WHOI Statistical Study of Acoustics in Shallow Water Surf Zones

1/c Sara Senser

Project Advisors: Dr. James Preisig (WHOI)
CAPT Richard Hartnett (USCGA)
Dr. Keith Gross (USCGA)
Dr. Joseph Wolcin (USCGA)
LT Greg Czerwonka (USCGA)

Projects in Electrical & Computer Engineering
United States Coast Guard Academy
May 2003

WHOI Statistical Studies of Acoustics in Shallow Water Surf Zones

1/c Sara Senser

Project Advisors: Dr. James Preisig (WHOI)
CAPT Richard Hartnett (USCGA)
Dr. Keith Gross (USCGA)
Dr. Joseph Wolcin (USCGA)
LT Greg Czerwonka (USCGA)

Projects in Electrical & Computer Engineering
United States Coast Guard Academy
May 2003

ABSTRACT

The end goal of this project is to design an accurate, straightforward statistical model predicting when bubbles from breaking waves will absorb acoustic communication signals in surf zones. This model will be designed based on results from a mathematical analysis of data from the surf zone, and will pinpoint which environmental variables have the largest effect on the communication channel's behavior. Technical approaches to this problem include using mutual information calculations and variance reduction ratios to lower the uncertainty in received signal strength. Work to date has determined that the three most significant environmental variables in trying to predict signal strength are; 1) wave height, 2) the ratio of wave height to water depth and 3) the wind direction. Using these three variables to study data collected from the surf zone south of Scripps Pier in San Diego, the variance reduction ratio averages 75.5%. When trying to lower uncertainty in signal outages, current results show the most important one-dimensional environmental variables are significant and average wave height, which can lower uncertainty in outage probability by an average of 10.87% and in outage length by an average of 28.76%. This project affects groups working in surf zones, such as those engaged in Navy mine clearing operations, because these groups must be able to maintain reliable communications in order to complete their missions. As a result, the model created must be universally applicable and based on environmental conditions easily measured by a vessel on site in the surf zone.

INTRODUCTION

Data Overview

The data for this project was collected in the surf zone just south of Scripps Pier in San Diego. Results presented in this paper are based on data collected during 49 different 18-minute time periods. These time periods come from 5 day long tapes of data, and these 5 days split easily into two sets. The first set of days is data for 3 days in early July, the second set includes the remaining two tapes, with data collected about 6 weeks later in mid September. In the 6 weeks between these two sets, the sand bar off of Scripps Pier moved and the slope of the surf zone changed significantly. Figure 1 below shows the slope difference by plotting the average depth at pressure sensors 2-9 (sensors 1 and 10 did not always get reliable data so they are not included in the figure) during data sets 1 and 2. During Set 1, there is a direct relationship between water depth and sensor number such that water generally gets deeper as sensor number increases, and sensor 10 is an average of 2.27 meters deeper in the water than sensor 2. During Set 2, the depth is fairly constant across all the sensors, with only 2 cm of difference between sensors 9 and 2. Sensor 9 is actually in water 2 cm shallower than the average depth for sensor 2 located 80 meters closer to shore. All analysis was done using depth data from pressure sensor number 6 because this sensor is at about a midpoint for both slopes and the average depth at this point is about the same for both sets.

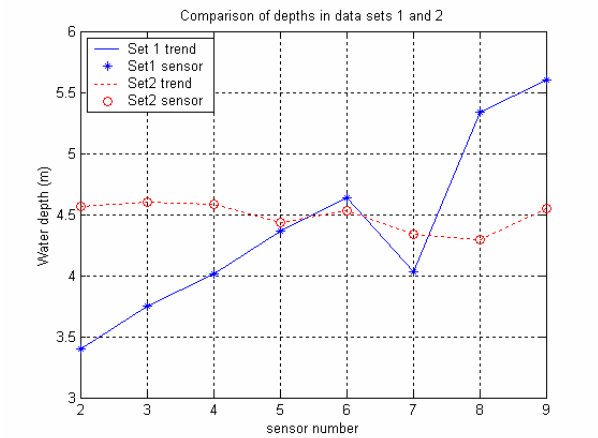

Figure 1: A comparison of sensor depths between Set 1 and Set 2, showing the change in bottom slope.

