{1.60}_{1.02}$	$.121\,^{.132}_{.111}$
A. phillipi	4	$4.43\,^{7.26}_{1.99}$	$.670\,^{1.019}_{.235}$
Hammer on metal	1	$2.41\,^{2.41}_{2.41}$	$.079\,^{.079}_{.079}$

Table 3 displayed statistics related to amplitude modulation. The first two of these statistics were not useful diagnostics in themselves because the absolute value of each was inherently tied to equipment gain settings during any stages of processing. However, the proportion of Amplitude Standard Deviation/Amplitude Mean (standard error) was a useful indication of amplitude modulation, with larger standard errors indicating more modulation.

Attack Fraction and Attack Proportion were negatively correlated because sound cuts were edited so that initial and terminal noise levels were approximately the same. One of these probably would have been sufficient.

Amplitude Skewness appeared to be a less robust statistic. However, some gross differences agreed with our experience. Among baleen whales, *Eubalaena glacialis* tended to start loudly and taper to silence (negative skew), while *B. acutorostrata* often started softly and swelled in level.

Table 3

Species	Amplitude Mean	Amplitude StdDev	Attack Fraction	Attack Proportion	Amplitude Skewness
B. mysticetus	$.095\,^{.265}_{.013}$	$.336\,^{.889}_{.041}$	$.51\,^{.57}_{.49}$	$.49\,^{.51}_{.43}$	$.439\,^{1.643}_{-.200}$
C. marginata	$.161\,^{.331}_{.028}$	$.291\,^{.548}_{.062}$	$.44\,^{.56}_{.38}$	$.56\,^{.63}_{.44}$	$-.140\,^{.016}_{-.289}$
E. glacialis	$.195\,^{.513}_{.011}$	$.501\,^{1.584}_{.057}$	$.51\,^{.57}_{.45}$	$.49\,^{.55}_{.43}$	$-1.127\,^{.486}_{-5.160}$
E. australis	$1.254\,^{5.651}_{.034}$	$2.263\,^{7.928}_{.125}$	$.50\,^{.69}_{.40}$	$.50\,^{.60}_{.31}$	$-.232\,^{2.658}_{-1.391}$
B. acutorostrata	$19.177\,^{38.858}_{4.948}$	$35.734\,^{64.134}_{12.505}$	$.45\,^{.56}_{.25}$	$.55\,^{.75}_{.44}$	$1.123\,^{1.648}_{.339}$
B. physalus	$1.612\,^{7.413}_{.000}$	$3.383\,^{12.925}_{.001}$	$.50\,^{.75}_{.29}$	$.56\,^{.92}_{.38}$	$.444\,^{3.828}_{-2.211}$
P. catodon	$.077\,^{.310}_{.002}$	$.442\,^{1.565}_{.010}$	$.46\,^{.57}_{.32}$	$.54\,^{.68}_{.42}$	$-.226\,^{.885}_{-2.052}$
D. leucas	$.010\,^{.010}_{.010}$	$.021\,^{.021}_{.021}$	$.52\,^{.52}_{.52}$	$.48\,^{.48}_{.48}$	$.678\,^{.678}_{.678}$
S. longirostris	$.143\,^{1.387}_{.000}$	$.427\,^{3.700}_{.000}$	$.50\,^{.56}_{.45}$	$.50\,^{.55}_{.44}$	$.131\,^{4.023}_{-3.049}$
S. long. + P. cat.	$.010\,^{.016}_{.004}$	$.046\,^{.058}_{.032}$	$.51\,^{.53}_{.48}$	$.49\,^{.52}_{.47}$	$-.392\,^{2.632}_{-5.964}$
S. bredanensis	$.113\,^{.383}_{.014}$	$.298\,^{1.023}_{.048}$	$.49\,^{.51}_{.48}$	$.51\,^{.52}_{.49}$	$.369\,^{3.189}_{-.575}$
C. commersonii	$.008\,^{.012}_{.004}$	$.043\,^{.056}_{.030}$	$.58\,^{.62}_{.54}$	$.42\,^{.46}_{.38}$	$-5.200\,^{-.840}_{-9.560}$
D. delphis	$.126\,^{.430}_{.004}$	$.484\,^{1.560}_{.022}$	$.50\,^{.54}_{.46}$	$.49\,^{.56}_{.39}$	$-.041\,^{4.272}_{-4.224}$
G. griseus	$.060\,^{.299}_{.001}$	$.163\,^{.622}_{.007}$	$.49\,^{.53}_{.45}$	$.51\,^{.55}_{.47}$	$.038\,^{2.064}_{-3.067}$
G. macrorhynchus	$.381\,^{1.079}_{.056}$	$.952\,^{2.785}_{.090}$	$.50\,^{.57}_{.46}$	$.50\,^{.54}_{.43}$	$.017\,^{2.366}_{-1.606}$
G. melaena	$.726\,^{1.757}_{.153}$	$1.653\,^{3.603}_{.377}$	$.50\,^{.57}_{.42}$	$.50\,^{.58}_{.43}$	$.192\,^{2.065}_{-2.607}$
O. orca	$.228\,^{.451}_{.011}$	$.384\,^{.610}_{.048}$	$.50\,^{.56}_{.46}$	$.50\,^{.54}_{.44}$	$.280\,^{1.348}_{-.360}$
P. crassidens	$.084\,^{.331}_{.001}$	$.231\,^{.682}_{.006}$	$.49\,^{.53}_{.46}$	$.51\,^{.54}_{.47}$	$.030\,^{.984}_{-.840}$
P. phocoena	$.171\,^{.933}_{.000}$	$.944\,^{5.185}_{.001}$	$.48\,^{.56}_{.39}$	$.42\,^{.56}_{.09}$	$-.106\,^{3.919}_{-5.415}$
I. geoffrensis	$.217\,^{.244}_{.190}$	$1.007\,^{1.140}_{.874}$	$.43\,^{.44}_{.42}$	$.57\,^{.58}_{.56}$	$-.992\,^{-.819}_{-1.166}$
A. phillipi	$.566\,^{1.138}_{.161}$	$1.359\,^{2.517}_{.585}$	$.51\,^{.54}_{.48}$	$.49\,^{.52}_{.46}$	$.081\,^{.836}_{-.301}$
Hammer on metal	$1.489\,^{1.489}_{1.489}$	$10.214\,^{10.214}_{10.214}$	$.40\,^{.40}_{.40}$	$.60\,^{.60}_{.60}$	$-.192\,^{-.192}_{-.192}$

Table 4 displays statistics relating to frequency modulation. Baleen whales and toothed species clearly had different magnitudes of Upsweep Mean, but much of this could have resulted from higher center frequencies of most toothed whale sounds.

Upsweep Fraction and Upsweep Proportion were not always negatively correlated, because the sounds from marine mammals could start or end at very different frequencies. The numbers for *E. glacialis* were instructive. Upsweep Mean was negative for this species, indicating that the sounds had lower frequencies toward the end, and Upsweep Fraction was nearly half indicating that about half of the block-to-block changes in frequency were positive. Thus, most of these downsweeps were greater in magnitude than the upsweeps, and indeed the Upsweep Proportion was less than one-half.

The Time Frequency and Time Upsweep Correlations showed considerable variation within larger samples. We know that the features they target are useful diagnostics, so we must seek better means of measuring them. Alternatives include non-parametric measures of correlation.

Table 4

Species	Upsweep Mean Hz/s.	Upsweep Fraction	Upsweep Proportion	Time-Freq. Correlation	Time-Upsweep Correlation
B. mysticetus	$-395\,^{18}_{-1149}$	$.47\,^{.58}_{.36}$	$.50\,^{.70}_{.38}$	$-.24\,^{.42}_{-.94}$	$.02\,^{.24}_{-.07}$
C. marginata	$-144\,^{-107}_{-191}$	$.00\,^{.01}_{.00}$	$.68\,^{.88}_{.34}$	$-.27\,^{.01}_{-.45}$	$.18\,^{.50}_{-.22}$
E. glacialis	$-1320\,^{145}_{-3879}$	$.54\,^{.99}_{.32}$	$.34\,^{.56}_{.03}$	$-.17\,^{.35}_{-.64}$	$.10\,^{.32}_{-.03}$
E. australis	$102\,^{553}_{-92}$	$.69\,^{1.00}_{.36}$	$.57\,^{1.00}_{.20}$	$.18\,^{.75}_{-.98}$	$.16\,^{.62}_{-.15}$
B. acutorostrata	$0\,^{98}_{-100}$	$.72\,^{.97}_{.41}$	$.35\,^{.50}_{.09}$	$-.08\,^{.73}_{-.63}$	$-.04\,^{.33}_{-.41}$
B. physalus	$-36\,^{12}_{-134}$	$.17\,^{.94}_{0.00}$	$.32\,^{.91}_{0.00}$	$-.10\,^{1.00}_{-1.0}$	$-.12\,^{.86}_{-.98}$
P. catodon	$45628\,^{799608}_{-962976}$	$.58\,^{1.00}_{0.00}$	$.52\,^{1.00}_{0.00}$	$-.13\,^{.74}_{-.76}$	$-.03\,^{.72}_{-.63}$
D. leucas	$-680\,^{-680}_{-680}$	$.18\,^{.18}_{.18}$	$.69\,^{.69}_{.69}$	$-.82\,^{-.82}_{-.82}$	$-.02\,^{-.02}_{-.02}$
S. longirostris	$172907\,^{3788800}_{-1082692}$	$.70\,^{1.00}_{.09}$	$.59\,^{1.00}_{.16}$	$-.00\,^{.89}_{-.89}$	$.04\,^{.87}_{-.32}$
S. long. + P. cat.	$81780\,^{477717}_{-115619}$	$.54\,^{.83}_{.40}$	$.29\,^{.41}_{.06}$	$-.40\,^{-.02}_{-.90}$	$-.03\,^{.09}_{-.13}$
S. bredanensis	$41495\,^{331392}_{-41335}$	$.56\,^{.63}_{.48}$	$.45\,^{.51}_{.36}$	$-.04\,^{.09}_{-.16}$	$-.02\,^{.05}_{-.13}$
C. commersonii	$-6998\,^{-790}_{-13207}$	$.17\,^{.26}_{.08}$	$.37\,^{.50}_{.25}$	$.10\,^{.42}_{-.22}$	$-.41\,^{-.30}_{-.51}$
D. delphis	$-70028\,^{81760}_{-788441}$	$.58\,^{1.00}_{.05}$	$.58\,^{1.00}_{.22}$	$.00\,^{.66}_{-.54}$	$.00\,^{.23}_{-.38}$
G. griseus	$4432\,^{52320}_{-9104}$	$.55\,^{.73}_{.35}$	$.49\,^{.80}_{.22}$	$.04\,^{.87}_{-.45}$	$-.02\,^{.16}_{-.27}$
G. macrorhynchus	$10434\,^{57709}_{-7471}$	$.51\,^{.74}_{.33}$	$.51\,^{.79}_{.28}$	$-.04\,^{.70}_{-.66}$	$-.01\,^{.11}_{-.35}$
G. melaena	$2431\,^{29658}_{-11684}$	$.51\,^{.88}_{.01}$	$.49\,^{.78}_{.30}$	$.01\,^{.96}_{-.90}$	$-.08\,^{.17}_{-.47}$
O. orca	$-129211\,^{2504}_{-896083}$	$.48\,^{.71}_{.09}$	$.43\,^{.68}_{.25}$	$-.05\,^{.41}_{-.40}$	$-.16\,^{.05}_{-.90}$
P. crassidens	$-161\,^{1082}_{-2131}$	$.55\,^{.65}_{.42}$	$.38\,^{.62}_{.10}$	$.45\,^{.79}_{.11}$	$-.00\,^{.04}_{-.11}$
P. phocoena	$8550\,^{46219}_{-2429}$	$.68\,^{1.00}_{0.00}$	$.70\,^{1.00}_{0.00}$	$.02\,^{1.00}_{-.99}$	$-.09\,^{.16}_{-.36}$
I. geoffrensis	$-43940\,^{-6242}_{-81638}$	$.35\,^{.41}_{.29}$	$.54\,^{.55}_{.53}$	$-.15\,^{-.09}_{-.21}$	$.04\,^{.10}_{-.02}$
A. phillipi	$-1633\,^{-6}_{-4367}$	$.23\,^{.44}_{.10}$	$.45\,^{.70}_{.22}$	$.06\,^{.26}_{-.02}$	$.02\,^{.09}_{-.02}$
Hammer on metal	$-503277\,^{-503277}_{-503277}$	$.01\,^{.01}_{.01}$	$.20\,^{.20}_{.20}$	$-.36\,^{-.36}_{-.36}$	$.48\,^{.48}_{.48}$

The short-term bandwidth statistics in Table 5, the aggregate bandwidth statistics in Table 6, and the center frequency statistics of Table 7 were the most diagnostic for this set of sound sequences. They apparently separated the sounds of different species. Bandwidths appeared to scale with frequency, suggesting that they could be expressed best in proportion to center frequency.

Table 5

Species	Short-term Bandwidth Mean Hz/s.	Short-term Spectral Concentration Hz/s.	Short-term Spectral Asymmetry
B. mysticetus	34502^{39597}_{32023}	3016^{3756}_{2592}	$.11^{.22}_{.03}$
C. marginata	881^{949}_{810}	106^{118}_{99}	$.17^{.25}_{.08}$
E. glacialis	7444^{19157}_{3200}	335^{572}_{200}	$-.15^{.36}_{0.00}$
E. australis	4378^{6773}_{3269}	417^{680}_{265}	$.22^{.40}_{0.00}$
B. acutorostrata	907^{1094}_{814}	102^{158}_{74}	$.20^{.54}_{0.00}$
B. physalus	464^{945}_{54}	40^{96}_{7}	$.06^{.29}_{0.00}$
P. catodon	$1898178^{6104578}_{244488}$	135440^{339051}_{31364}	$.41^{.76}_{.07}$
D. leucas	52254^{52254}_{52254}	4652^{4652}_{4652}	$.13^{.13}_{.13}$
S. longirostris	$1189241^{2556271}_{821087}$	102772^{316991}_{51781}	$.12^{.31}_{.00}$
S. long. + P. cat.	$3387894^{6005715}_{891902}$	221269^{403662}_{54720}	$.22^{.42}_{.00}$
S. bredanensis	$3454478^{7664267}_{1336456}$	166331^{312275}_{106172}	$.22^{.31}_{.15}$
C. commersonii	30991^{33563}_{28419}	2471^{2765}_{2176}	$.04^{.05}_{.04}$
D. delphis	$1266464^{3159867}_{132364}$	102935^{336716}_{13007}	$.13^{.32}_{.01}$
G. griseus	$508874^{2462444}_{57730}$	26081^{67239}_{5404}	$.24^{.64}_{.08}$
G. macrorhynchus	$721235^{1932495}_{129749}$	38673^{102974}_{11581}	$.25^{.49}_{.10}$
G. melaena	327154^{653042}_{220740}	27724^{48173}_{17660}	$.18^{.41}_{.05}$
O. orca	214199^{322328}_{14074}	17242^{27372}_{1113}	$.19^{.46}_{.09}$
P. crassidens	200492^{430008}_{113113}	20372^{47612}_{9063}	$.14^{.23}_{.09}$
P. phocoena	39577^{126412}_{12800}	3815^{10007}_{800}	$.17^{.47}_{0.00}$
I. geoffrensis	$1102209^{1576500}_{627918}$	47188^{56594}_{37782}	$.49^{.52}_{.47}$
A. phillipi	137311^{164278}_{119179}	12491^{14300}_{11469}	$.18^{.20}_{.17}$
Hammer on metal	$3383956^{3383956}_{3383956}$	119674^{119674}_{119674}	$.36^{.36}_{.36}$

Table 6

Species	Total Spectrum			Modal Spectrum		
	Upper Frequency Hz	Lower Frequency Hz	Spectrum Concentration Hz	Upper Frequency Hz	Lower Frequency Hz	Spectrum Concentration Hz
B. mysticetus	821_{250}^{1841}	545_{187}^{1438}	149_{41}^{302}	903_{250}^{2375}	558_{188}^{1438}	114_{32}^{212}
C. marginata	93_{90}^{99}	79_{76}^{80}	15_{11}^{19}	91_{90}^{94}	76_{70}^{80}	10_{5}^{17}
E. glacialis	361_{200}^{717}	156_{20}^{300}	48_{12}^{119}	314_{40}^{660}	158_{20}^{310}	27_{12}^{55}
E. australis	174_{40}^{240}	124_{20}^{208}	38_{17}^{69}	164_{40}^{240}	132_{20}^{219}	22_{10}^{44}
B. acutorostrata	89_{40}^{130}	70_{28}^{92}	15_{7}^{24}	88_{40}^{130}	72_{30}^{100}	9_{5}^{15}
B. physalus	74_{30}^{140}	58_{20}^{129}	6_{3}^{15}	74_{30}^{140}	67_{25}^{130}	4_{2}^{10}
P. catodon	6463_{320}^{17100}	1411_{160}^{8480}	1385_{156}^{4044}	3093_{320}^{10240}	1919_{160}^{9734}	215_{82}^{610}
D. leucas	2563_{2563}^{2563}	2266_{2266}^{2266}	298_{298}^{298}	2517_{2517}^{2517}	2266_{2266}^{2266}	251_{251}^{251}
S. longirostris	11329_{640}^{27200}	6987_{320}^{25280}	1009_{174}^{2569}	11050_{640}^{26243}	7166_{320}^{25920}	757_{161}^{2209}
S. long. + P. cat.	20036_{10384}^{30080}	8640_{4480}^{10560}	2105_{181}^{3778}	11258_{10207}^{13120}	8640_{4480}^{10560}	674_{179}^{1256}
S. bredanensis	15250_{7045}^{22400}	4118_{320}^{7360}	2237_{751}^{5306}	8849_{3955}^{15040}	5517_{320}^{8202}	798_{295}^{1315}
C. commersonii	3574_{3398}^{3750}	2785_{1886}^{3683}	77_{46}^{109}	3574_{3398}^{3750}	3503_{3322}^{3685}	47_{36}^{57}
D. delphis	16484_{5120}^{37153}	9127_{1753}^{15040}	1284_{368}^{3212}	13598_{5120}^{27200}	9097_{2286}^{15306}	807_{284}^{2003}
G. griseus	8682_{960}^{13945}	4185_{117}^{12344}	903_{79}^{2221}	7815_{345}^{13594}	4440_{117}^{12344}	525_{54}^{1279}
G. macrorhynchus	5941_{2880}^{11840}	2186_{160}^{3447}	659_{163}^{1373}	5064_{320}^{11680}	2586_{160}^{6553}	356_{109}^{979}
G. melaena	5707_{2880}^{8715}	4586_{1992}^{8451}	556_{142}^{1037}	5599_{2880}^{8640}	4729_{1992}^{8442}	317_{113}^{542}
O. orca	1936_{800}^{3394}	949_{560}^{2064}	336_{72}^{557}	1635_{664}^{2488}	1067_{480}^{2062}	189_{60}^{389}
P. crassidens	6891_{234}^{13359}	5184_{117}^{7266}	473_{111}^{698}	6706_{234}^{13359}	5219_{117}^{7284}	313_{89}^{579}
P. phocoena	1522_{40}^{3080}	610_{0}^{2480}	110_{25}^{231}	660_{40}^{2560}	619_{0}^{2520}	28_{23}^{39}
I. geoffrensis	5291_{2847}^{7734}	1875_{1758}^{1992}	1143_{738}^{1548}	3379_{2813}^{3946}	2184_{1875}^{2493}	278_{256}^{301}
A. phillipi	234_{234}^{234}	117_{117}^{117}	100_{92}^{113}	234_{234}^{234}	117_{117}^{117}	67_{60}^{76}
Hammer on metal	1440_{1440}^{1440}	640_{640}^{640}	800_{800}^{800}	1111_{1111}^{1111}	800_{800}^{800}	210_{210}^{210}

27

Table 7

Species	Median Frequency Mean Hz	Total Spectrum Median Hz	Modal Spectrum Median Hz
B. mysticetus	$698\,^{1723}_{200}$	$642\,^{1586}_{187}$	$634\,^{1591}_{188}$
C. marginata	$83\,^{86}_{81}$	$82\,^{84}_{80}$	$82\,^{86}_{80}$
E. glacialis	$345\,^{593}_{138}$	$317\,^{622}_{139}$	$226\,^{330}_{28}$
E. australis	$157\,^{235}_{39}$	$145\,^{218}_{27}$	$141\,^{219}_{25}$
B. acutorostrata	$78\,^{110}_{28}$	$75\,^{103}_{28}$	$73\,^{100}_{30}$
B. physalus	$66\,^{129}_{25}$	$66\,^{129}_{25}$	$67\,^{130}_{25}$
P. catodon	$4906\,^{11283}_{525}$	$4340\,^{11578}_{234}$	$2606\,^{10069}_{182}$
D. leucas	$2058\,^{2058}_{2058}$	$2279\,^{2279}_{2279}$	$2312\,^{2312}_{2312}$
S. longirostris	$9451\,^{26124}_{349}$	$9702\,^{26160}_{320}$	$9664\,^{26018}_{321}$
S. long. + P. cat.	$16107\,^{22928}_{10992}$	$16848\,^{26772}_{10116}$	$10874\,^{12547}_{10013}$
S. bredanensis	$11038\,^{16942}_{7180}$	$8837\,^{17320}_{3849}$	$7230\,^{9013}_{2832}$
C. commersonii	$3236\,^{3664}_{2809}$	$3473\,^{3683}_{3262}$	$3503\,^{3685}_{3322}$
D. delphis	$13538\,^{26735}_{3285}$	$13571\,^{30097}_{1753}$	$11917\,^{19455}_{2286}$
G. griseus	$6518\,^{11305}_{939}$	$6763\,^{12346}_{869}$	$6408\,^{12397}_{286}$
G. macrorhynchus	$4339\,^{9384}_{2630}$	$3859\,^{8766}_{1573}$	$3398\,^{8568}_{198}$
G. melaena	$5195\,^{7203}_{2530}$	$5198\,^{8451}_{2534}$	$5136\,^{8442}_{2518}$
O. orca	$1850\,^{3357}_{938}$	$1480\,^{2064}_{622}$	$1290\,^{2062}_{583}$
P. crassidens	$6558\,^{9636}_{2091}$	$5623\,^{7964}_{165}$	$5459\,^{7955}_{148}$
P. phocoena	$1423\,^{2537}_{536}$	$1346\,^{3017}_{14}$	$625\,^{2520}_{9}$
I. geoffrensis	$3622\,^{4250}_{2994}$	$3387\,^{4072}_{2702}$	$2840\,^{3129}_{2551}$
A. phillipi	$263\,^{393}_{185}$	$149\,^{161}_{141}$	$125\,^{135}_{118}$
Hammer on metal	$2718\,^{2718}_{2718}$	$1217\,^{1217}_{1217}$	$840\,^{840}_{840}$

28

Table 8 displays statistics relating to amplitude-frequency interactions. Although relatively less robust than many of the previous statistics, these measures also appeared to be useful. For example, the statistics for *C. marginata* had positive Amplitude Frequency Correlation suggesting that the higher frequency sections were louder than low frequency ones, and the negative Amplitude Correlation indicated that sections with downsweeps (or smaller upsweeps) tended to be the loudest.