There are several other drastic environmental differences between the two sets. Compared to Set 2, Set 1 had bigger waves, a longer wave period and lighter winds. Wave heights during the first set are an average of 22% bigger and the average wave period was 22% longer than those seen in the second set. Wind speeds in Set 2 are 20% greater than those during set 1. Both sets have variable wind directions, winds from the North are most common. More detailed summary information is provided in Attachment A.

Experiment Set-up

The equipment set-up, located in the Surf Zone just south of Scripps Pier in San Diego is shown in Figure 2 below. The vertical line on the far right side of the figure is Scripps Pier, the circles on this line are the stanchions. The signal we are studying is transmitted from the colored circle at the top of the figure (located even with the end of Scripps pier). The signal is received 100 meters away at the other colored ball, located about halfway up the figure below the 2 horizontal lines. The line connecting the transmitter and receiver is an array of pressure sensors. The line itself is just a cable, the boxes on this line show the 10 pressure sensors with about 10 meters between them. The horizontal lines are long shore pressure arrays, the data they collected are not studied during this project. Wind data used in this project came from a separate source in the surf zone.

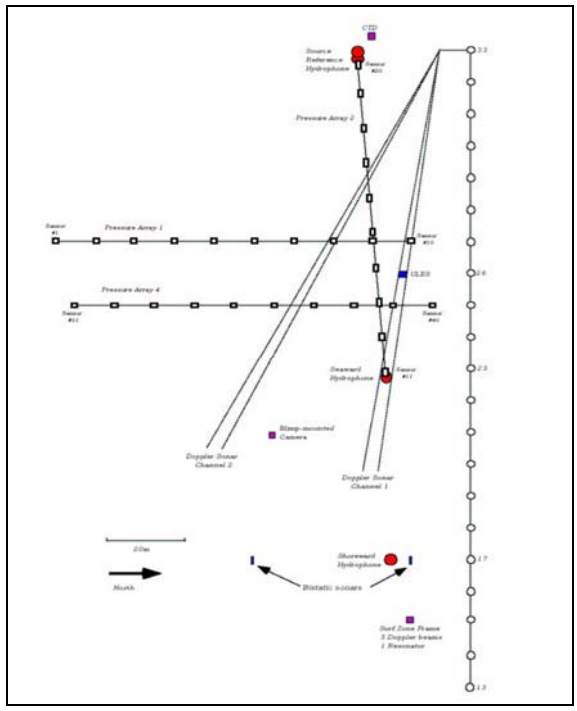

Figure 2: The equipment set up South of Scripps pier to record pressure and energy data.

Oceanography Background

Water is usually a fairly benign medium, where communication signals with frequencies less than 33 kHz can be transmitted with a relatively low frequency of error. However, in a surf zone, traveling waves can act as transient caustics and the bubbles introduced into the medium by breaking waves can absorb a communications signal completely. Transient caustics occur when a wave takes the right shape to reflect a signal without dissipating it. When a transient caustic is present, very little energy arrives at the expected initial arrival, instead a strong signal arrives at a delayed time. In Figure 3 below, a caustic occurs at about 2 seconds on the x axis, with a delay of 1.5 msec on the left hand y-axis. (It shows up as a dark circle). This problem is being studied at Woods Hole Oceanographic Institute.

The other problem, signal outages, is the focus of this project. Figure 3 below shows a signal outage when the signal strength shown with the curve on the figure falls to below –50 dB (right hand y-axis) between seconds 3.5 and 4 (x-axis). This is fairly undisruptive outage in comparison to others seen in the data, because it is short and the signal strength stays above –52dB. Many outages last several minutes during which time the signal strength often falls below –60 dB.

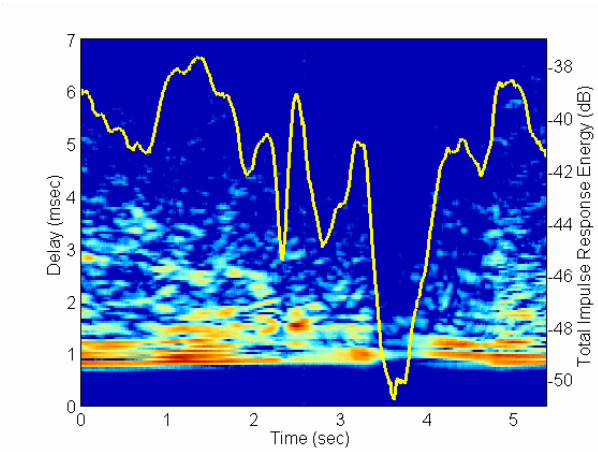

Figure 3: A figure displaying the problems sometimes seen in surf zone environments.

Oceanography Definitions

Wave Definitions

Tide is an average water depth. In this project, tide at a given time was calculated by averaging all the depth samples from 9 minutes before the time to 9 minutes after the time.

Figure 4 below shows a 2 minute wave sample. A wave height is defined as the maximum value between intersections with the mean water level. Wave heights on the figure are depicted with crosses. The average or RMS wave height is the average of all of these wave heights. The significant wave height is the average of the biggest third of the waves. These waves have circles at the top of them in the figure below. The maximum wave height is the biggest wave seen in the designated period of time. In the example shown below, the maximum wave is the third one, occurring about 30 seconds into the sample.

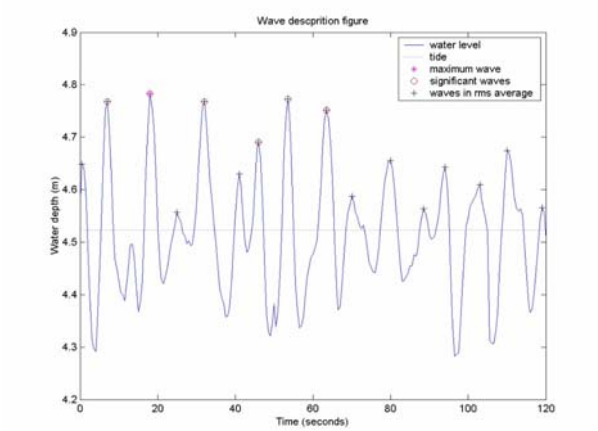

Figure 4: A 2 minute sample of wave data used to explain wave definitions

Wind Definitions

The most important wind terminology talks about the wind direction. An onshore wind blows from sea to land, and onshore wind blows from land to sea. This is important because these winds will change the behavior of breaking waves. Onshore winds cause waves to break in deeper water, and are more likely to put a lot of bubbles in the water column. Offshore winds cause waves to break in shallower water with fewer bubbles (Reference 1).

For the purposes of this paper, variables called wind speed relative to North are the component of the wind coming from the North, based on the provided wind speed and direction.

Outage Definition

An outage is any time the signal strength (dB) falls below a defined threshold. Outage thresholds have been defined starting at –35 dB and continuing every 5 dB down until –60dB. At some point, the signal strength dissipates so far that the receiver does not get enough energy to recognize the signal. This critical outage level varies based on the environment's signal to noise ratio, but for the purpose of this paper, the critical outage level where a receiver can not get enough strength from the signal to complete the mission is defined as –50dB.

Previous research

This project is an active area of research at Woods Hole Oceanographic Institute and at Naval Research Labs. In the past, research has used communication theory techniques like entropy and mutual information calculations to compare individual environmental variables to the amount of energy received. The most important relationship discovered was a threshold in the ratio of significant wave height to water depth. Below this threshold breaking waves will not disrupt a communications signal. Above this threshold, the signal is sometimes lost. Figure 5 below shows a plot displaying the ratio of wave height to water depth versus total energy received (dB). Different shapes depict different time periods.