Note that all of these biological sounds except *A. phillipi* tended to have positive Amplitude Frequency Correlations (higher frequency sections were louder).

Table 8

Species	Amplitude Frequency Correlation	Amplitude Upsweep Correlation
B. mysticetus	$.27^{.40}_{.07}$	$-.01^{.01}_{-.04}$
C. marginata	$.51^{.91}_{.13}$	$-.28^{-.17}_{-.40}$
E. glacialis	$.11^{.37}_{-.04}$	$-.12^{.00}_{-.20}$
E. australis	$.24^{.92}_{-.15}$	$.04^{.21}_{-.06}$
B. acutorostrata	$.13^{.74}_{-.40}$	$.01^{.12}_{-.05}$
B. physalus	$.65^{1.00}_{-.22}$	$-.37^{.08}_{-.98}$
P. catodon	$.22^{.65}_{-.02}$	$-.01^{.43}_{-.98}$
D. leucas	$.69^{.69}_{.69}$	$-.10^{-.10}_{-.10}$
S. longirostris	$.34^{.90}_{-.02}$	$.09^{1.00}_{-.24}$
S. long. + P. cat.	$.51^{.57}_{.42}$	$-.00^{.11}_{-.07}$
S. bredanensis	$.28^{.47}_{.11}$	$.00^{.06}_{-.02}$
C. commersonii	$.62^{.63}_{.60}$	$-.26^{-.22}_{-.31}$
D. delphis	$.37^{.77}_{.10}$	$-.02^{.13}_{-.37}$
G. griseus	$.44^{.78}_{.19}$	$.03^{.15}_{-.05}$
G. macrorhynchus	$.19^{.48}_{.02}$	$.01^{.08}_{-.04}$
G. melaena	$.28^{.43}_{.17}$	$.01^{.07}_{-.05}$
O. orca	$.40^{.72}_{.07}$	$-.09^{.02}_{-.63}$
P. crassidens	$.40^{.68}_{.03}$	$-.00^{.00}_{-.02}$
P. phocoena	$.32^{.95}_{.11}$	$-.02^{.56}_{-1.00}$
I. geoffrensis	$.24^{.25}_{.24}$	$-.09^{-.01}_{-.16}$
A. phillipi	$-.04^{-.02}_{-.06}$	$-.01^{.01}_{-.02}$
Hammer on metal	$.47^{.47}_{.47}$	$-.63^{-.63}_{-.63}$

4.4 Principal Component Analysis

To obtain a better perspective on the overall distribution of sounds as measured by our statistics, we performed a principal components analysis on the numerical results. The first two principal components provided axes for scatter plots that expressed about half of the total variability in our statistics. As discussed earlier, segregation of the principal

component scores by species was accomplished by linking to the SOUND table in our PARADOX structure. The baleen whales and toothed species were plotted separately in Figures 38 and 39. These two groups generally could be separated by frequency alone, and the separate plots focus attention on the variation within these groups. Each sound cut is represented by a colored symbol on the plot; color and symbol type redundantly code species identity.

In both Figure 38 and Figure 39, the data for most species tended to cluster in a relatively discrete portion of the plot. Some regions were shared by a few neighboring species with similar sound types. These results appeared to confirm that acoustic features could be analyzed so as to compute sound statistics that would be useful to classify an unknown biological sound as one of a few potential candidates.

This result was particularly remarkable because the sound data used for the tests included both specific sounds produced by individuals and sounds produced by many animals, with temporally overlapping sounds by two or more animals. If we distinguish between the individual and group recordings, our acoustic classifier will perform much better.

A goal for such analyses has been the development of a system for automatic diagnosis of marine animal sounds based on acoustic criteria. The statistical problem for classification of an unknown relative to known groups is usually addressed by forming estimates of the distance between the unknown and the known, with the most likely classification being the one that minimizes this distance. The distance measure that is usually employed is the Mahalanobis distance (e. g. Morrison 1976, p. 241). Alternatively, we may find that methods which make fewer assumptions (Efron and Tibshirani 1991) will be better suited to developing the classifier. Non-acoustic criteria also could be incorporated from the text databases for additional refinement of these judgements.

5 Discussion and Summary

Software tools have been developed and tested for statistical analysis of marine animal sounds. The preliminary results suggest that such sounds can be classified by means of relatively simple statistical algorithms. Three areas of this research are planned for particular attention.

The SOUNDC database will need to be modified to identify additional information about the sound cuts. These additions will include identification of choruses of overlapping sounds and notes regarding the sequential organization of discrete sound elements. Sequences also needed to be identified relative to the usual pattern of sound production for that species or behavior, and noted if atypical.

Continued evolution of our sound statistics is inevitable. All frequency statistics should be re-scaled, such that the other information is not swamped by gross differences in center frequency, for example. Some correlations may need to be replaced with more robust (perhaps non-parametric) alternatives. Methods for expressing some of these statistics may need to be investigated to increase their independence from each other.

A comprehensive analysis of noise sensitivity and compensation techniques is also important. The specificity of our acoustic measurements is improved by removing the influence of ambient noise, but our ability to classify could be critically impaired if we erroneously discard portions of the signal.

Acknowledgements

6 LITERATURE CITED

Caldwell, M. C., D. K. Caldwell and P. L. Tyack 1990, *The Bottlenose Dolphin: recent progress in research*, Academic Press, San Diego.

Efron, B. and R. Tibshirani 1991, Statistical data analysis in the computer age, SCIENCE 253:390-395.

Morrison, D. F. 1976, *Multivariate Statistical Analysis*, 2nd ed., McGraw-Hill Inc., New York.

Watkins, William A., Kurt Fristrup, and Mary Ann Daher 1991. Marine animal SOUND database. Technical Report WHOI-91-21, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, 51 pp.

Watkins, William A., and Douglas Wartzok 1985. Sensory biophysics of marine mammals. Marine Mammal Science 1:219-260.
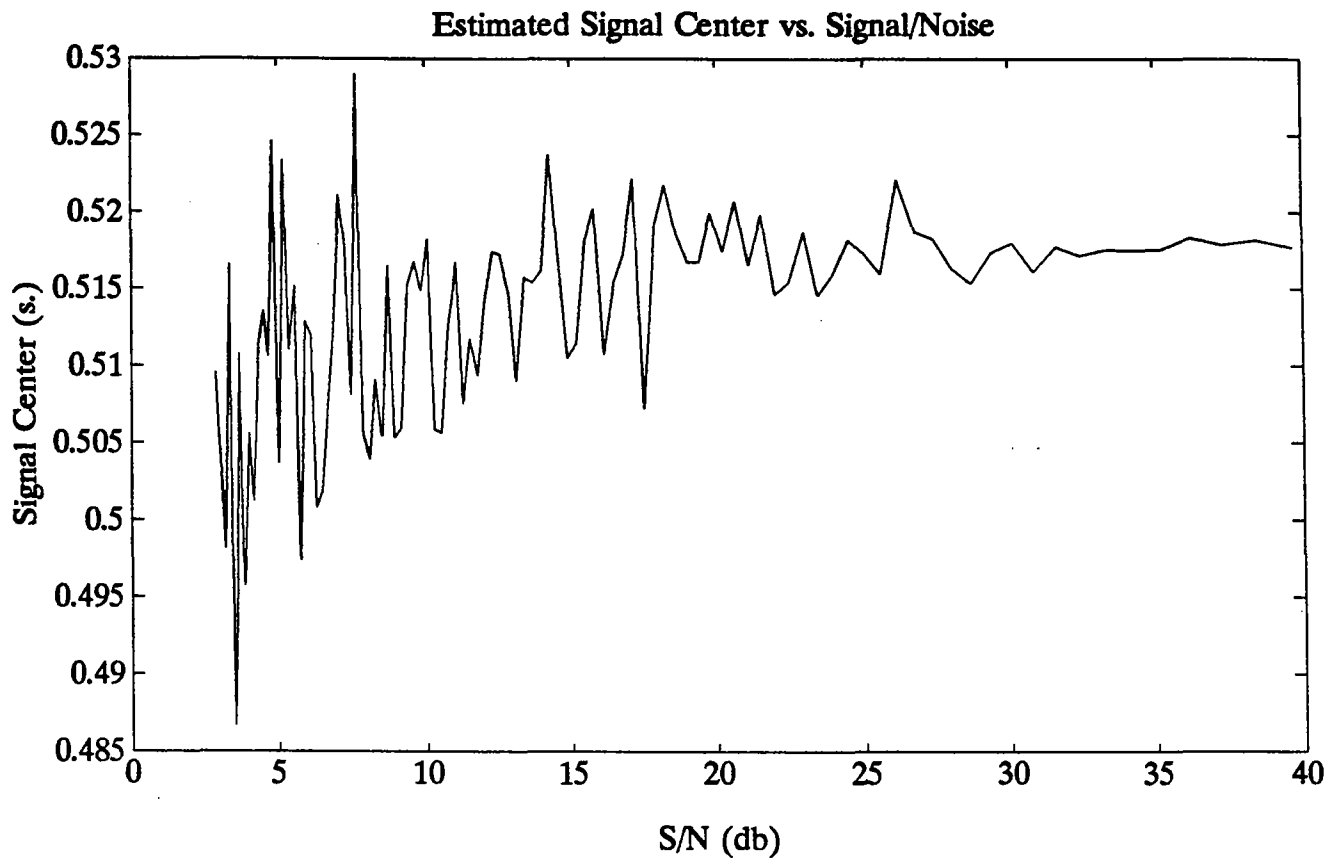
Figure 1. Noise performance of Signal Center (eq. 1). The vertical axis scores Signal Center in seconds; the horizontal axis scores signal/noise. Note the small range of values on the vertical axis. We subjectively label this as low variance, small trend.

Estimated Signal Duration vs. Signal/Noise
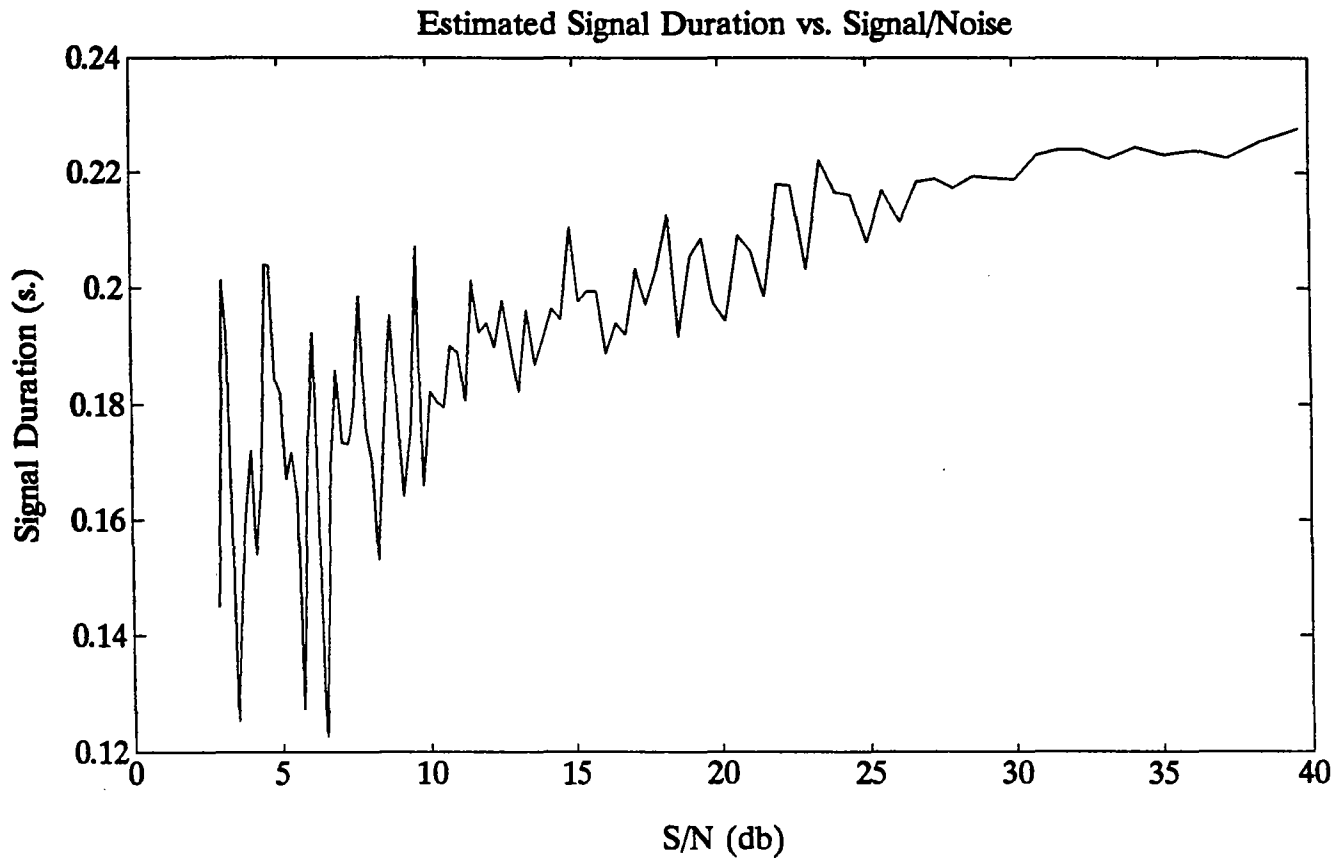
Figure 2. Noise performance of Signal Duration (eq. 2). The vertical axis scores Signal Duration in seconds; the horizontal axis scores signal/noise. We subjectively label this as low variance, large trend.

Estimated Signal Concentration vs. Signal/Noise
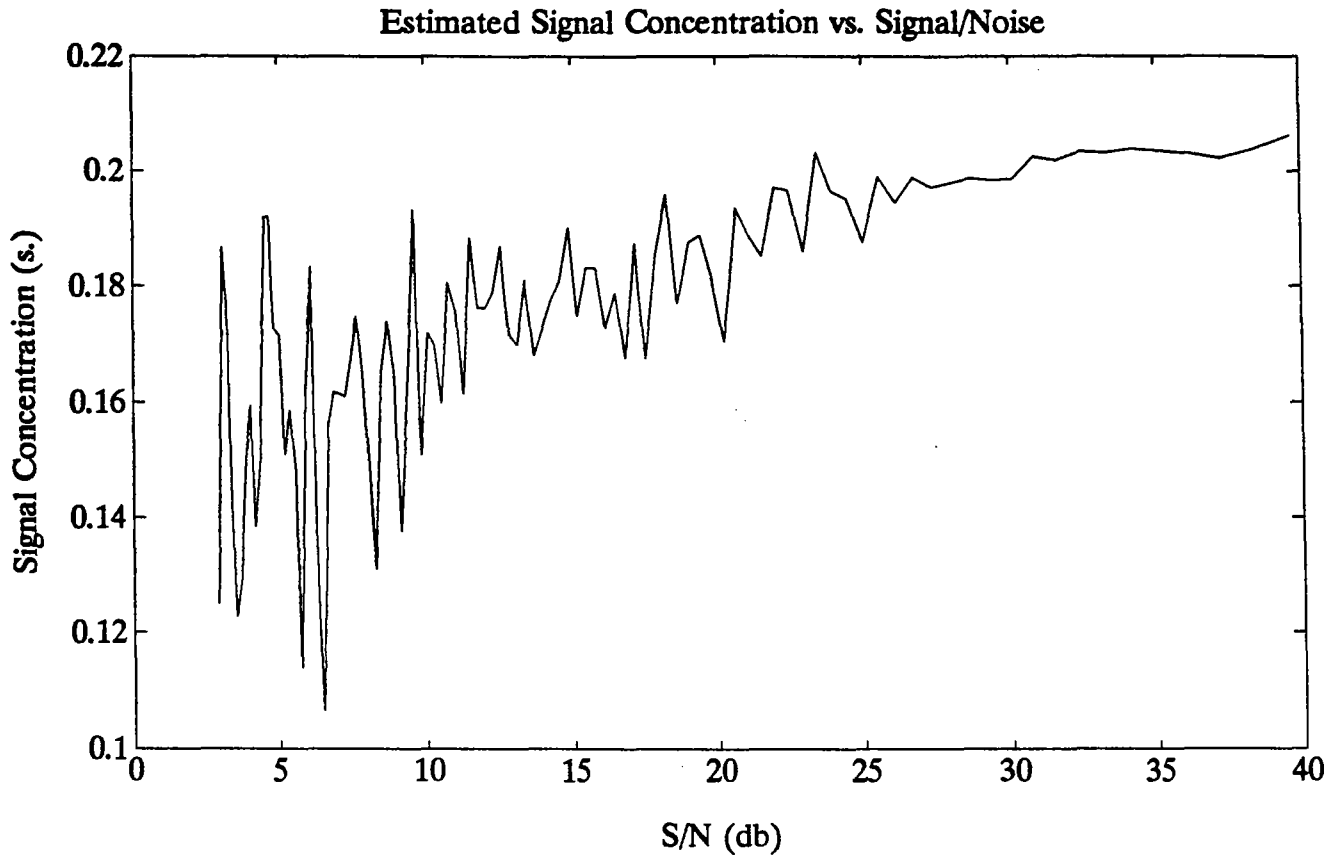
Figure 3. Noise performance of Signal Concentration (eq. 3). The vertical axis scores Signal Concentration in seconds; the horizontal axis scores signal/noise. We subjectively label this as low variance, large trend.

Estimated Amplitude Mean vs. Signal/Noise
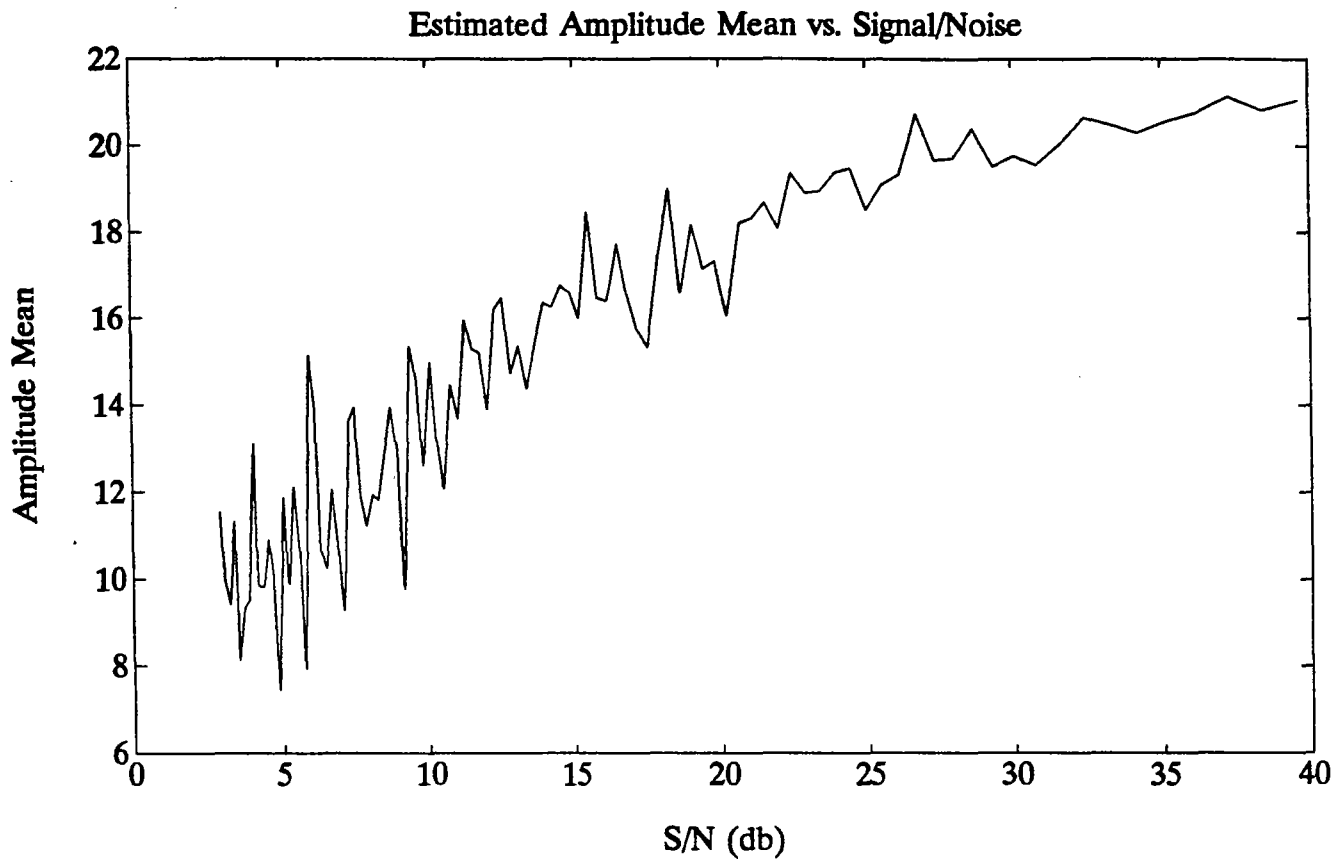
Figure 4. Noise performance of Amplitude Mean (eq. 4). The vertical axis scores Amplitude Mean in arbitrary units; the horizontal axis scores signal/noise. We subjectively label this as low variance, large trend.
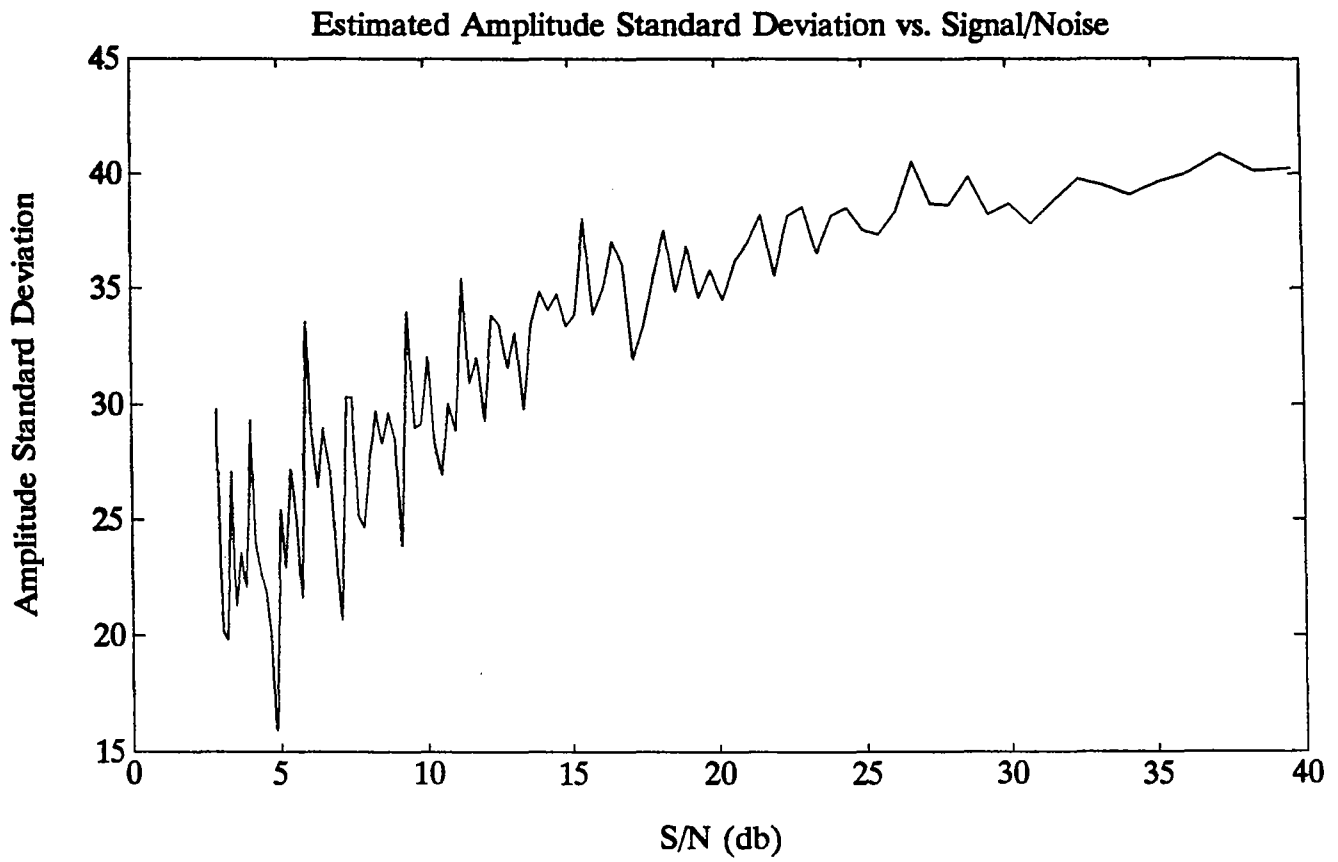
Figure 5. Noise performance of Amplitude Standard Deviation (eq. 5). The vertical axis scores Amplitude Standard Deviation, in arbitrary units; the horizontal axis scores signal/noise. We subjectively label this as low variance, large trend.