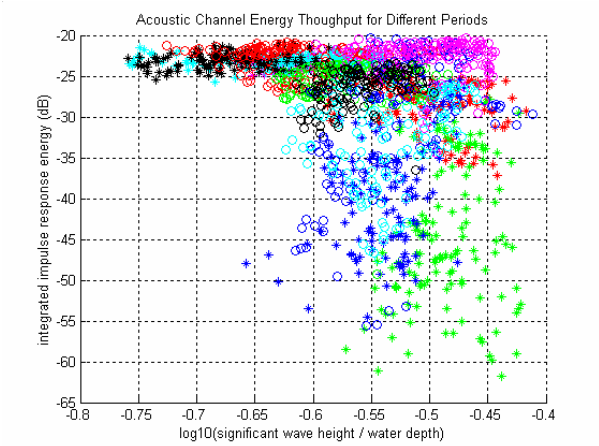

Figure 5: The relationship between ratio of significant wave height to water depth and the energy received. A threshold is seen in the ratio, below this threshold the signal will not be lost.

In addition to the threshold discussed above, the data depicted in Figure 5 was encouraging because all the data points in each time period show consistent behavior. This means that while the total variance in energy values above the threshold increases relative to the energy variance below the threshold, the variance for each time period tends to stay small. For example, there is a time period displayed with purple circles, during which the signal strength remains good despite a high ratio. The time period depicted with green stars has a similar ratio, but during this time very little energy was received. The relationships shown in Figure 5 prompted more research because it led scientists to believe there was something in the environment, hopefully something measurable, which changed between time periods like the two discussed that led to the differences in the signal behavior. This research was aimed at answering two questions. The first goal was to study the threshold itself, and find factors that would affect it's location to help predict the threshold accurately. Oceanographers have an equation known as McGowen's breaking criteria and while it considers many factors, the only factor which can change on the time periods we are studying is the wave height and water depth. Therefore, there might be an equation that can help pinpoint what the threshold should be, aiding future predictions for different environments. The second and more important goal was to determine which environmental factors gave two time periods with similar ratio values different energy values. More generally, this goal aimed at finding out which environmental factors had the greatest effect on received signal strength.

OBJECTIVE

The long-term objective of this project is to create a general model predicting when communication signals will be lost in surf zones. This knowledge is especially important for groups that work in surf zones, where the breaking waves create bubbles that can destroy communication signals for minutes at a time. Scientifically, the goal is to determine which environmental factors affect the signal's ability to travel through the surf zone, then to use these factors to predict how often the signal will dissipate and how long it will stay out. The model has to be universally applicable to surf zones with different bottom types, weather and wave behavior.

Operationally, this study will teach people what environmental readings they need to take in order to predict the stability of their communications channel. When completed, the model will be simple enough to use when on site in the surf zone, only requiring people to pay special attention to readings a prudent mariner would already record, such as water depth, wave height, wind direction and wind speed. One specific application of this project is Navy mine clearing operations where a ship must be able to communicate with equipment working in the shallow water surf zones. The model will give them the information they need to make solid decisions about their mine clearing duties. Knowing facts about their channel stability will help people determine the best communication method based on environmental conditions. For example, on a day when conditions are likely to create frequent signal outages in the shallow water, mine clearing robots can be programmed to enter the shallow water, complete a task then return to deeper water where communication channels are more stable.

SYSTEM DESIGN

Field creation

The first step in finding relationships between environmental variables and the energy received was to modify raw data using calculations that would change the data into variables we wanted to study. These variables were then entered into data fields where they could easily be stored and studied.