Estimated Attack Fraction vs. Signal/Noise
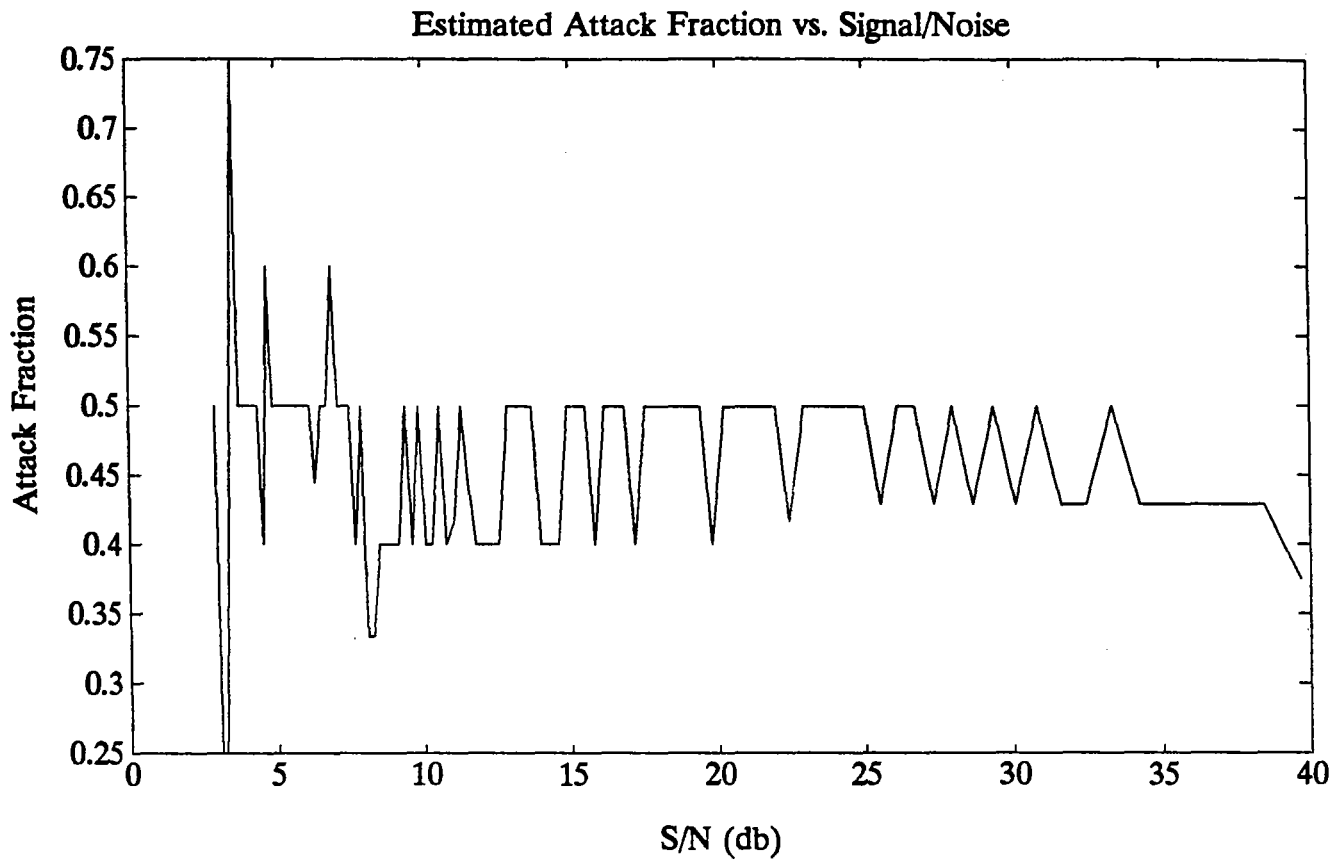
Figure 6. Noise performance of Attack Fraction (eq. 6). The vertical axis scores Attack Fraction (values of 0 ↔ 1 are possible); the horizontal axis scores signal/noise. We subjectively label this as high variance, small trend.

Estimated Attack Proportion vs. Signal/Noise
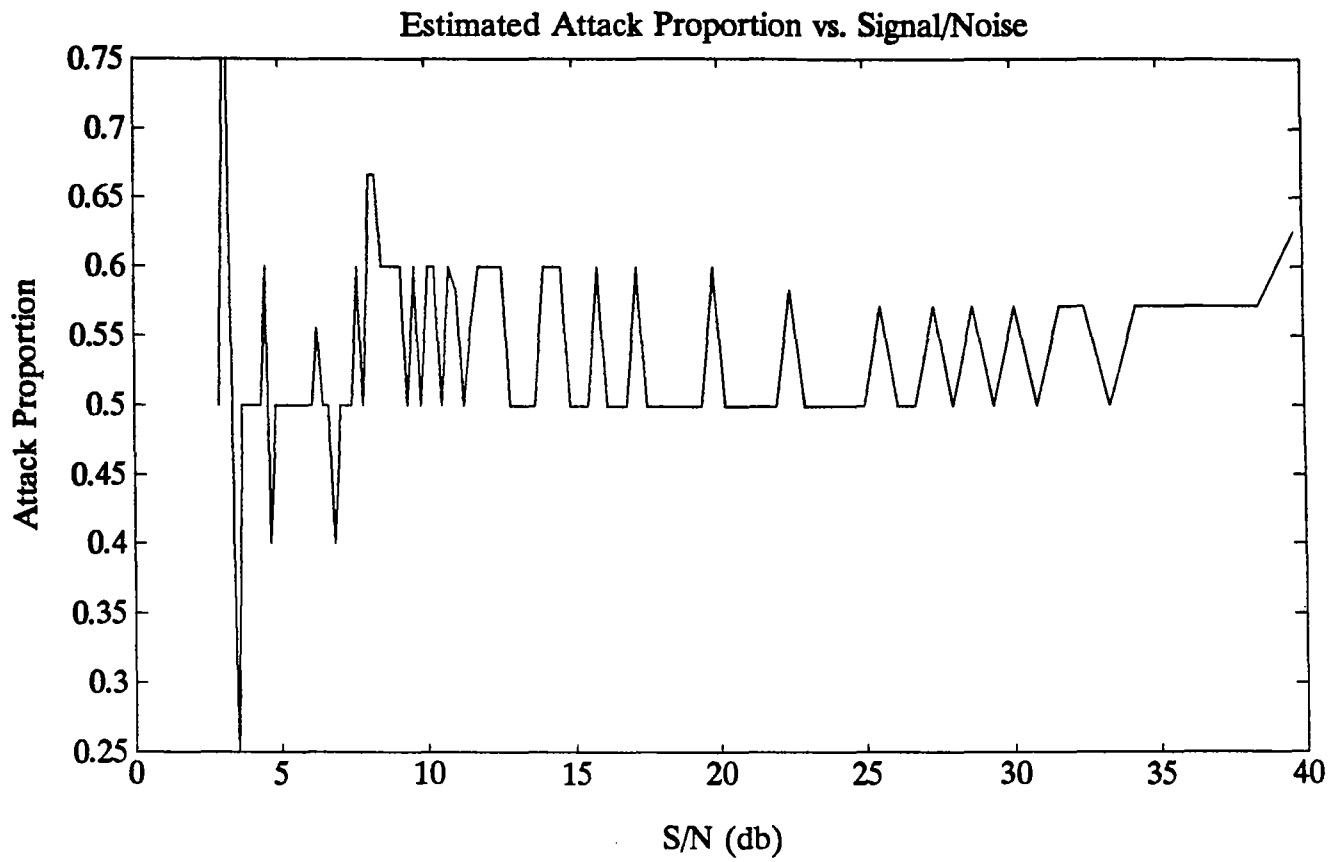
Figure 7. Noise performances of Attack Proportion (eq. 7). The vertical axis scores Attack Proportion (values of $0 \leftrightarrow 1$ are possible); the horizontal axis scores signal/noise. We subjectively label this as high variance, small trend.

Estimated Amplitude Skewness vs. Signal/Noise
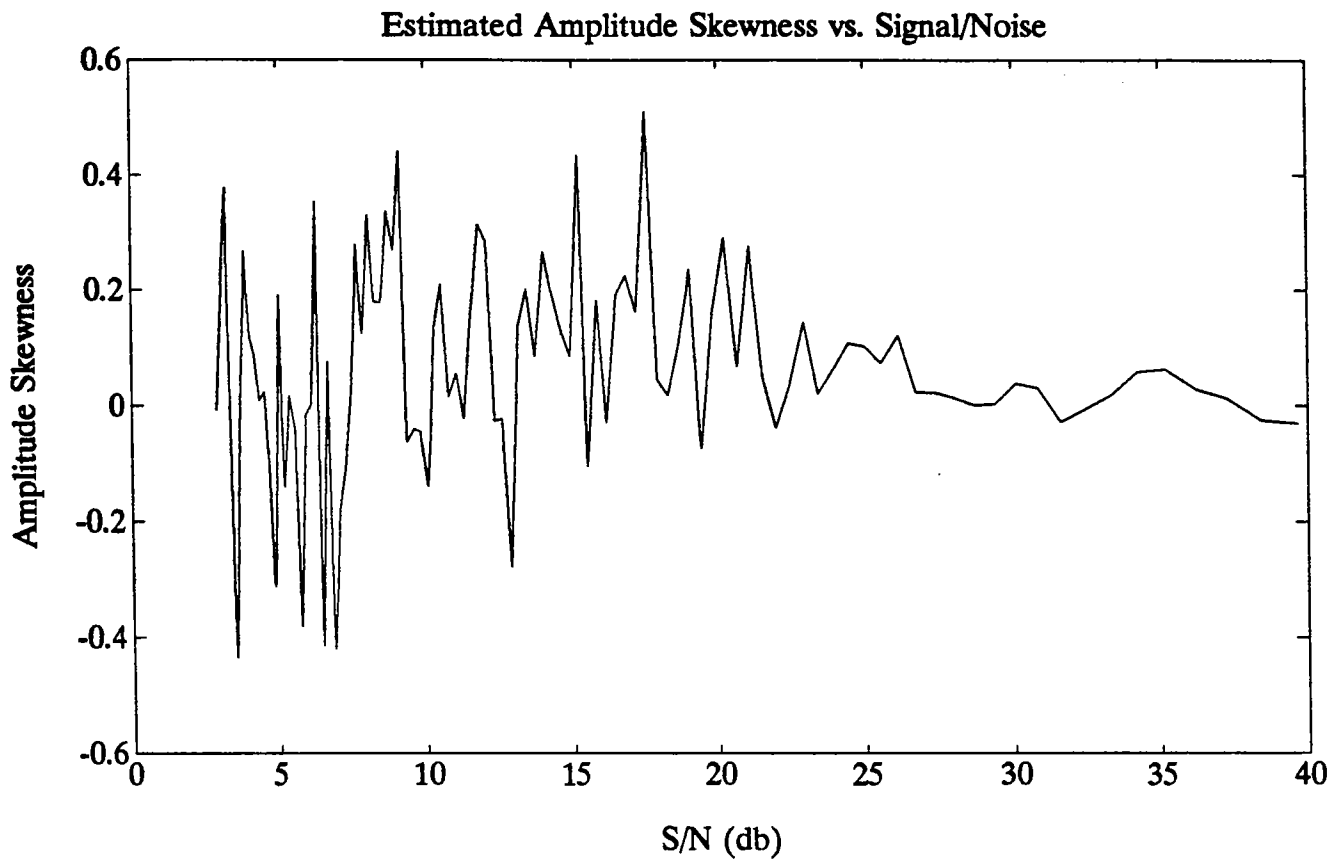
Figure 8. Noise performance of Amplitude Skewness (eq. 8). The horizontal axis scores Amplitude Skewness (scale independent); the horizontal axis scores signal/noise. We subjectively label this as high variance, small trend.

Estimated Upsweep Mean vs. Signal/Noise
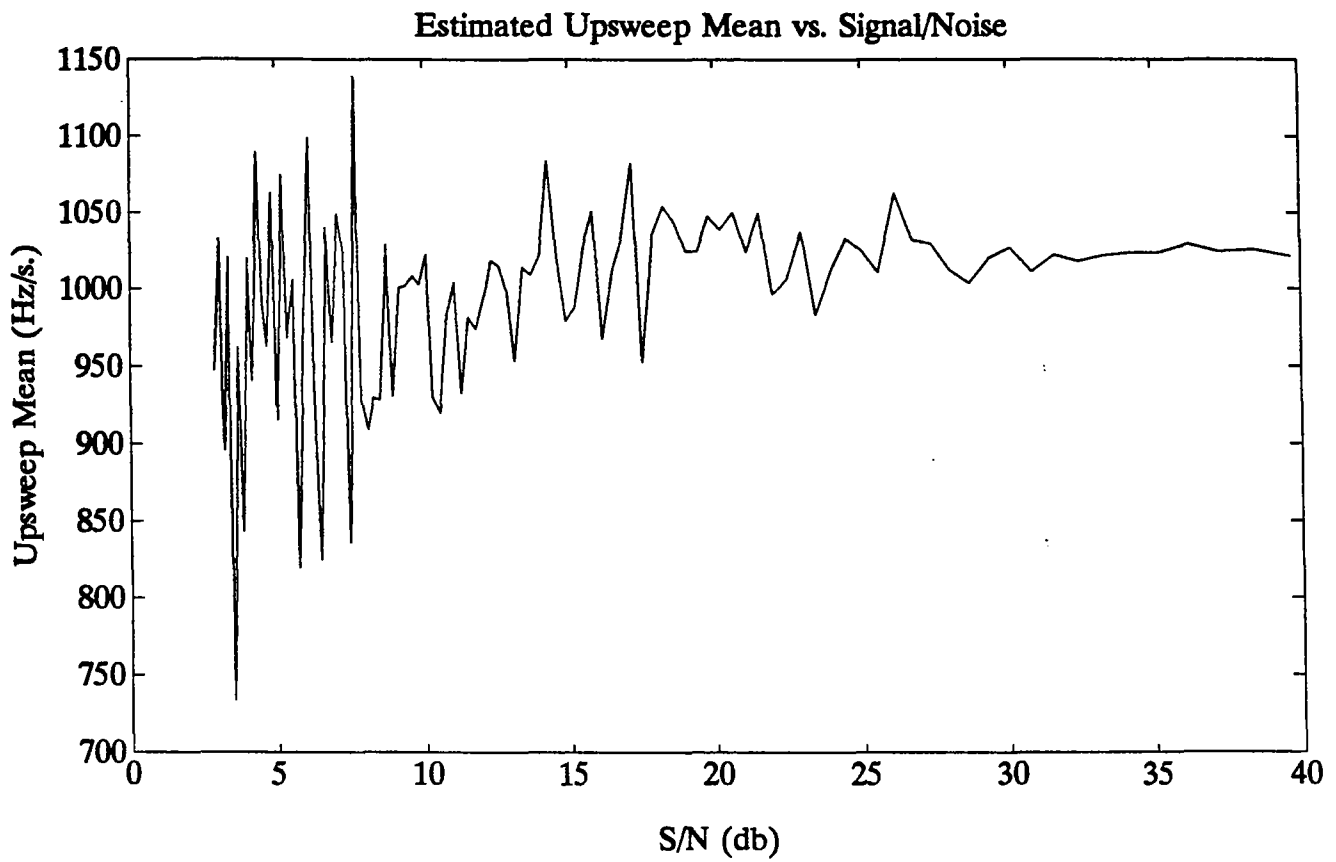
Figure 9. Noise performance of Upsweep Mean (eq. 9). The vertical axis scores Upsweep Mean in Hz/s.; the horizontal axis scores signal/noise. We subjectively label this as low variance, small trend.

44

Estimated Upsweep Fraction vs. Signal/Noise
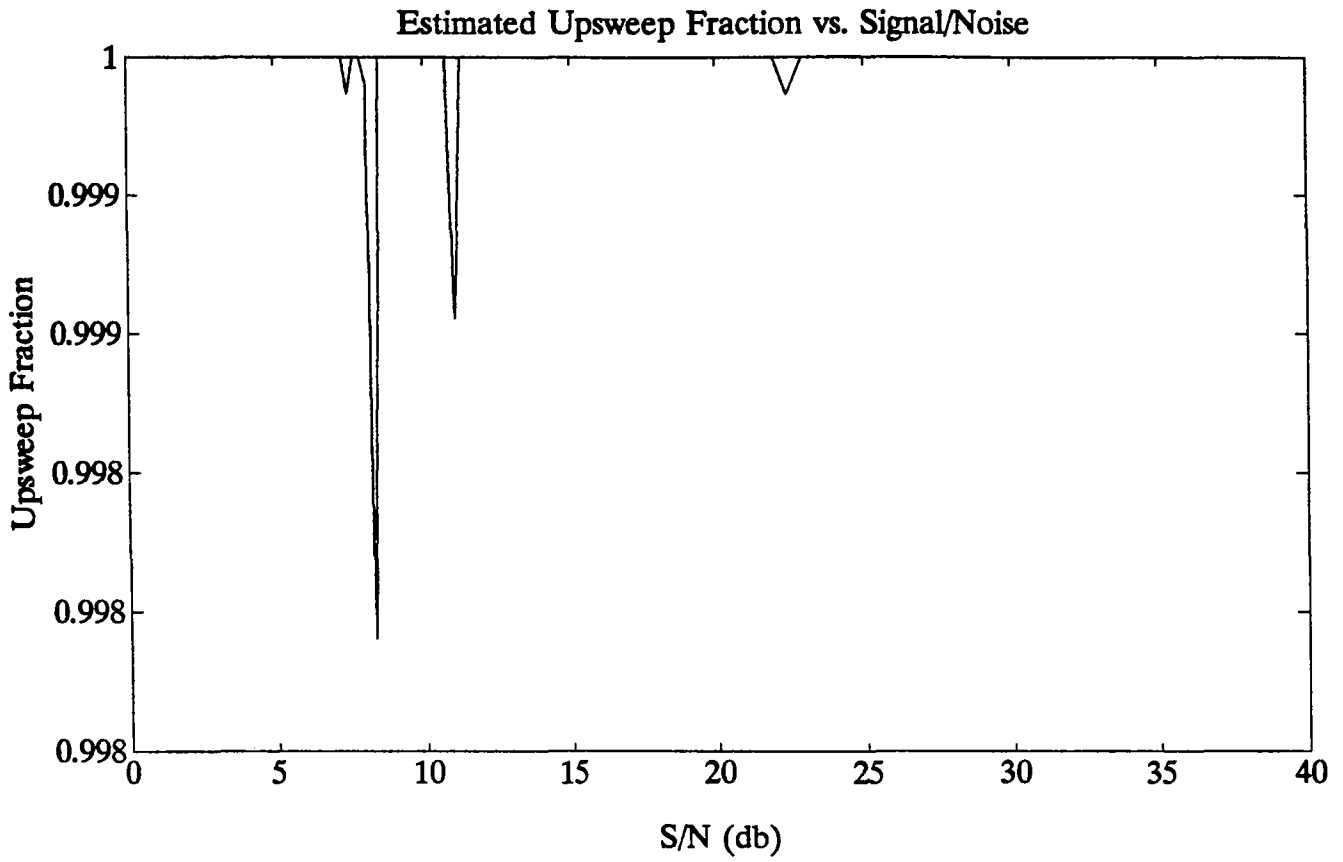
Figure 10. Noise performance of Upsweep Fraction (eq. 10). The vertical axis scores Upsweep Fraction (values of 0 ↔ 1 are possible); the horizontal axis scores signal/noise. We subjectively label this as low variance, small trend.