Energy Data

The energy data was originally in a form where it presented energy received at each time and at delay values up to 7 msec. This data was summed along the delay axis so that the matrix became of vector showing total energy received. Looking back to Figure 3, this calculation is the equivalent of changing the data

depicted by the colors in the background into the data shown by the curve going across the figure. Energy was stored in the fields as absolute energy with sample rate of 100Hz. For most calculations, absolute energy was later changed to energy in dB using the following equation:

$$\boxed{Pwr(dB)=10*\log_{10}(Pwr(Absolute))} \quad (1)$$

Wave Data

The pressure data provided gave water depth in meters at each of the 10 sensors (at a rate of 5Hz for the first three days of data and 2Hz for the last two days). Each depth sample had a time associated with it, stating when the depth sample was taken. These times were changed from ticks to a Julian day, hour, minute and second. This calculation required a set of two coefficients provided with the raw data.

As stated in the definitions, raw depth data was changed into a tide by averaging all the samples from 9 minutes before a set sample to 9 minutes after. Other depth variables were depth averaged over 2 minutes, 1 minute and 10 seconds. These variables were calculated by taking the depth samples from 2 minutes, 1 minute or 10 seconds respectively before the sample in question and averaging them together. The final depth variable was instant depth, the depth at the sleected pressure sample.

The data for the shallowest sensor (sensor 1 during set 1 and sensor 10 during set 2) sometimes had some low frequency noise included with the initial depth values, so the depth averages discussed so far do not always provide reliable values when looking at sensors 1 and 10. This does not effect the results as analysis was not done on these sensors.

Before any other wave calculations were performed, the pressure data for all sensors was filtered to remove any noise included with the data. The filter used was a band pass, Kaiser window, FIR filter with a pass band between 0.05 Hz and .4Hz and stop band attenuation of 30 dB. This led to a 384th order filter for the 5Hz data and a 154th order filter for the 2Hz data.

The filter removed the DC component of the wave data, so once the data is filtered, the remaining data shows depth fluctuations about the mean water level (tide). An algorithm was found to find wave heights by locating maximum values in the data between intersections with the mean water level.

Once wave heights were found, the average wave height was calculated over intervals of 2 minutes, 1 minute and 10 seconds preceding each depth sample by averaging all the wave heights found in the given time period. Significant wave height was found by averaging the largest third of the waves in each time period and the maximum wave height in each time period was also recorded. Another variable recorded was the number of waves seen in 10 seconds, 1 minute and 2 minutes prior to the current pressure sample. Some data was recorded about the most recent waves, including the height of the last wave, the time since the last wave and the time between the previous two waves. The final wave variable in the fields is instant wave height, the wave height (relative to mean water depth) at the pressure sample in question.

Wind Data

Wind data did not require many calculations. In raw form, it gave 5-minute averages for average and maximum wind speeds, wind direction, wave period and wave height. These values were transferred directly into data fields.

Summary

In summary, when fields are complete, they include the following data (NOTE: all pressure related variables area vector of 10 entries, 1 for each sensor):

- time: The time associated with the pressure data sample, given with Julian day, hour, minute and second
- instdepth: The depth at the pressure data sample.
- depth10sec: Depth averaged for 10 seconds prior to the pressure sample
- depth1min: Depth averaged for 1 minute prior to the pressure sample
- depth2min: Depth averaged for 2 minutes prior to the pressure sample
- tide: Depth averaged from 9 minutes before to 9 minutes after the current pressure sample.
- instwave: The wave height (relative to mean water level) at the current pressure sample
- rmspeak10sec: RMS wave height averaged for 10 seconds prior to the current sample
- rmspeak1min: RMS wave height, averaged for 1 minute prior to the current sample
- rmspeak2min: RMS wave height, averaged for 2 minutes prior to the current sample
- sigwave10sec: Significant wave height, averaged for 10 seconds prior to the current sample
- sigwave1min: Significant wave height, averaged for 1 minute prior to the current sample