45
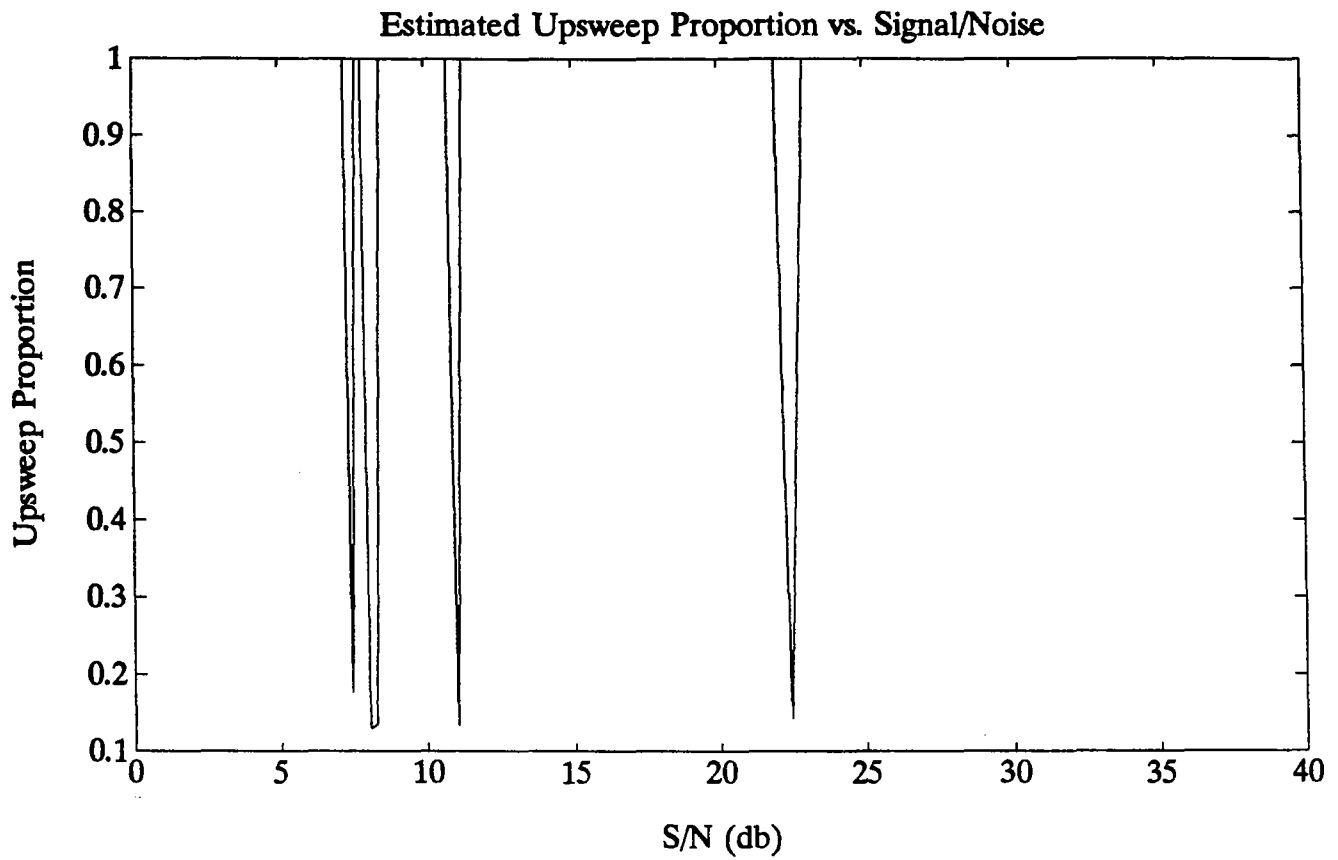
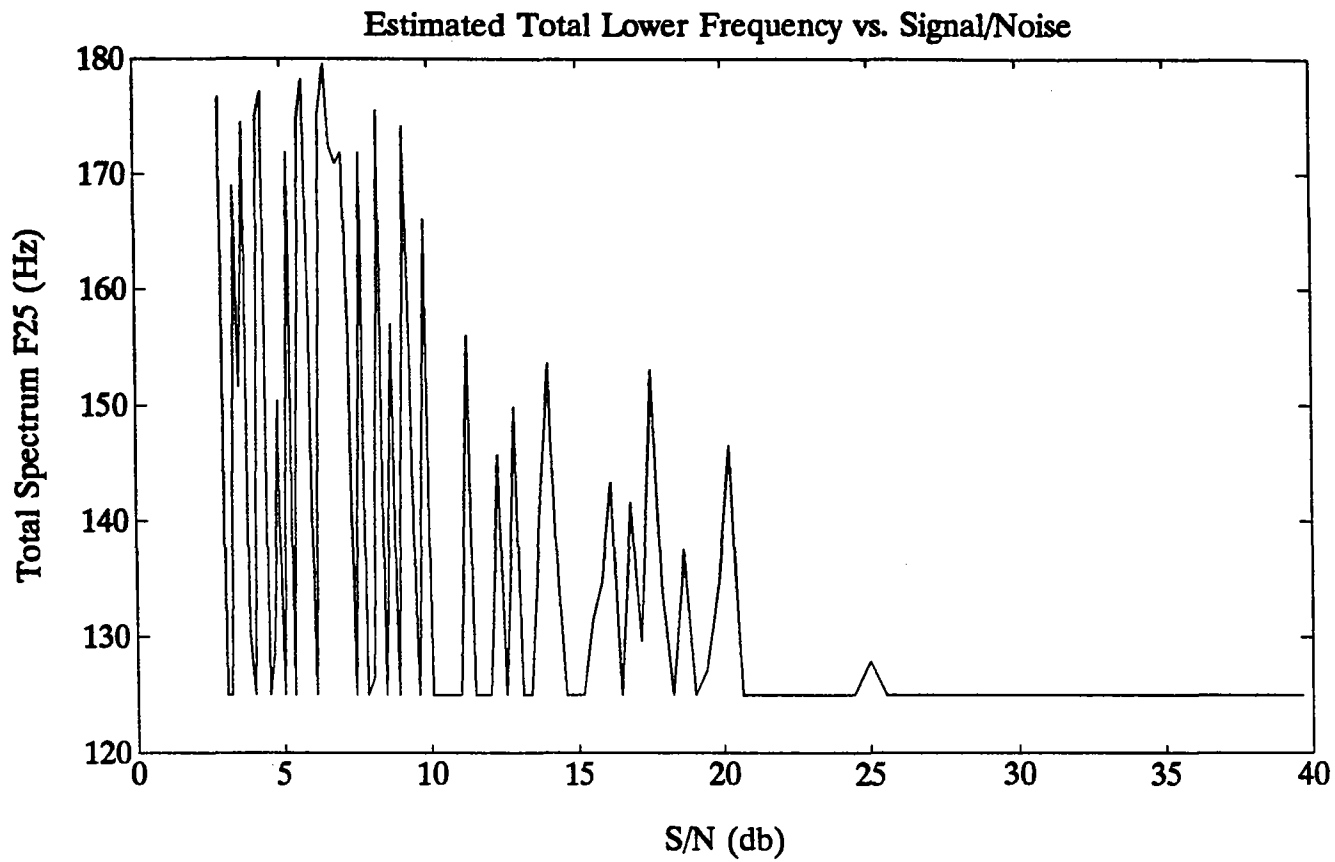Figure 11. Noise performance of Upsweep Proportion (eq. 11). The vertical axis scores Upsweep Proportion (values of 0 ↔ 1 are possible); the horizontal axis scores signal/noise. We subjectively label this as high variance, small trend.
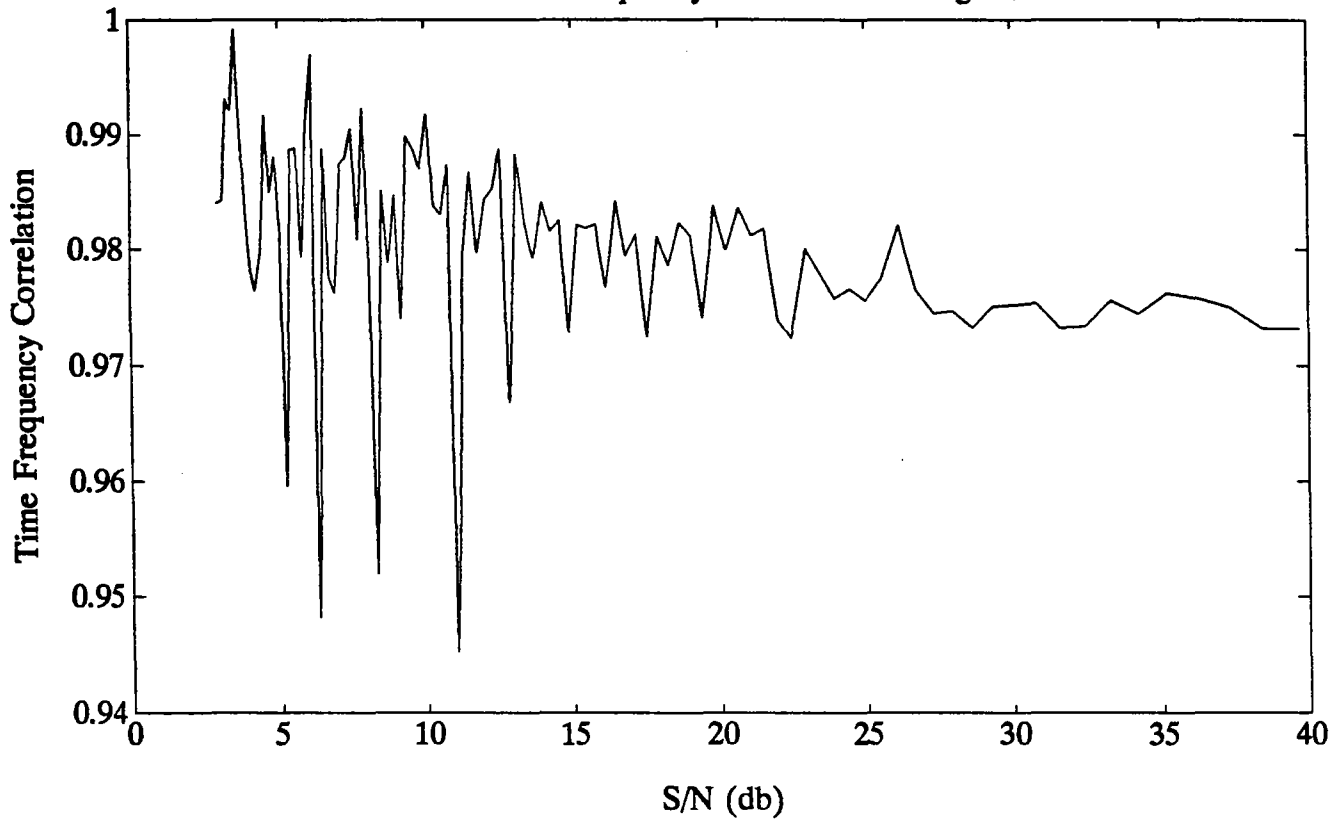
Figure 12. Noise performance of Time Frequency Correlation (eq. 12). The vertical axis scores Time Frequency Correlation coefficients (values $-1 \leftrightarrow 1$ possible); the horizontal axis scores signal/noise. We subjectively label this as low variance, small trend.
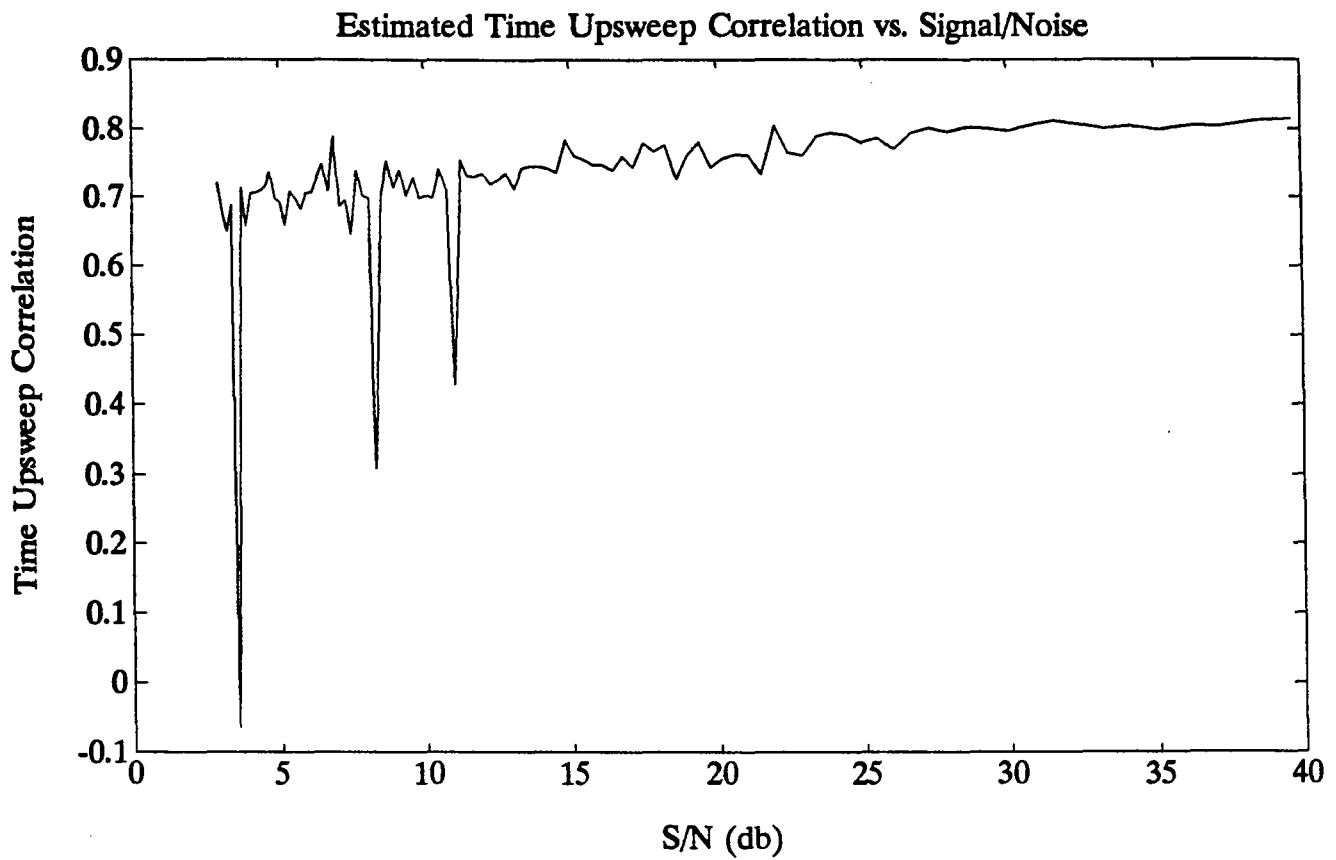
Figure 13. Noise performances of Time Upsweep Correlation (eq. 13). The vertical axis scores Time Upsweep Correlation coefficients (values of −1 ↔ 1 possible); the horizontal axis scores signal/noise. We subjectively score this as low variance, small trend.

Estimated Short-term Bandwidth Mean vs. Signal/Noise
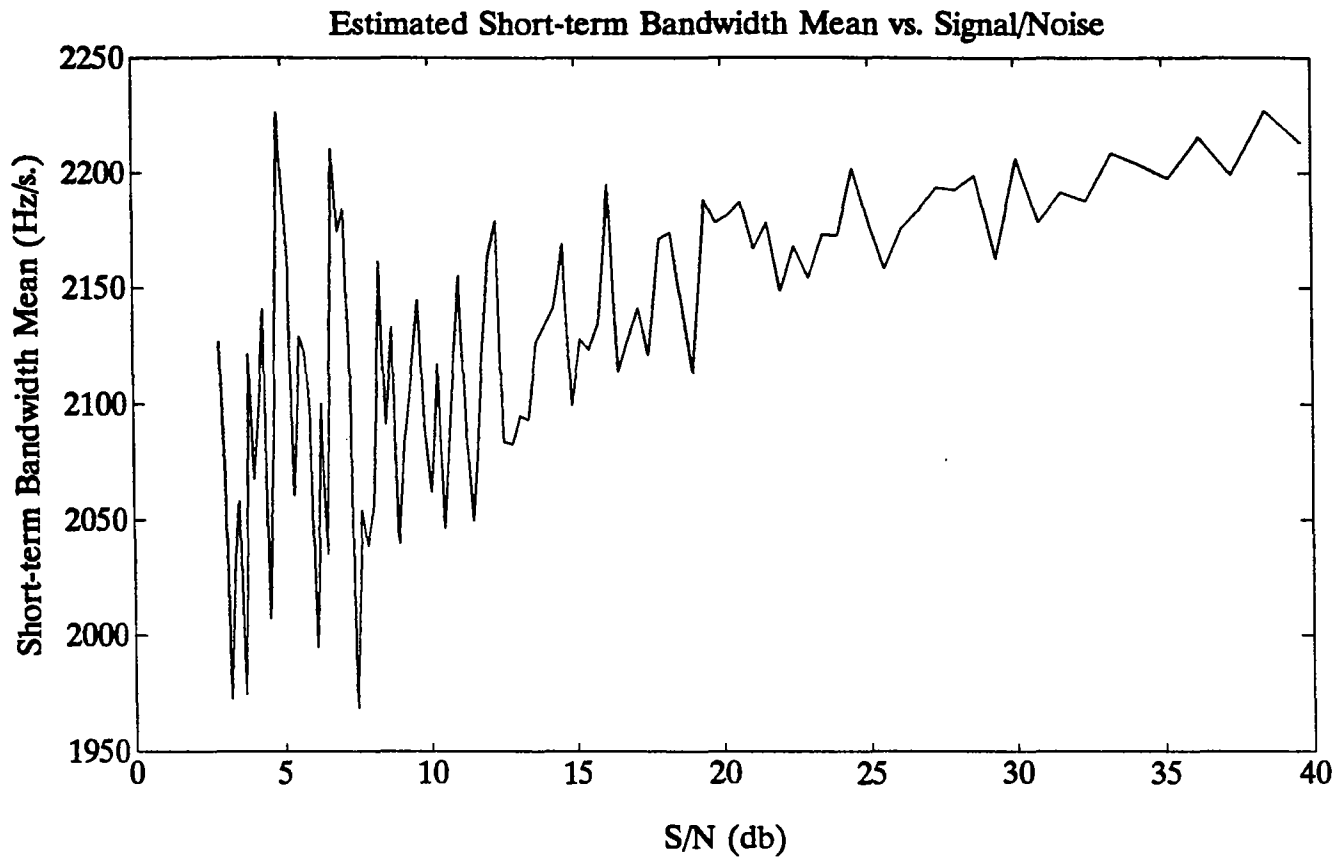
Figure 14. Noise performance of Short-term Bandwidth Mean (eq. 14). The vertical axis is Short-term Bandwidth Mean in Hz/s.; the horizontal axis scores signal/noise. We subjectively label this as low variance, large trend.

49

Estimated Short-term Spectral Concentration vs. Signal/Noise
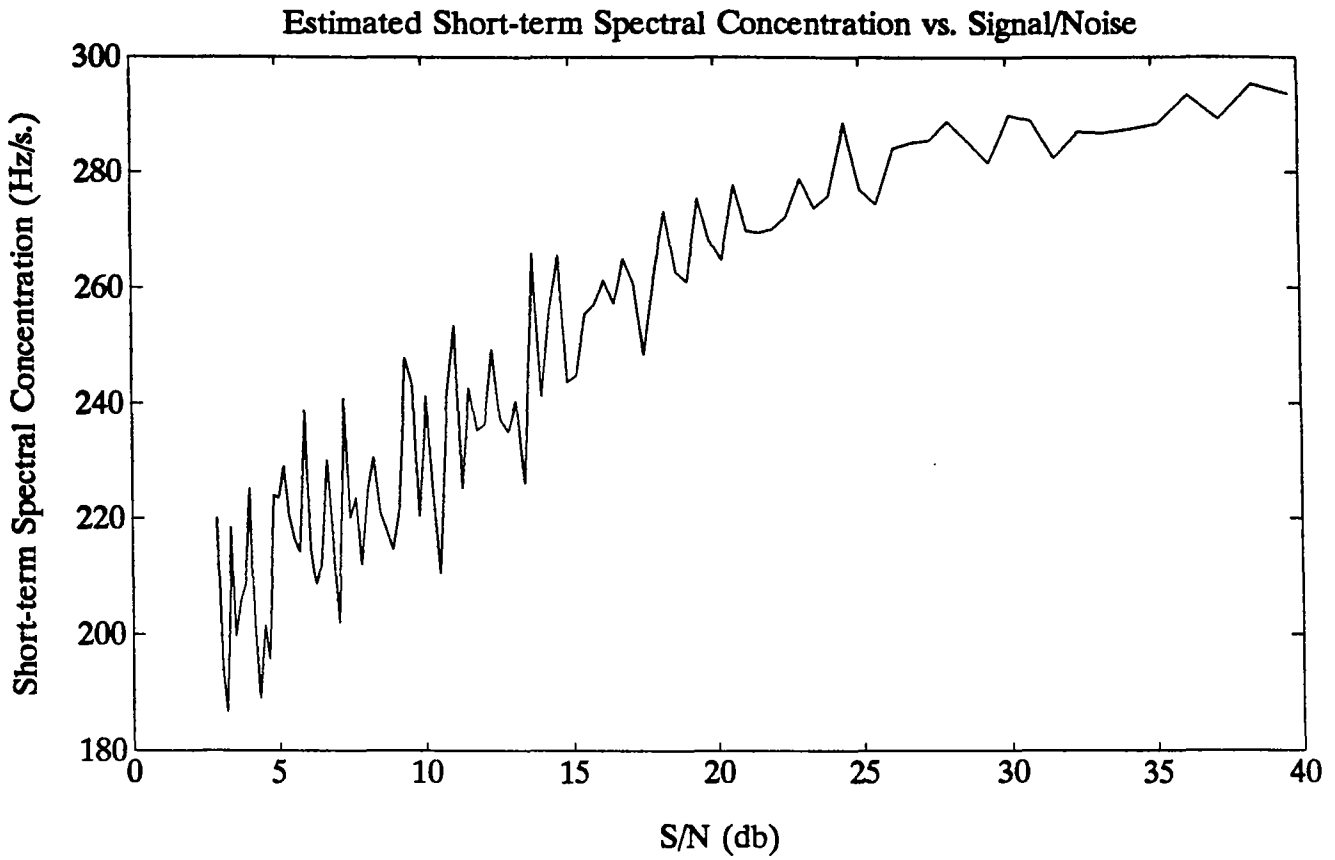
Figure 15. Noise performance of Short-term Spectral Concentration. The vertical axis scores Short-term Spectral Concentration in Hz/s.; the horizontal axis scores signal/noise. We subjectively label this as low variance, large trend.

50

Estimated Short-term Spectral Asymmetry vs. Signal/Noise
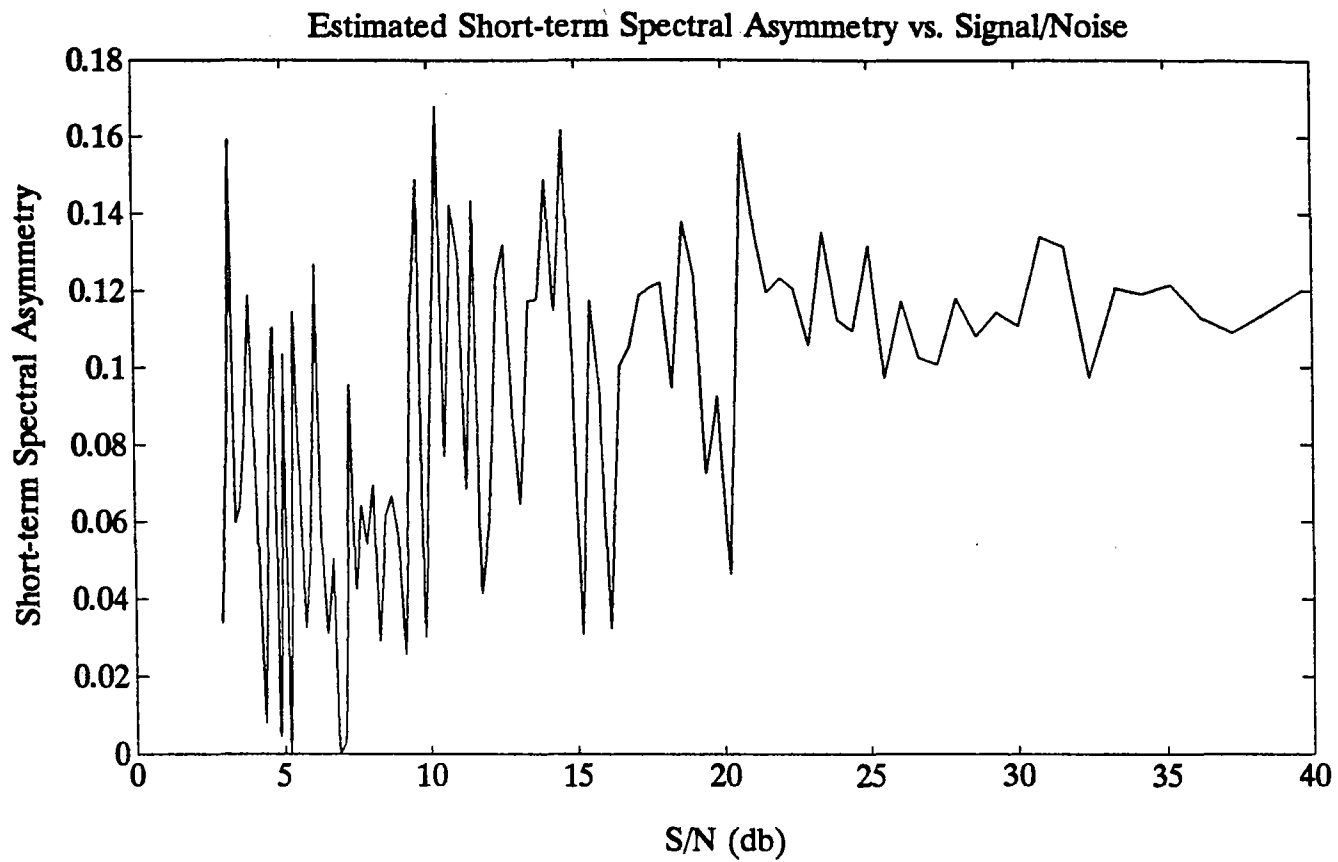
Figure 16. Noise performance of Short-term Spectral Asymmetry (eq. 16). The vertical axis scores Short-term Spectral Asymmetry (values of $0 \leftrightarrow 1$ are possible); the horizontal axis scores signal/noise. We subjectively label this as high variance, small trend.