- sigwave2min: Significant wave height, averaged for 2 minutes prior to the current pressure sample
- maxwave10sec: Maximum wave seen in 10 seconds prior to the current pressure sample
- maxwave1min: Maximum wave height seen in the minute prior to the current pressure sample
- maxwave2min: Maximum wave height seen in 2 minutes prior to the current pressure sample
- numwave10sec: Number of waves seen in 10 seconds prior to the current pressure sample
- numwave1min: Number of waves seen minute prior to the current pressure sample
- numwave2min: Number of waves seen in 2 minutes prior to the current pressure sample
- timelastpeak: The number of seconds since the last wave
- timebetweenpeaks: The time (seconds) between the previous two waves
- heightlastpeak: The height of the most recent wave
- energy: The energy values received between the time of this pressure sample and the next one. Since energy was recorded at 100 Hz and pressure was recorded at 5Hz or 2Hz, these energy vectors usually include 20 or 50 energy samples respectively.
- windtime: The time associated with the wind data, included as a check that wind was added to the right field
- windavespd: The average wind speed (5 minute average)
- windtopspd: The maximum wind speed (5 minute average)
- winddir: The wind direction (5 minute average)
- windtide: The tide values provided with wind data, included as a reality check for pressure data tides
- wvhgt: The wave height provided with wind data, included as a reality check for pressure data wave heights
- wvprd: The wave period provided with wind data

Outage Fields

The second phase of this project used fields with outage information and some basic environmental variables to begin finding relationships between outages and the environment. One outage field was created to match up with each of the 18-minute long environmental data fields discussed above. Variables included in these outage fields include:

- timestart: The time of the first entry entered in the field
- timeend: The time of the last entry in the field
- energy: All the energy values between the beginning and end times
- out35: The number of times the energy level fell below the –35 dB threshold
- thresh35: a vector of energy values, which when plotted on top of the initial energy matrix will show where the signal falls below the –35 dB threshold
- out35mean: The mean signal level while the energy stays below the –35dB threshold
- out35std: The standard deviation of the energy values while the signal level stays below the –35dB level
- out35length: The length of the outages on the –35dB threshold level
- out35lenmean: The average length of outages on the –35dB threshold
- out35lenstd: The standard deviation of the length of outages on the –35dB threshold
- **Note:** The variables, out35, thresh35, out35mean, out35std, out35length, out35lenmean and out35lenstd are repeated for thresholds at –40dB, -45dB, -50dB, -55dB and –60dB
- Rmswavemean: The average of all rmspeak2min entries at sensor 6 between timestart and timeend
- Rmswavestd: The standard deviation of all rmspeak2min entries at sensor 6 between timestart and timeend
- Sigwavemean: The average of all sigwave2min entries at sensor 6 between timestart and timeend
- Sigwavestd: The standard deviation of all sigwave2min entries at sensor 6 between timestart and timeend
- Maxwavemean: The average of all maxwave2min entries at sensor 6 between timestart and timeend
- Maxwavestd: The standard deviation of all maxwave2min entries at sensor 6 between timestart and timeend
- Wtsmean: The average of all windtopspd entries between timestart and timeend
- Wtsstd: The standard deviation of all windtopspd entries between timestart and timeend
- Wdirmean: The average of all winddir entries between timestart and timeend
- Wdirstd: The standard deviation of all winddir entries between timestart and timeend
- Wvprdmean: The average of all wvprd entries between timestart and timeend
- Wprdstd: The standard deviation of all wvprd entries between timestart and timeend
- Maxwav: The biggest wave in the time period between timestart and timeend
- Sigwav: The significant wave height calculated for the interval from timestart to timeend
- Rmswav: The average wave height calculated for the interval from timestart to timeend

- Windavespd: The average wind speed calculated for the interval timestart to timeend

Technical Approach

There are three approaches that have been used to find relationships between variables in this project. In fairly early stages of analysis, figures were used to find relationships between signal strength and the environmental variables visually.

The first mathematical approach used was communication theory techniques like entropy and mutual information values. These helped determine how well environmental variables could decrease the uncertainty in received energy values. This approach was helpful in determining which environmental variables were the most important, but calculating entropy required putting energy values into bins and the bin size chosen had a significant effect on the results.

The second mathematical approach was to use variance calculations and variance reduction ratios to lower the variance in the values for energy received. The results using these calculations were fairly consistent with those from mutual information calculations. In addition, variance reduction calculations provided results that would be more helpful in designing a model because the energies did not have to be placed into bins.

Visual Inspection of data

Since humans are visually oriented, to determine a starting place for our mathematical analysis of the data, we used MATLAB ® figures to plot the data and look for environmental variables that could group outages together into tight groups. If a figure could show a clear relationship between a variable and the amount of energy received, we could expect the mathematical analysis to show a similar relationship. Figure 5 above provides an example of the type of figures created in this phase of analysis. Other figures grouped data by signal level rather than by time period. As analysis progressed, figures increased in dimensions, sometimes graphing energy values versus 3 independent environmental variables.

Mutual Information

The approach taken during the first part of this project was based on communications concepts, specifically on information theory. Data complexity is measured in bits of entropy, calculated from probabilities determined by putting data into bins to create a histogram. Once the energy values are in bins, entropy is calculated with the equation:

$$H(x)= -\Sigma p(x)*\log_2 p(x) \qquad (2)$$

where H(x) is the entropy of the data, and p(x) is the probability of falling into a given bin on the histogram.

Normalized mutual information values are calculated to determine how closely the energy received is related to different environmental factors. Histograms are used to calculate mutual information, and this value is normalized to limit the effect of the histogram's bin size on the final percentage of mutual information. The equation used to calculate normalized information is:

$$\frac{H(x)-H(x|y)}{H(x)} \qquad (3)$$

where H(x) is the entropy of the energy data and H(x|y) is the conditional entropy of energy given some environmental factor y. The numerator of equation (3) is mutual information between energy and environmental variable y. This conditional entropy was found by sorting energy into sets based on the value of the environmental factor and sending the energy values in each set to the histogram.

Variance Reduction Ratios

The second mathematical way of looking at the data was variance values and variance reduction ratios. This calculation gave a measure of how well knowledge of an environmental variable could decrease the overall uncertainty in the received energy values.

Variance is a measure of how far individual data points fall from the mean of the data set. Once the variance is known, the process for calculating conditional variance includes sorting energy values into sets based on the value of one or more environmental variables. The variance of energies in these sets is calculated and weighted by the probability of falling into the bin. The equation looks like:

$$\sigma_{e|y}^2=\Sigma p_i*\sigma_{e|y=yi}^2 \qquad (4)$$

where $\sigma_{e|y}^2$ is the conditional variance of energy given one or more environmental factors, p_i is the probability of the energy falling into a given set "i", and $\sigma_{e|y=yi}^2$ is the variance of the energy values in set "i".

Once the conditional variance is known, the variance reduction ratio is determined with the following equation:

$$\text{Variance Reduction Ratio} = \frac{\sigma_{e|y}^2 - \sigma_{e|y}^2}{\sigma_{e|y}^2} \quad (6)$$

This ratio provides a measure of the reduction in variance as a percent of the original variance. Functions have been written to calculate the variance reduction ratio based on up to three environmental variables.

RESULTS PART 1: Relationships Between Signal Strength and Environment

Environmental Variables Discarded

Variables with a 1 dimensional normalized mutual information value less than 6% were discarded at an early stage in the analysis. Table 1 below shows which variables were discarded and their normalized mutual information values averaged across all the 49 fields.

Table 1 shows stuff variables discarded after calculating 1 dimensional normalized mutual information values. Results are listed in order of decreasing normalized mutual information.