51

Estimated Total Upper Frequency vs. Signal/Noise
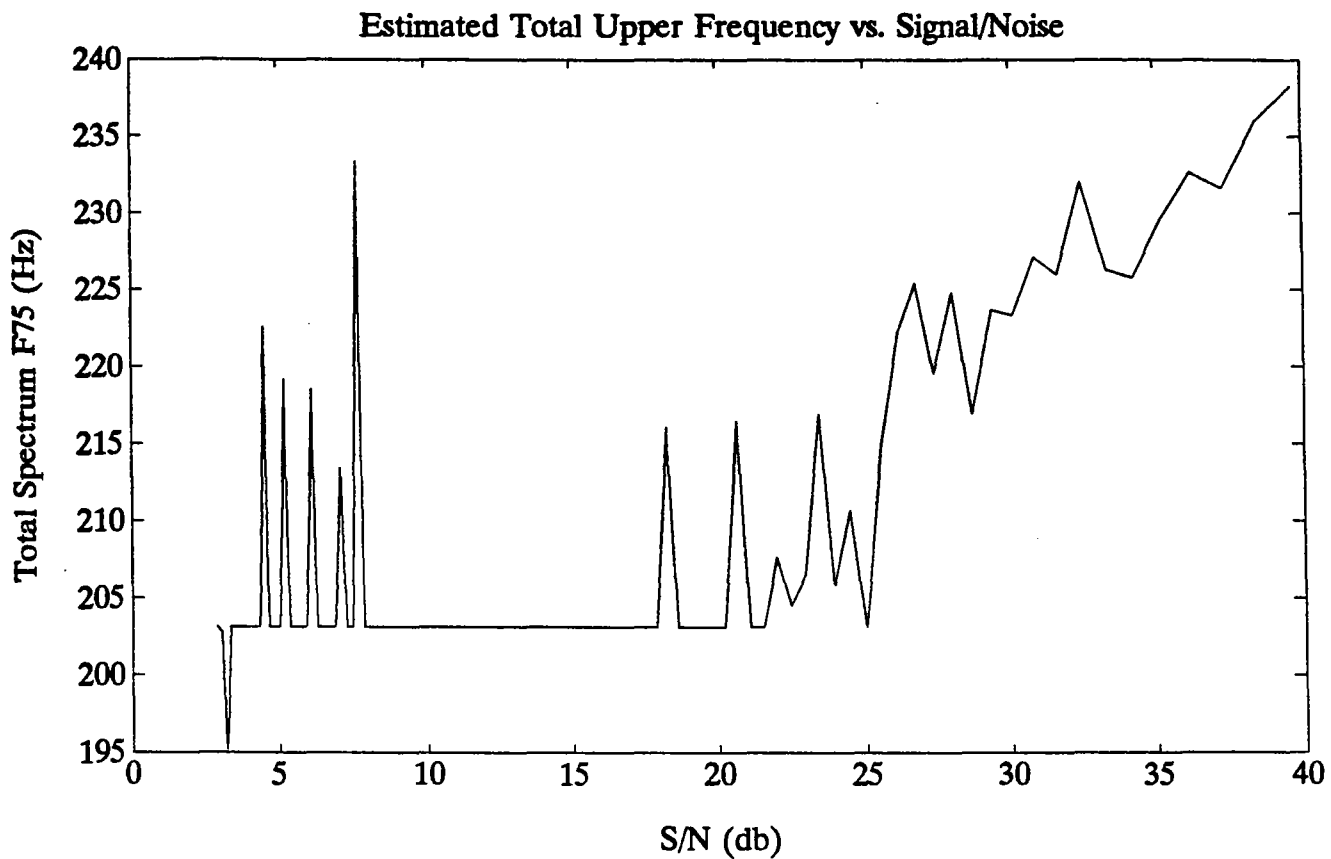
Figure 17:. Noise performance of Total Upper Frequency (see p. 5). The vertical axis scores Total Upper Frequency in Hz; the horizontal axis scores signal/noise. We subjectively label this as low variance, large trend.

Estimated Total Lower Frequency vs. Signal/Noise

Total Spectrum F25 (Hz)

180
170
160
150
140
130
120

0   5   10   15   20   25   30   35   40

S/N (db)

Figure 18. Noise performance of Total Lower Frequency (see p. 5). The vertical axis score Total Lower Frequency in Hz; the horizontal axis scores signal/noise. We subjectively label this as low variance, large trend.

Estimated Total Spectrum Concentration vs. Signal/Noise
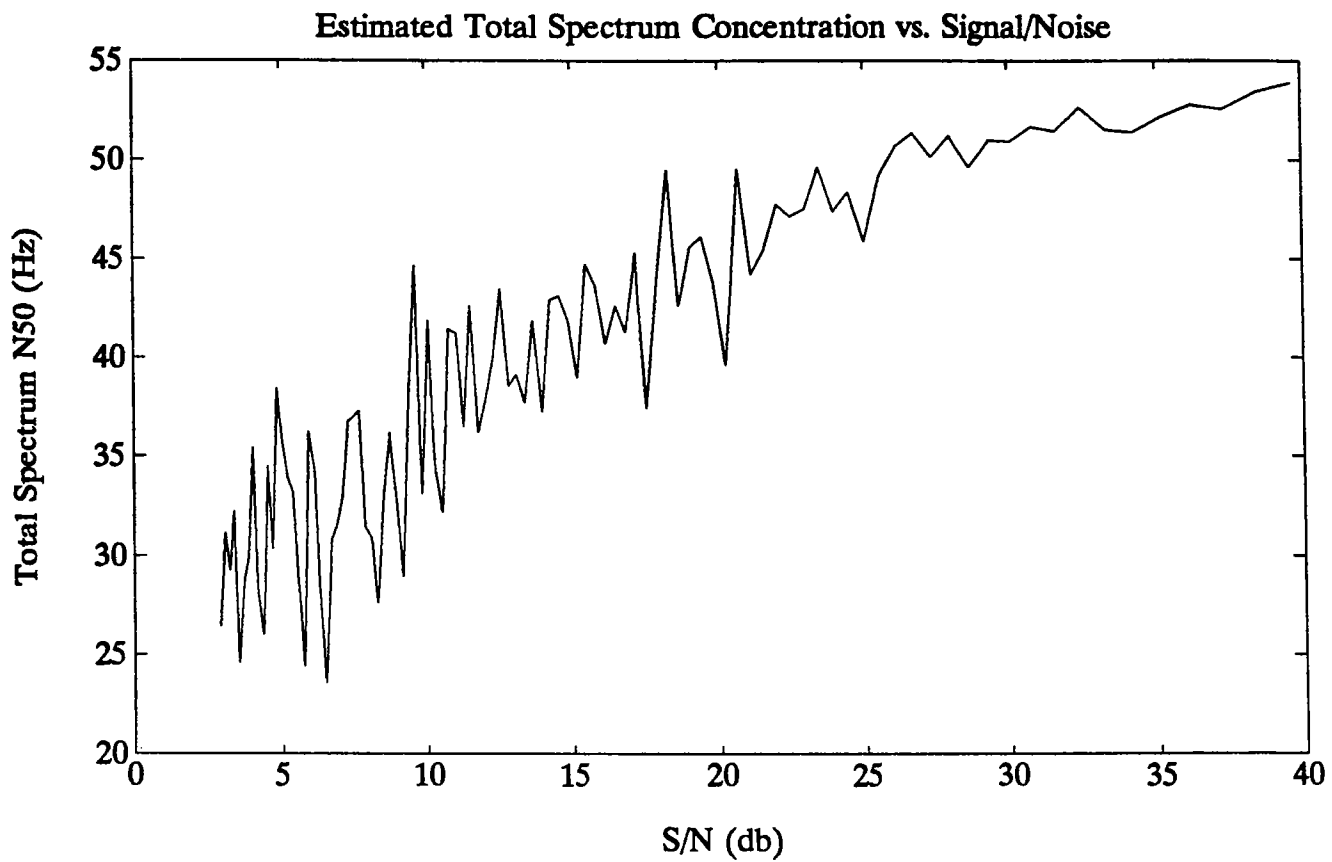
Figure 19. Noise performance of Total Spectrum Concentration (see p. 5). The vertical axis scores Total Spectrum Concentration in Hz; the horizontal axis scores signal/noise. We subjectively label this low variance, large trend.

Estimated Modal Upper Frequency vs. Signal/Noise
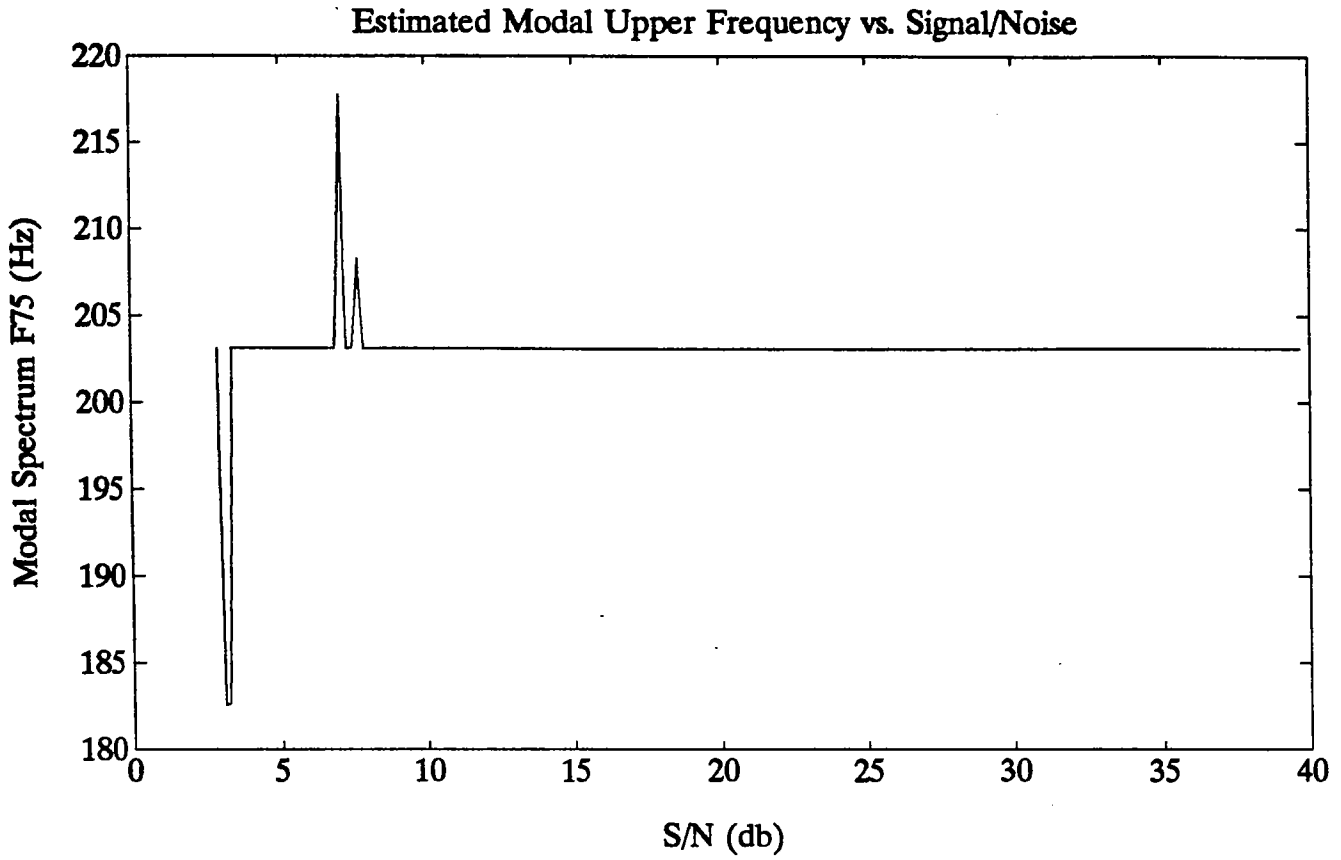
Figure 20. Noise performance of Modal Upper Frequency (see p. 5). The vertical axis scores Modal Upper Frequency in Hz; the horizontal axis scores signal/noise. We subjectively label this low variance, small trend.

Estimated Modal Lower Frequency vs. Signal/Noise
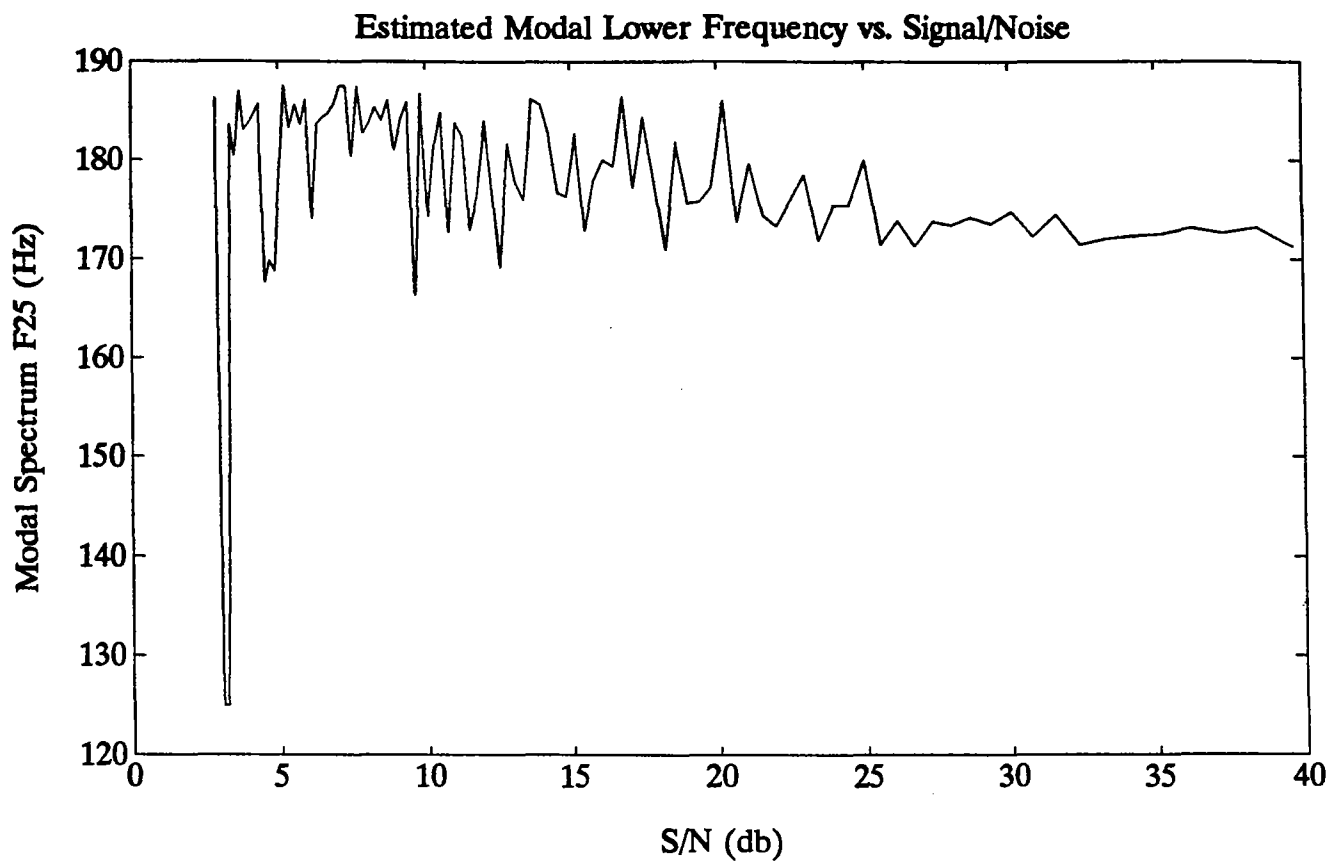
Figure 21. Noise performance of Modal Lower Frequency (see p. 5). The vertical axis scores Modal Lower Frequency in Hz; the horizontal axis scores signal/noise. We subjectively label this low variance, small trend.
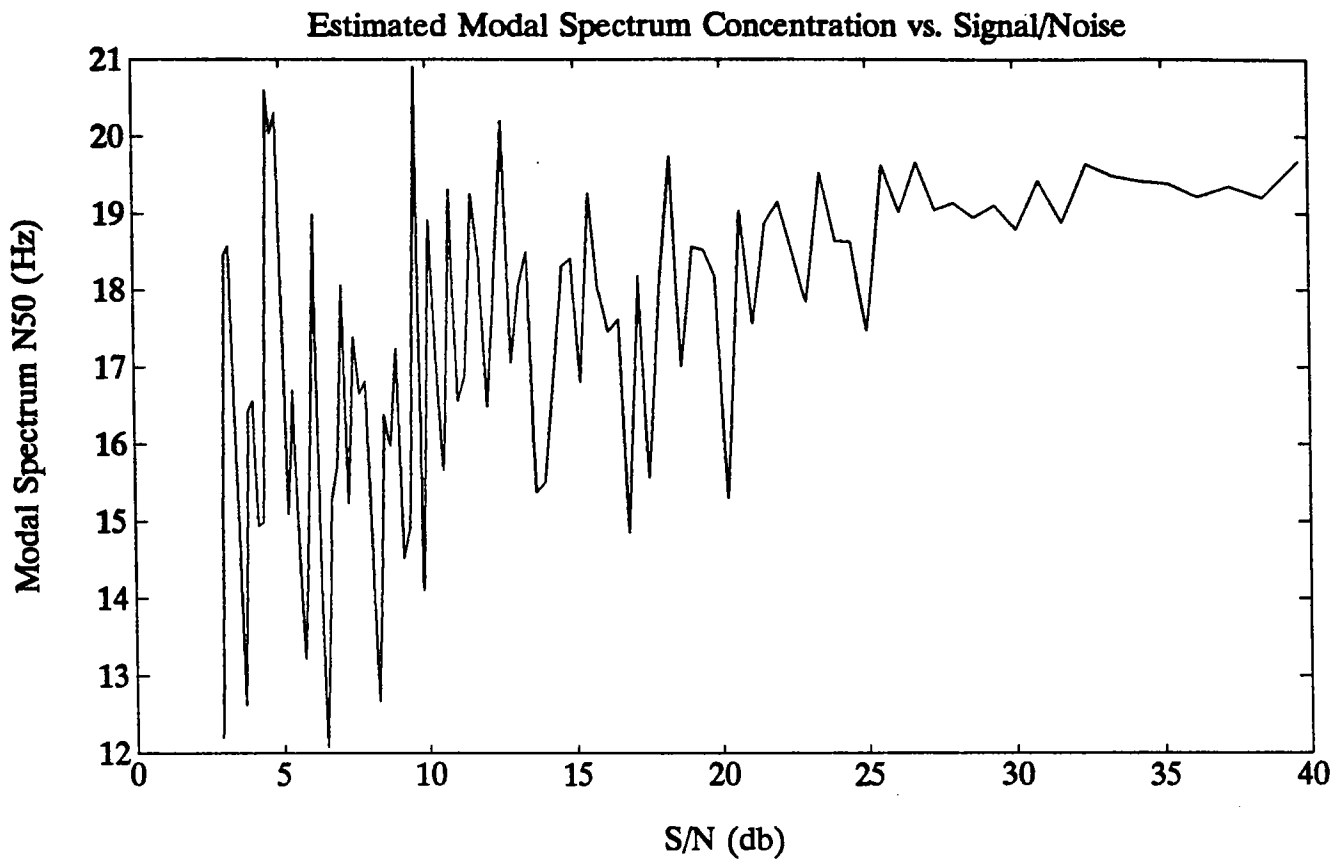
Figure 22. Noise performance of Modal Spectrum Concentration (see p. 5). The vertical axis scores Modal Spectrum Concentration in Hz; the horizontal axis scores signal/noise. We subjectively label this low variance, small trend.

57

Estimated Total Spectrum Median vs. Signal/Noise
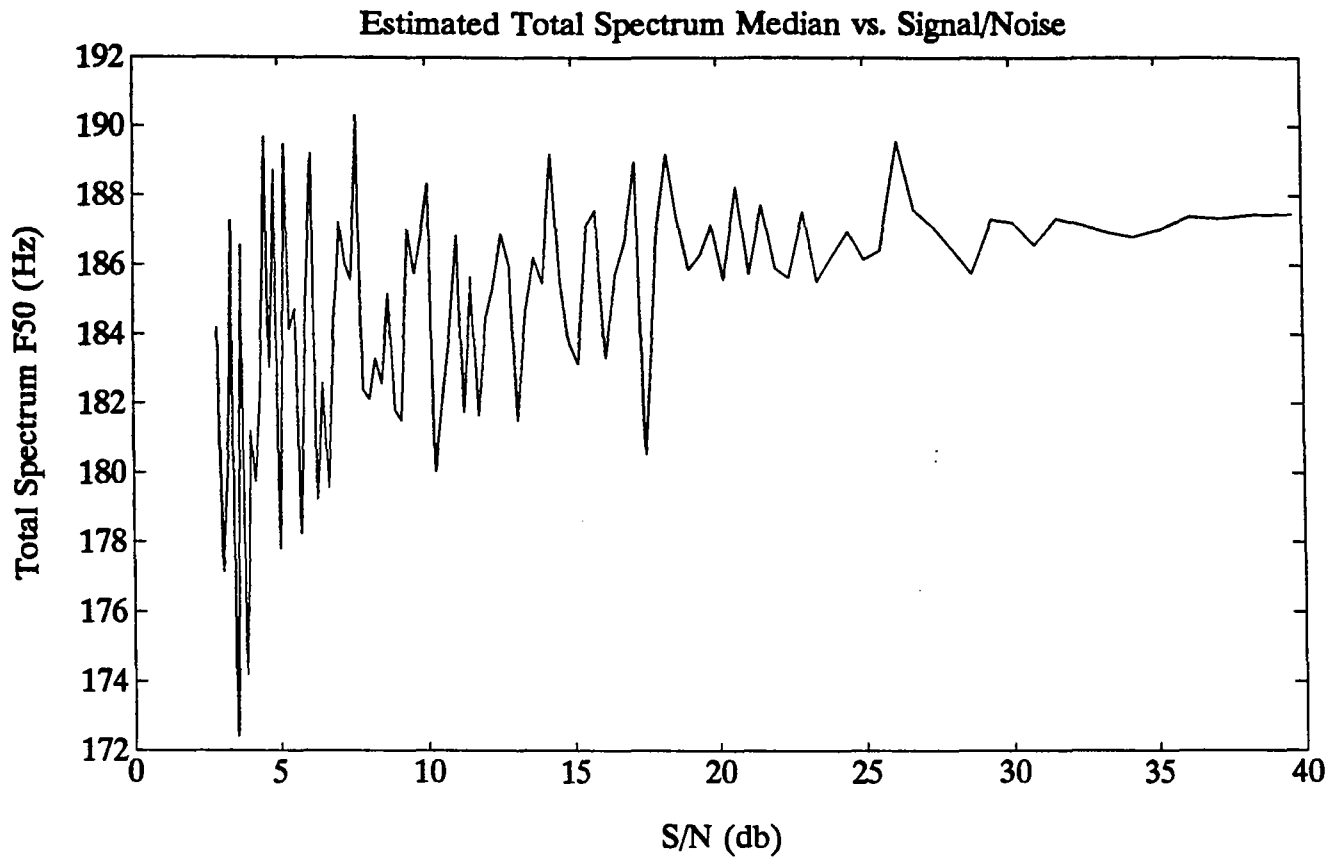
Figure 23. Noise performance of Total Spectrum Median (see p. 5). The vertical axis scales Total Spectrum Median in Hz; the horizontal axis scales signal/noise. We subjectively label this low variance, small trend.

Estimated Modal Spectrum Median vs. Signal/Noise
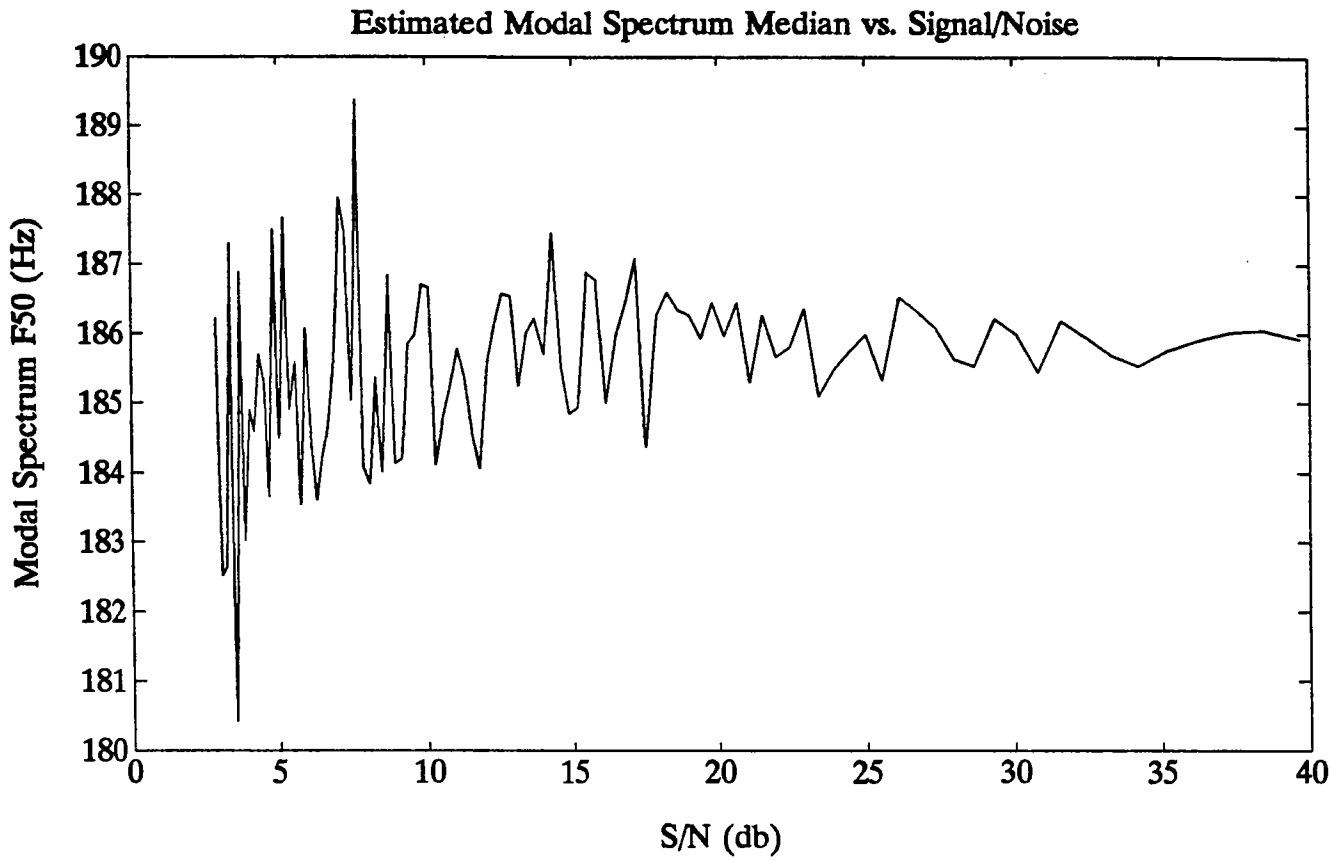
Figure 24. Noise performance of Modal Spectrum Median (see p. 5). The vertical axis scales Modal Spectrum Median in Hz; the horizontal axis scales signal/noise. We subjectively label this low variance, small trend.

59
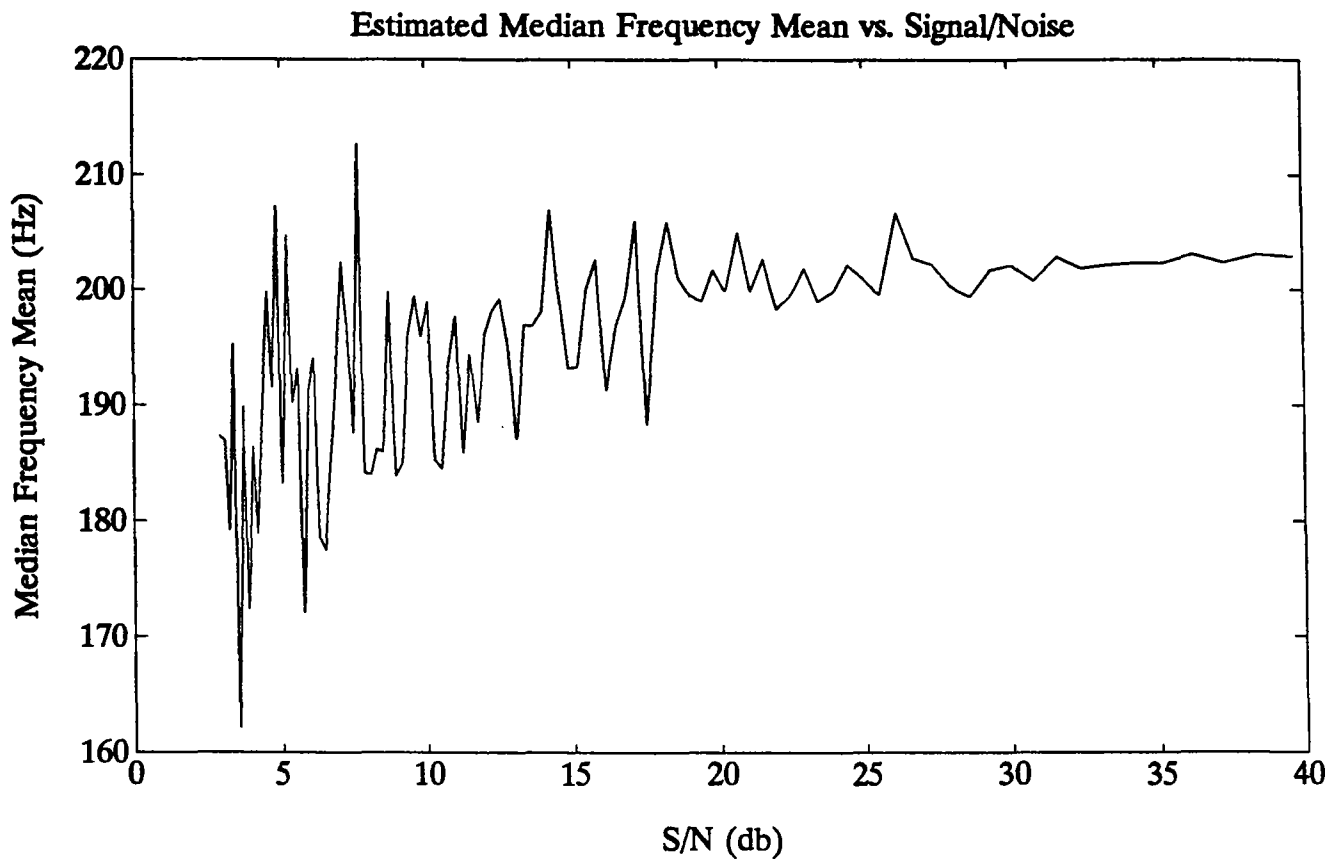
Figure 25. Noise performance of Median Frequency Mean (eq. 17). The vertical axis scales Median Frequency Mean in Hz; the horizontal axis scales signal/noise. We subjectively label this as low variance, small trend.

Estimated Amplitude Frequency Correlation vs. Signal/Noise
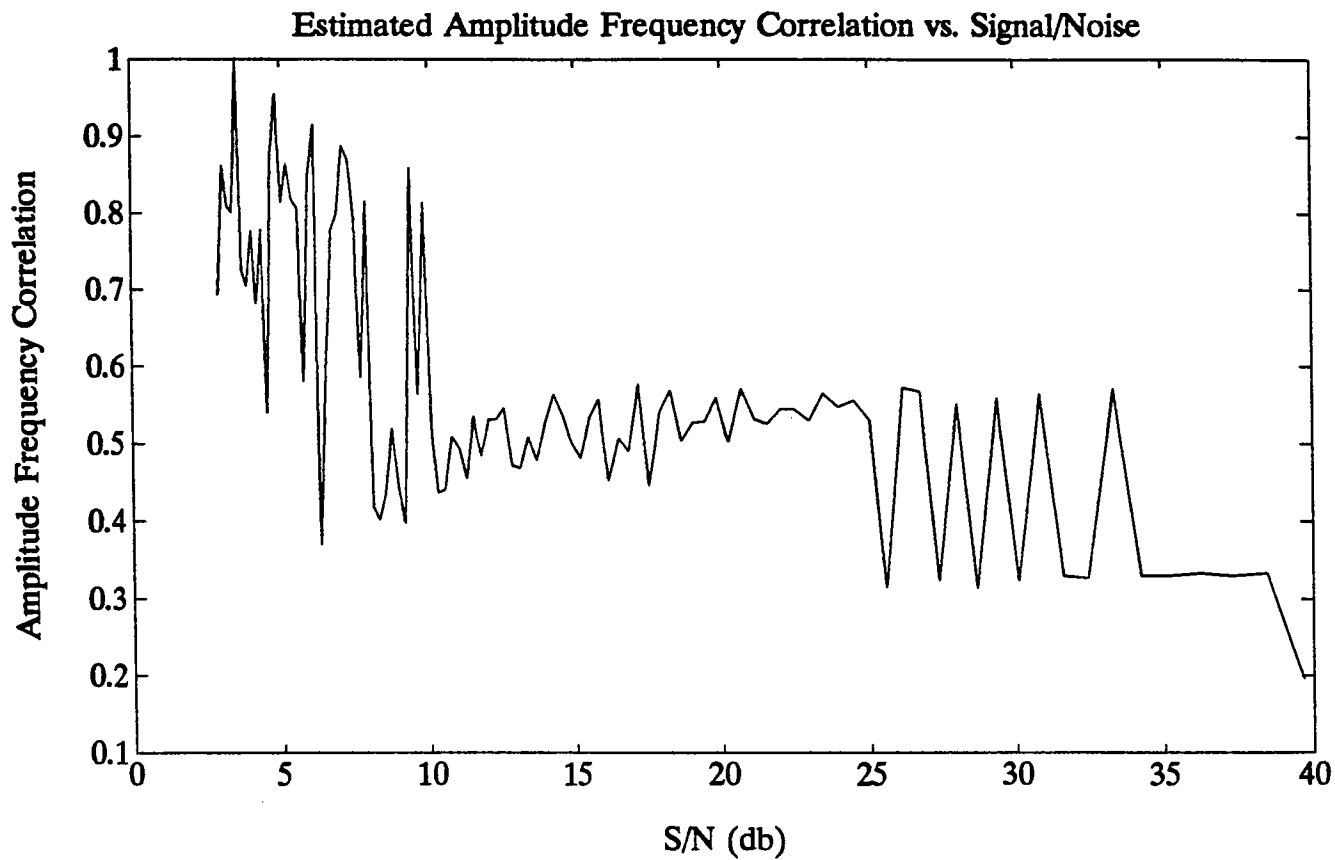
Figure 26. Noise performance of Amplitude Frequency Correlation (eq. 18). The vertical axis scales Amplitude Frequency Correlation coefficients (values of $-1 \leftrightarrow 1$ are possible); the horizontal axis scales signal/noise. We subjectively label this as high variance, large trend.
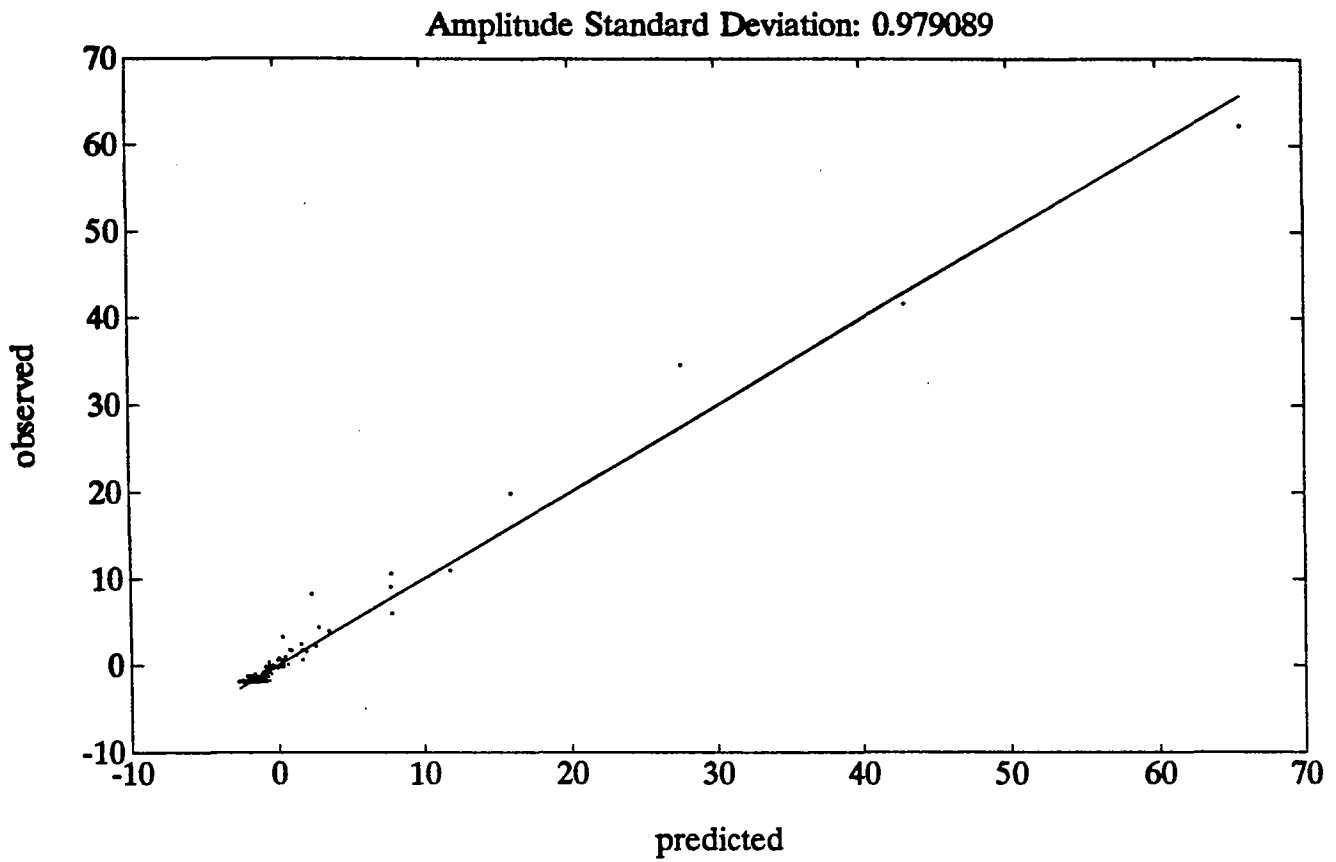
Estimated Amplitude Upsweep Correlation vs. Signal/Noise
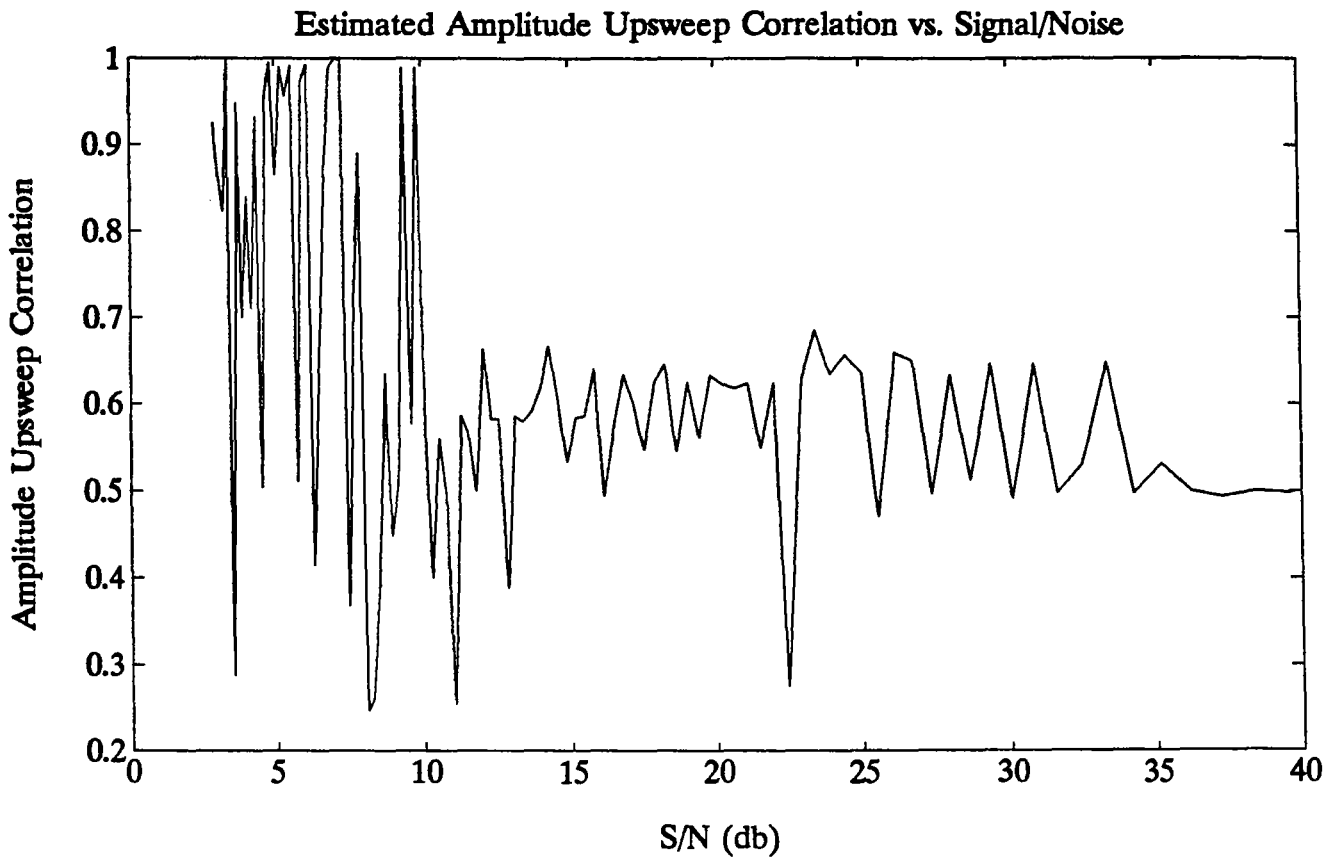
Figure 27. Noise performance of Amplitude Upsweep Correlation (eq. 19). The vertical axis scales Amplitude Upsweep Correlation coefficients (values of $-1 \leftrightarrow 1$ are possible); the horizontal axis scales signal/noise. We subjectively label this as high variance, large trend.

Figure 28. Linear prediction of Amplitude Standard Deviation based on 25 other variables (Signal Center not used for any regressions). The axes are in arbitrary units. The percent of variance explained by the linear regression is displayed in the title.
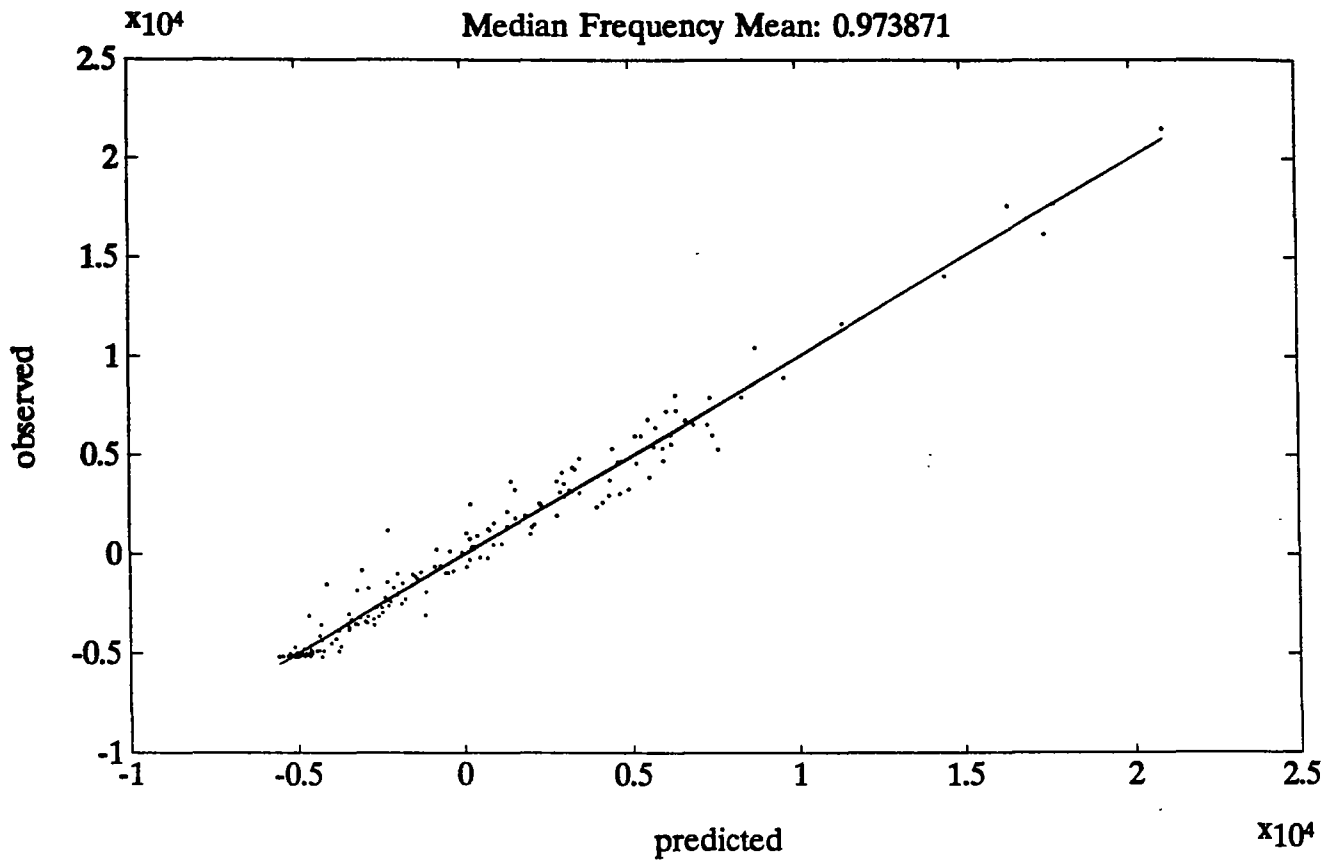
Figure 29. Linear prediction of Median Frequency Mean based on 24 other variables (Amplitude Standard Deviation already removed). The axes are in Hz, but the data are shifted such that the means are zero. The percent of variance explained by the linear regression is displayed in the title.
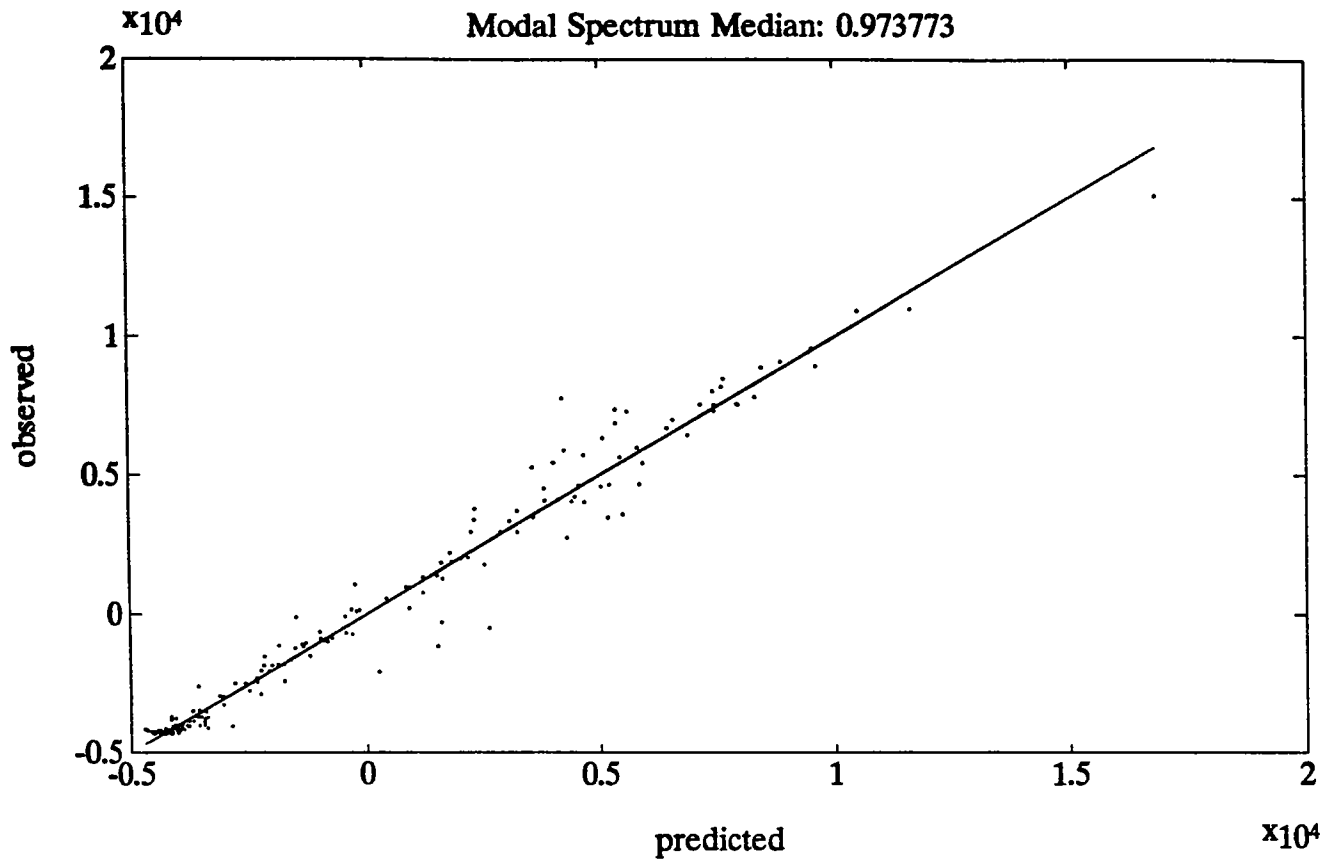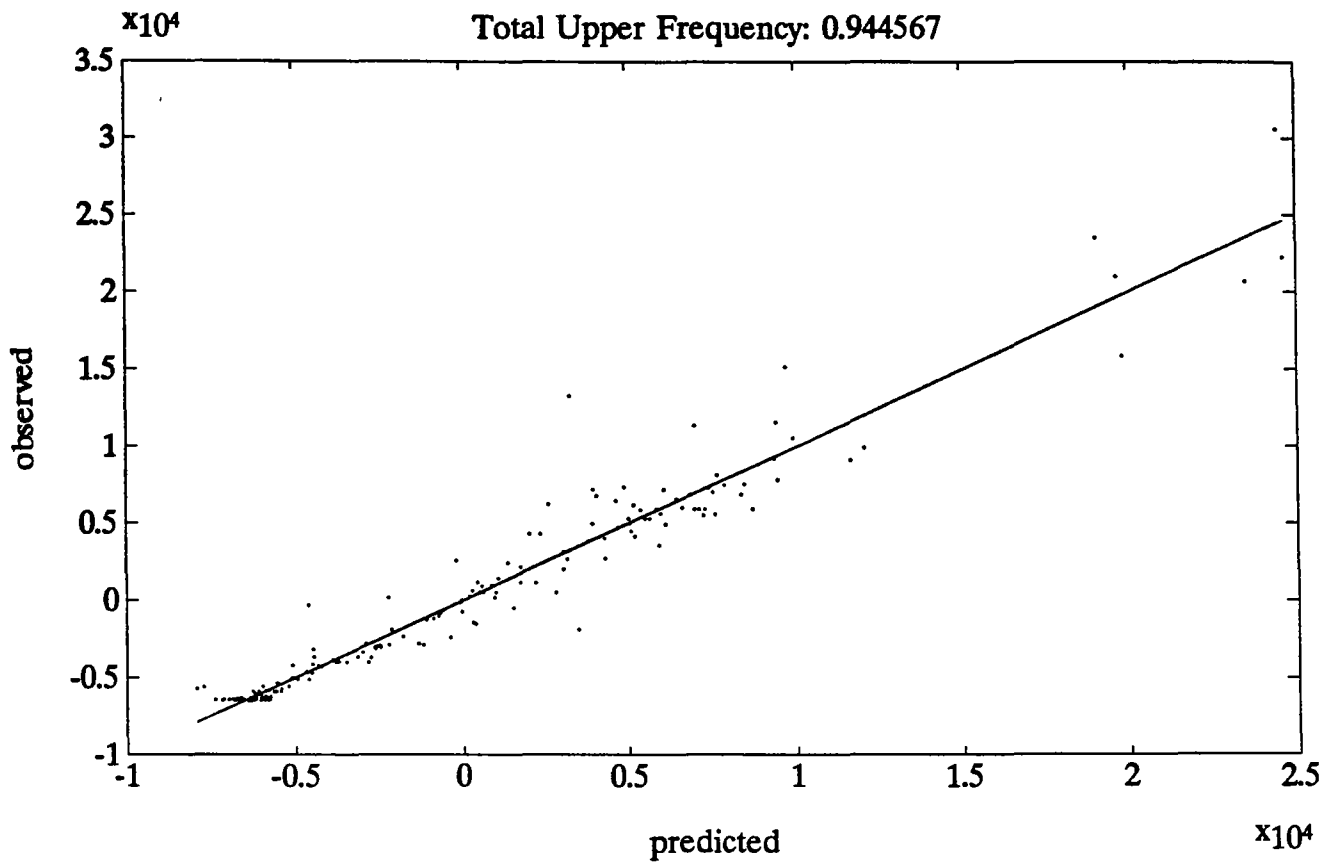
Figure 30. Linear prediction of Modal Spectrum Median based on 23 other variables (two variables previously removed). The axes are in Hz, but the data are shifted such that the means are zero. The percent of variance explained by the linear regression is displayed in the title.

Total Upper Frequency: 0.944567

Figure 31. Linear prediction of Total Upper Frequency based on 22 other variables (three variables previously removed). The axes are in Hz, but the data are shifted such that the means are zero. The percent of variance explained by the linear regression is displayed in the title.

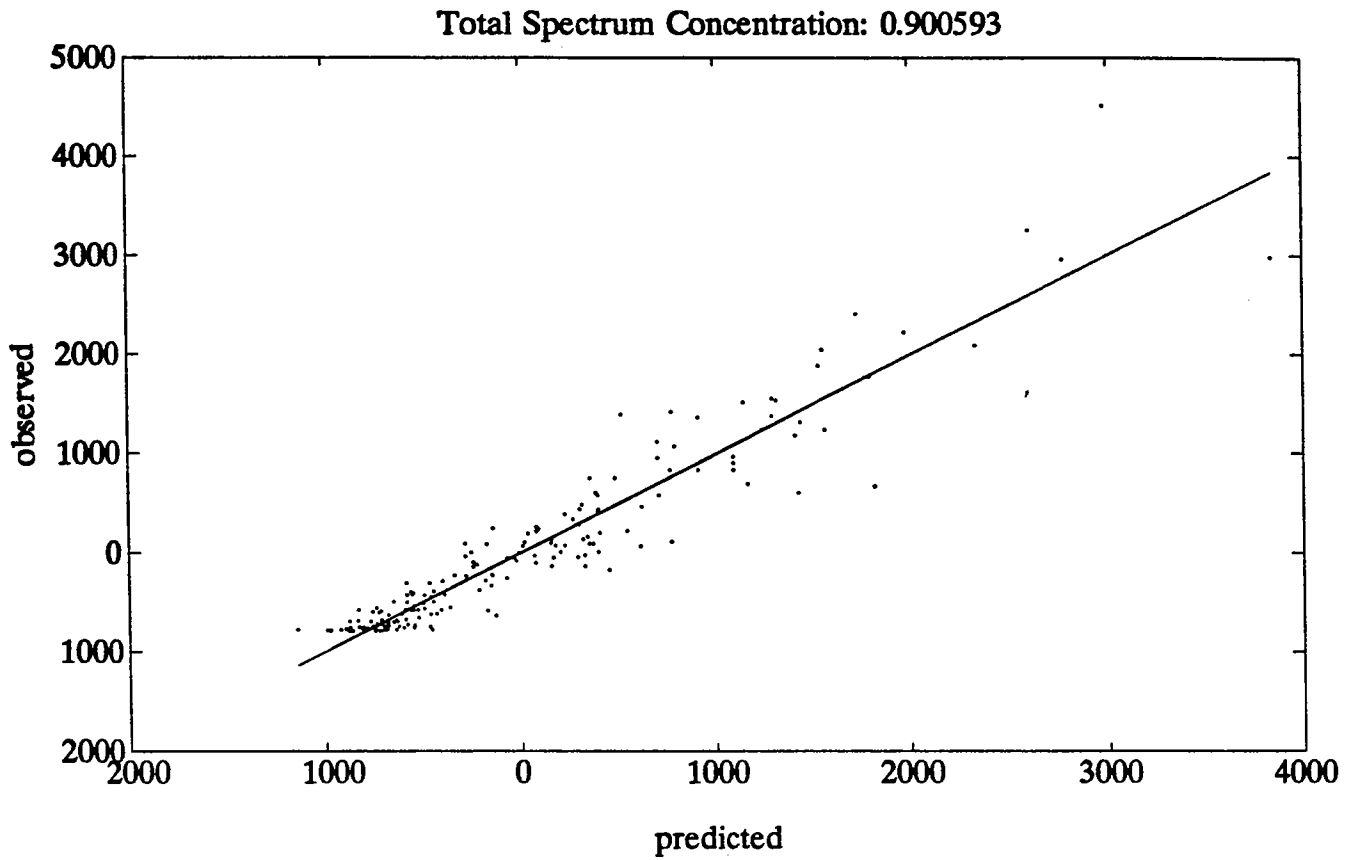**Total Spectrum Concentration: 0.900593**

Figure 32. Linear prediction of Total Spectrum Concentration based on 21 other variables (four variables previously removed). The axes are in Hz, but the data are shifted such that the means are zero. The percent of variance explained by the linear regression is displayed in the title.
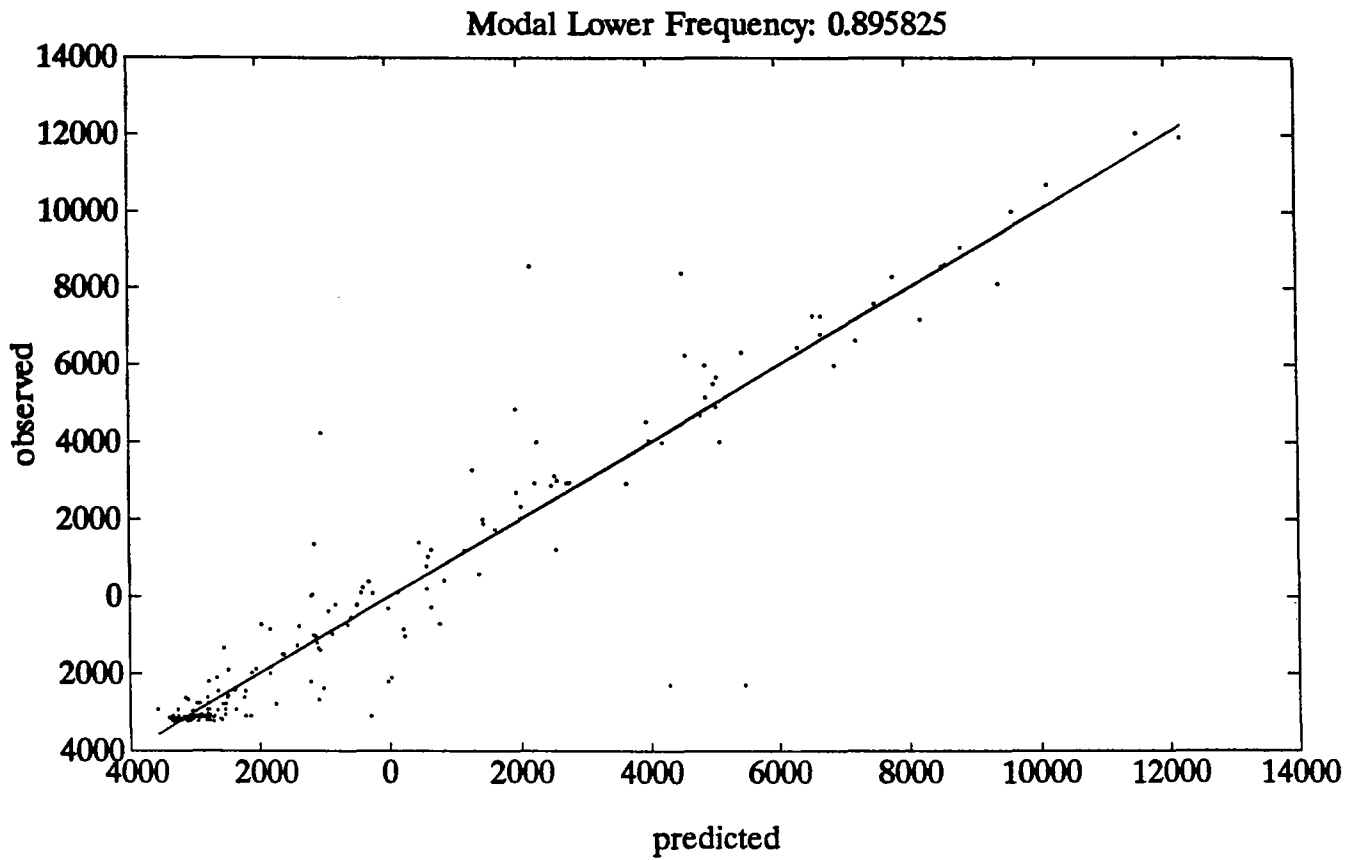
Figure 33. Linear prediction of Modal Lower Frequency based on 20 other variables (five variables previously removed). The axes are in Hz, but the data are shifted such that the means are zero. The percent of variance explained by the linear regression is displayed in the title.
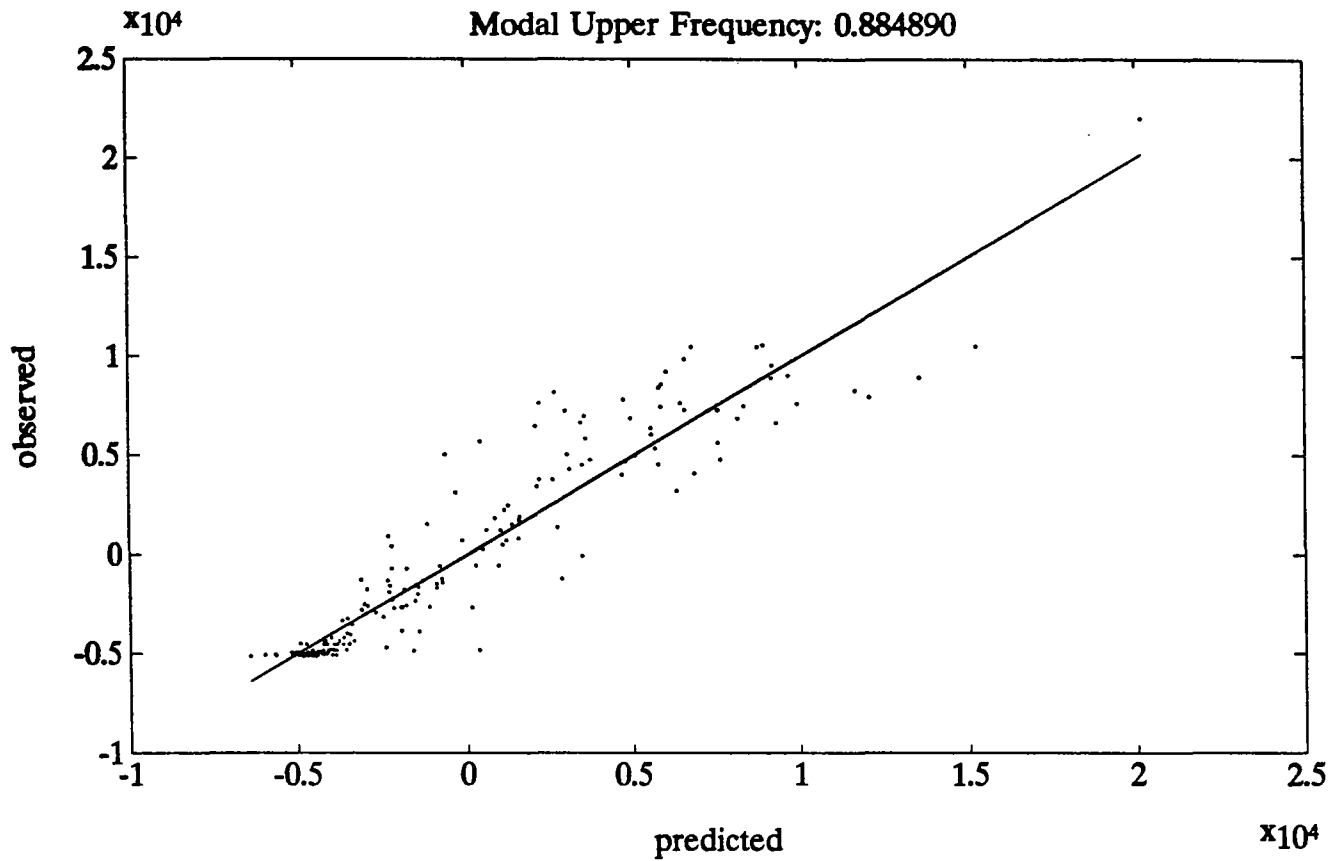
Figure 34. Linear prediction of Modal Upper Frequency based on 19 other variables (six variables previously removed). The axes are in Hz, but the data are shifted such that the means are zero. The percent of variance explained by the linear regression is displayed in the title.
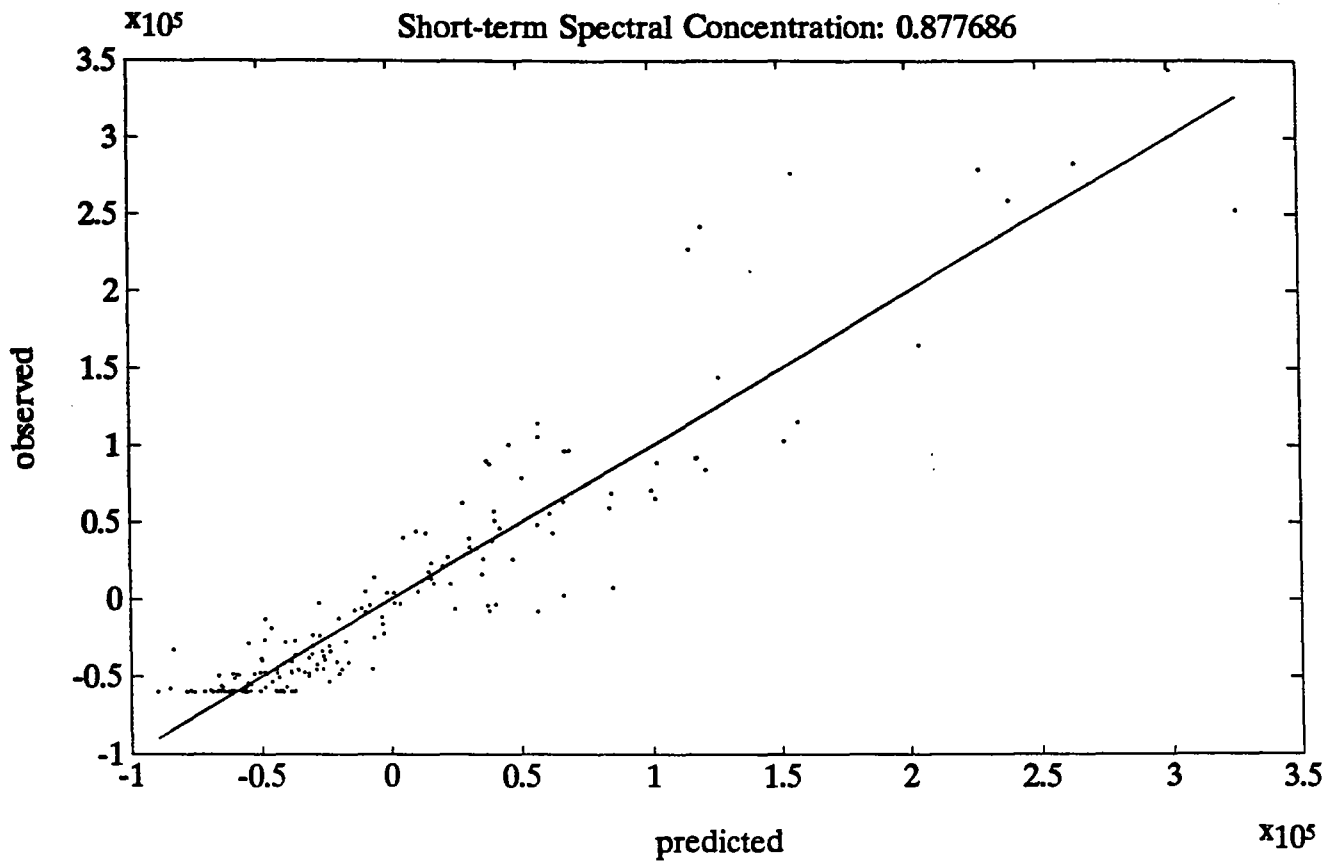
Figure 35. Linear prediction of Short-term Spectral Concentration based on 18 other variables (seven variables previously removed). The axes are in Hz, but the data are shifted such that the means are zero. The percent of variance explained by the linear regression is displayed in the title.
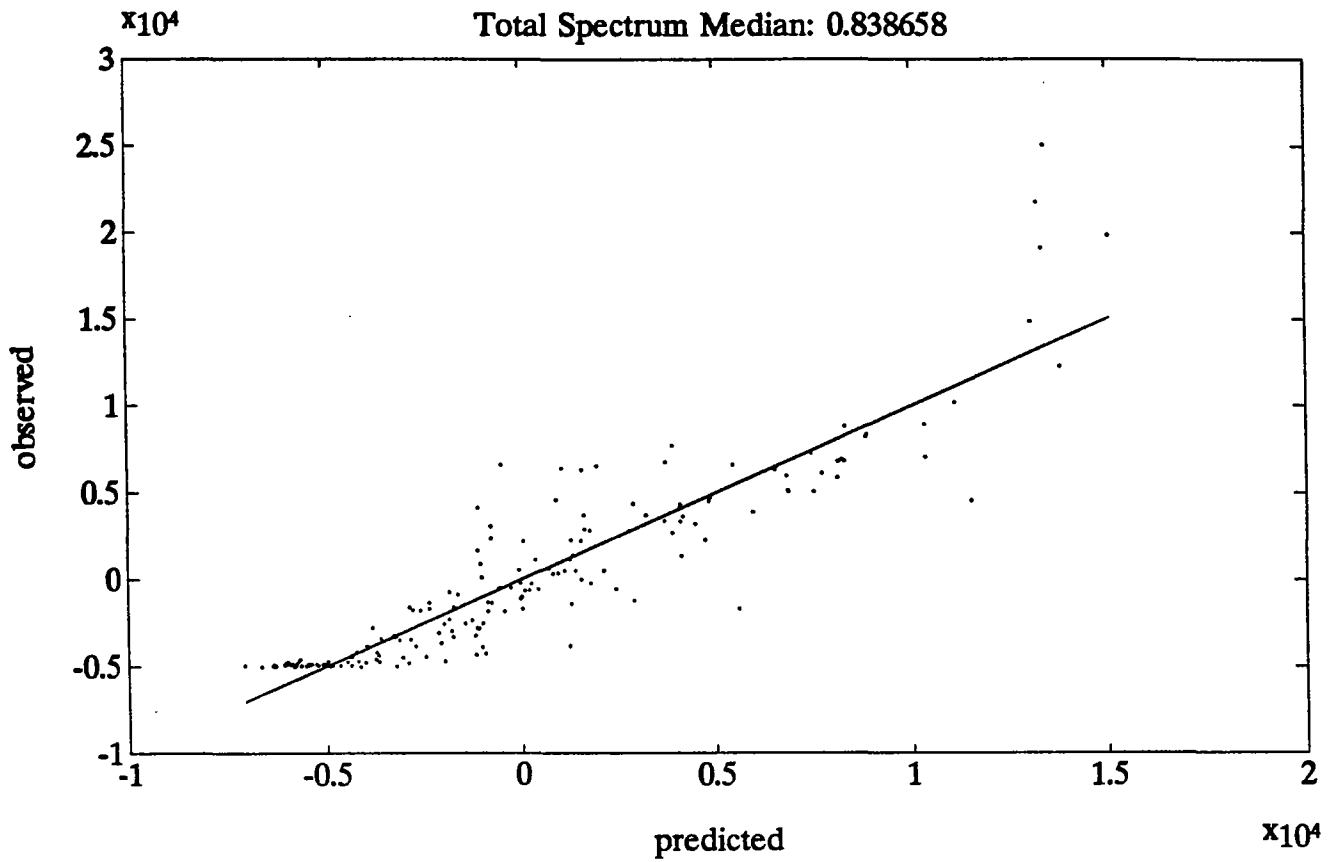
Figure 36. Linear prediction of Total Spectrum Median based on 17 other variables (eight variables previously removed). The axes are in Hz, but the data are shifted such that the means are zero. The percent of variance explained by the linear regression is displayed in the title.
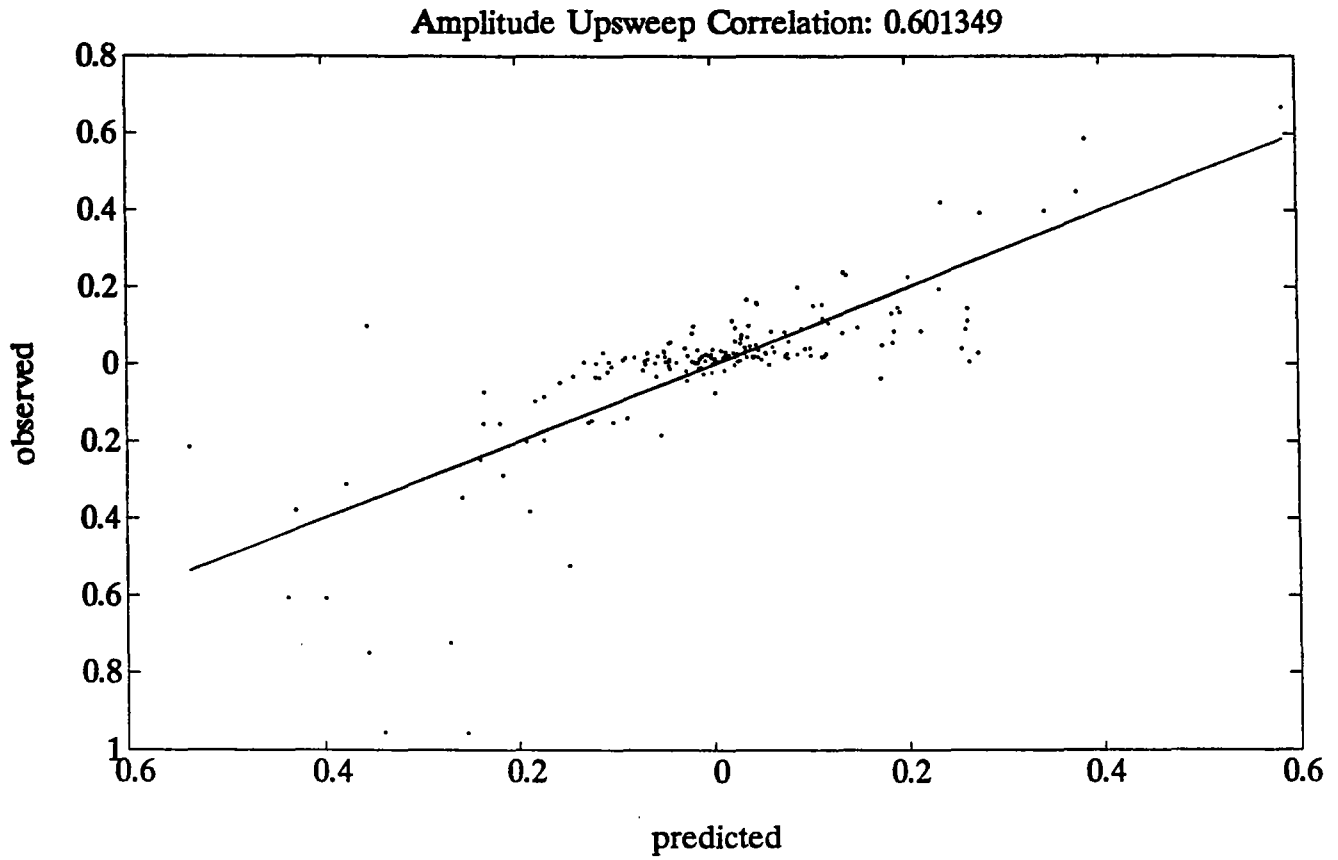
Amplitude Upsweep Correlation: 0.601349

Figure 37. Linear prediction of Amplitude Upsweep Correlation based on 16 other variables (nine variables previously removed). The axes are correlation coefficients (values $-1 \leftrightarrow 1$ possible), but the data are shifted such that the means are zero. The percent of variance explained by the linear regression is displayed in the title.

Figure 38. Plot of sound samples from six baleen whale species. The horizontal axis scales each sample's score on the first principal component (which basically reflects center frequency). The vertical axis scales each sample's score on the second principal component (not easily interpreted).

73

Acoustic Feature Analysis: Baleen Whales

Legend:
- ▷ B. acutorostrata
- ○ B. physalus
- × B. mysticetus
- + C. marginata
- □ E. glacialis
- ✳ E. australis

Principal Component 2 (vertical axis): 0.10, 0.05, 0.00, -0.05, -0.10, -0.15, -0.13

Principal Component 1 (horizontal axis): -0.13, -0.11, -0.09, -0.07, -0.05
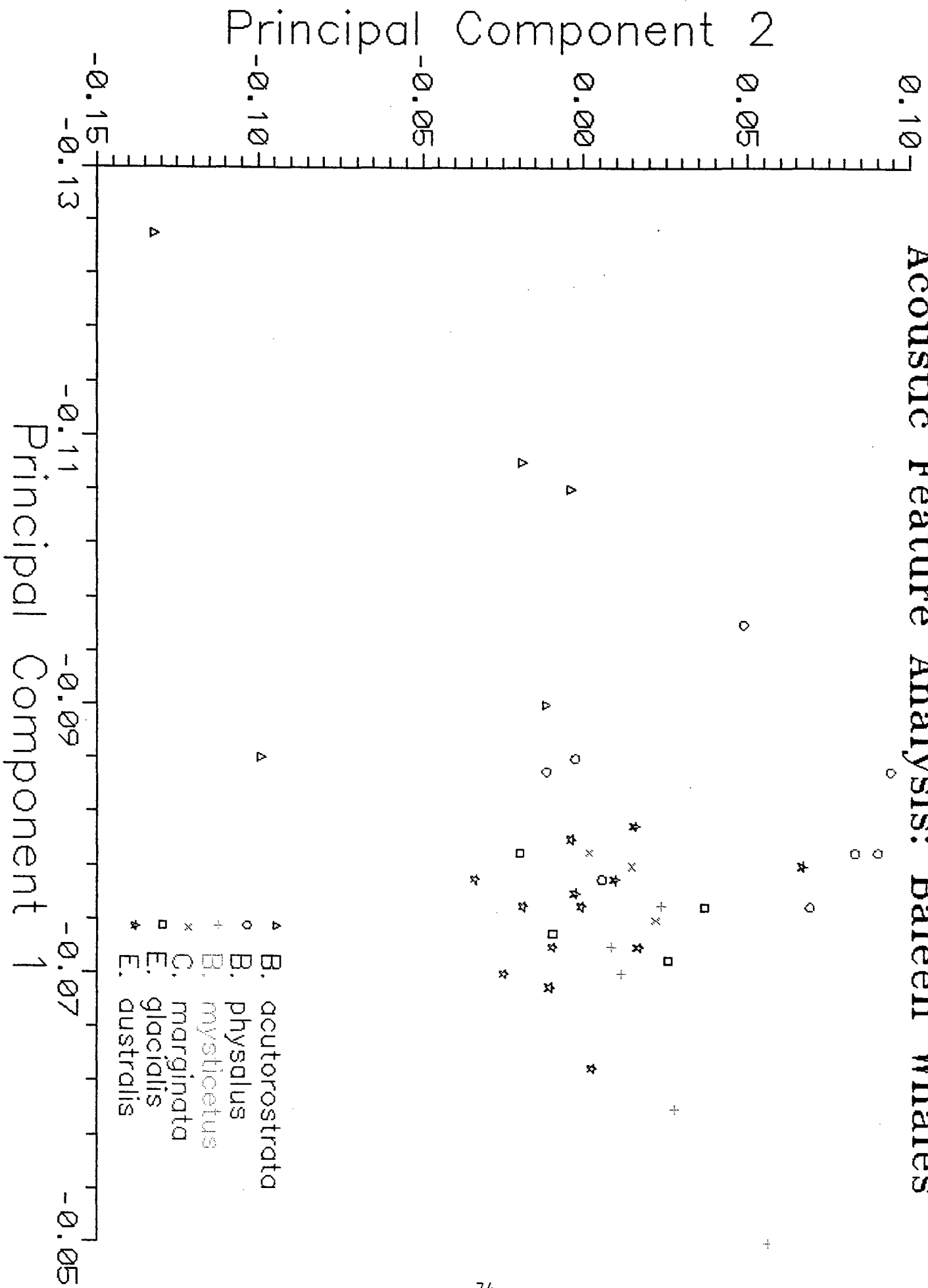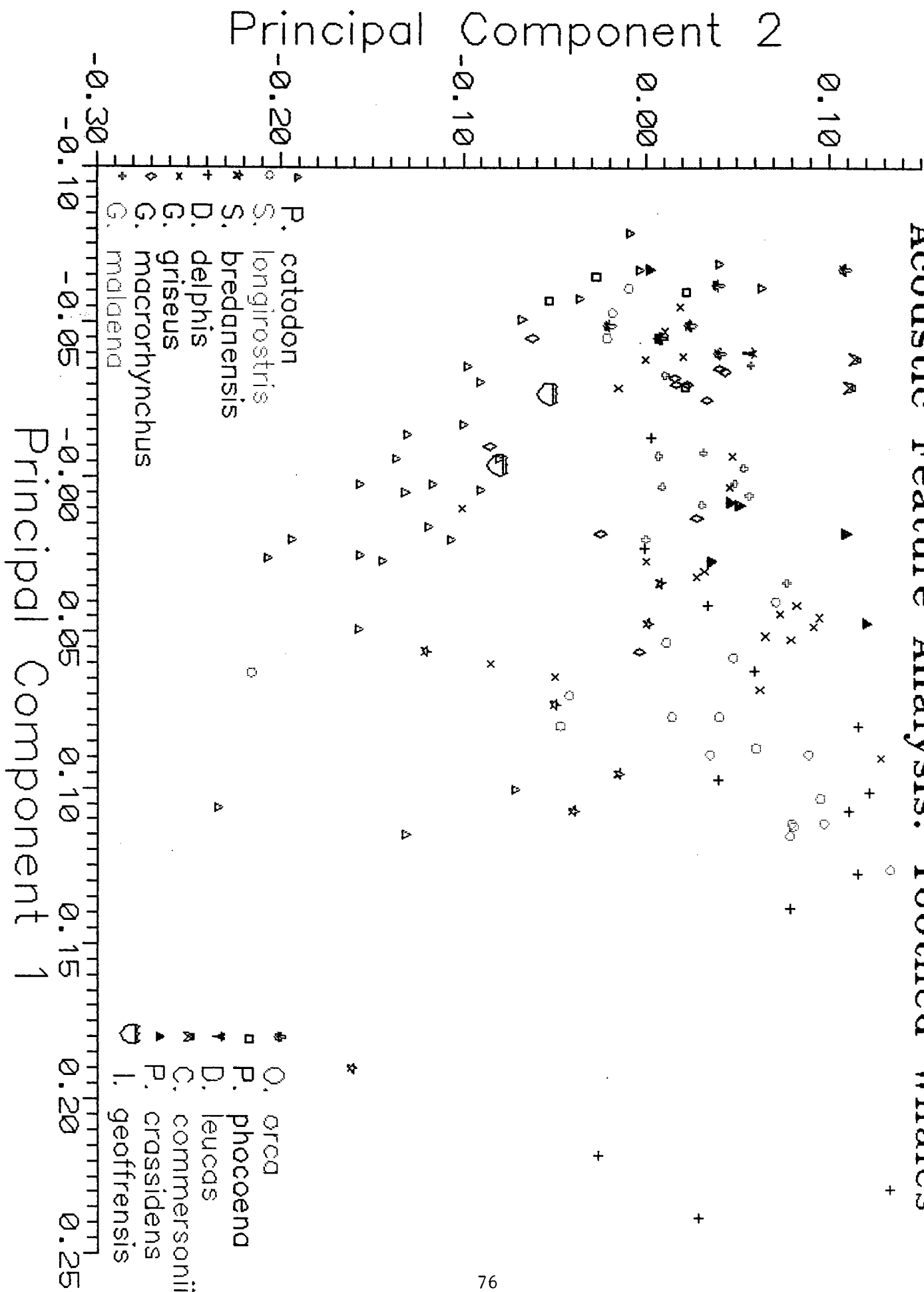
Figure 39. Plot of sound samples from thirteen toothed whale species. The horizontal axis scales each sample's score on the first principal component (which basically reflects center frequency). The vertical axis scales each sample's score on the second principal component (not easily interpreted).

Acoustic Feature Analysis: Toothed Whales

Legend:
- ▷ P. catodon
- ○ S. longirostris
- ✳ S. bredanensis
- + D. delphis
- × G. griseus
- ◇ G. macrorhynchus
- ✦ G. malaena
- ✸ O. orca
- □ P. phocoena
- ⋈ D. leucas
- ✦ C. commersonii
- ▶ P. crassidens
- ◖ I. geoffrensis

DOCUMENT LIBRARY

March 11, 1991

Distribution List for Technical Report Exchange

Attn: Stella Sanchez-Wade
Documents Section
Scripps Institution of Oceanography
Library, Mail Code C-075C
La Jolla, CA 92093

Hancock Library of Biology &
  Oceanography
Alan Hancock Laboratory
University of Southern California
University Park
Los Angeles, CA 90089-0371

Gifts & Exchanges
Library
Bedford Institute of Oceanography
P.O. Box 1006
Dartmouth, NS, B2Y 4A2, CANADA

Office of the International
  Ice Patrol
c/o Coast Guard R & D Center
Avery Point
Groton, CT 06340

NOAA/EDIS Miami Library Center
4301 Rickenbacker Causeway
Miami, FL 33149

Library
Skidaway Institute of Oceanography
P.O. Box 13687
Savannah, GA 31416

Institute of Geophysics
University of Hawaii
Library Room 252
2525 Correa Road
Honolulu, HI 96822

Marine Resources Information Center
Building E38-320
MIT
Cambridge, MA 02139

Library
Lamont-Doherty Geological
  Observatory
Columbia University
Palisades, NY 10964

Library
Serials Department
Oregon State University
Corvallis, OR 97331

Pell Marine Science Library
University of Rhode Island
Narragansett Bay Campus
Narragansett, RI 02882

Working Collection
Texas A&M University
Dept. of Oceanography
College Station, TX 77843

Library
Virginia Institute of Marine Science
Gloucester Point, VA 23062

Fisheries-Oceanography Library
151 Oceanography Teaching Bldg.
University of Washington
Seattle, WA 98195

Library
R.S.M.A.S.
University of Miami
4600 Rickenbacker Causeway
Miami, FL 33149

Maury Oceanographic Library
Naval Oceanographic Office
Stennis Space Center
NSTL, MS 39522-5001

Marine Sciences Collection
Mayaguez Campus Library
University of Puerto Rico
Mayaguez, Puerto Rico 00708

Library
Institute of Oceanographic Sciences
Deacon Laboratory
Wormley, Godalming
Surrey GU8 5UB
UNITED KINGDOM

The Librarian
CSIRO Marine Laboratories
G.P.O. Box 1538
Hobart, Tasmania
AUSTRALIA 7001

Library
Proudman Oceanographic Laboratory
Bidston Observatory
Birkenhead
Merseyside L43 7 RA
UNITED KINGDOM

Mac 90-32

50272-101

REPORT DOCUMENTATION PAGE	1. REPORT NO. WHOI-92-04	2.	3. Recipient's Accession No.
4. Title and Subtitle Characterizing Acoustic Features of Marine Animal Sounds			**5. Report Date** January 1992
			6.
7. Author(s) Kurt M. Fristrup and William A. Watkins			**8. Performing Organization Rept. No.** WHOI-92-04
9. Performing Organization Name and Address Woods Hole Oceanographic Institution Woods Hole, Massachusetts 02543			**10. Project/Task/Work Unit No.**
			11. Contract(C) or Grant(G) No. (C) N00140-90-D-1979 (G) N00014-91-J-1445
12. Sponsoring Organization Name and Address NAVSEA, Office of Naval Research, NOARL and ORINCON/DARPA			**13. Type of Report & Period Covered** Technical Report
			14.

15. Supplementary Notes

This report should be cited as: Woods Hole Oceanog. Inst. Tech. Rept., WHOI-92-04.

16. Abstract (Limit: 200 words)

Software tools were designed to characterize the acoustic features of marine animal sounds. These have resulted in a set of calculated measurements that summarize particular aspects of sound sequences. The specificity of these measurements was enhanced by adjusting calculations to compensate for ambient noise. The sound measures included statistics for Aggregate Bandwidth, Intensity, Duration, Amplitude Modulation, Frequency Modulation, Short-term Bandwidth, Center Frequency, and Amplitude Frequency Interaction. The efficacy of noise compensation was tested for each statistic. Then, the sound measures were tested on a subset of 200 sequences of marine animal sounds, including sequences from 20 species: six baleen whales, 13 toothed species, and one seal. The statistics were reviewed for each species and a graphical comparison of all species was generated using principal components analysis. Preliminary results confirm that such sounds can be classified by means of relatively simple statistical algorithms, and we are encouraged to continue toward a system for automatic classification of marine animal sounds.

17. Document Analysis   a. Descriptors

sound database
marine animal sounds
automatic characterization of sounds

b. Identifiers/Open-Ended Terms

c. COSATI Field/Group

18. Availability Statement Approved for public release; distribution unlimited.	**19. Security Class (This Report)** UNCLASSIFIED	**21. No. of Pages** 82
	20. Security Class (This Page)	**22. Price**

(See ANSI-Z39.18)   *See Instructions on Reverse*   OPTIONAL FORM 272 (4-77)
(Formerly NTIS-35)
Department of Commerce