Variable Name	Normalized Mutual Information
Ratio of maxwave10sec to tide	5.57%
Maxwave10sec	5.53%
Ratio of sigwave10sec to tide	5.46%
Sigwave10sec	5.39%
Ratio of rmspeak10sec to tide	4.34%
Rmspeak10sec	4.33%
Numwave2min	4.18%
Heightlastpeak	3.00%
Numwave1min	2.46%
Timebetweenpeaks	2.00%
Timelastpeak	1.17%
Numwave10sec	0.83%

More detailed information about 1 dimensional normalized mutual information values is included in Attachment B.

Significant Environmental Variables

Mutual Information Results

Initially, variables with a 1 dimensional normalized mutual information value of more than 8.5% were identified as significant environmental variables. Table 2 below shows these variables and their normalized mutual information values, averaged over all 49 fields.

Table 2 shows the 1 dimensional variables initially as significant as a result of high normalized mutual information results. These variables are listed in order of decreasing normalized mutual information variables.

Variable Name	Normalized Mutual Information
Tide	18.88%
Ratio of maxwave2min to tide	13.94%
Maxwave2min	13.68%
Ratio of maxwave1min to tide	12.08%
Maxwave1min	11.89%
Depth2min	11.64%
Windavespd relative to North	9.48%
Ratio of sigwave2min to tide	9.32%
Sigwave2min	9.32%
Depth1min	9.01%
Windtopspd relative to North	8.98%
wvprd	8.85%
Windavespd	8.78%
Ratio of rmspeak2min to tide	8.72%
Rmspeak2min	8.60%

As time went on, several other values were reintroduced into the analysis despite their low 1 dimensional normalized mutual information values. Table 3 shows these additional variables with their normalized mutual information values, averaged over all 49 data fields.

Table 3 shows the average normalized mutual information values for variables that were reintroduced into analysis despite relatively low normalized mutual information values. Results are listed in order of decreasing normalized mutual information.

Variable Name	Normalized Mutual Information
Ratio of sigwave1min to tide	7.99%
Sigwave1min	7.99%
Winddir	7.62%
Rmspeak1min	7.11%
Ratio of rmspeak1min to tide	7.06%
Windtopspd	6.40%

Trends in the above data that will be discussed further are the 1. the tendency for the ratio of wave height to water depth to have better results than the wave height alone 2. the tendency for wave variables averaged over 2 minutes to have more mutual information than 1 minute averages and 10 second averages 3. the tendency for relative wind speeds to have more normalized mutual information that wind speed by itself.

Table 4 below shows mutual information results summarized by Tape and also gives information about the number of outages seen at the –50dB level for each tape, showing a generally direct relationship between the number of outages and normalized mutual information.

Table 4 shows the number of outages at the –50 dB level and the average normalized mutual information values for each tape.

Tape Number	Number of outages at –50dB level	Average Normalized Mutual Information for Tape
226	7	14.57%
227	481	13.66%
228	128	9.12%
240	3	4.53%
241	0	4.47%

Variance Reduction Results

1 Dimensional

Table 5 below shows the 1 dimensional variance reduction ratios for significant environmental variables. These results are calculated taking all 49 fields into account.

Table 5 shows 1 dimensional variance reduction ratios for significant environmental variables

Variable Name	Variance Reduction Ratio (%)
Wind direction	31.55%
Ratio of maxwave2min to tide	15.93%
Average wind speed relative to north	15.18%
Top Wind speed relative to North	13.53%
Ratio of maxwave1min to tide	12.32%
Sigwave2min	11.94%
Ratio of rmswave2min to tide	11.89%
Rmswave2min	11.88%
Maxwave2min	11.60%
Ratio of sigwave2min to tide	10.71%
Ratio of sigwave1min to tide	9.53%
Sigwave1min	9.22%
Rmswave1min	9.21%
Ratio of rmswave1min to tide	9.09%
Maxwave1min	8.61%
Top Wind speed	7.89%
Average wind speed	6.70%

VARIANCE REDUCTION RATIO RESULTS AND DIFFERENT TYPES OF POWER

Variance reduction calculations were done for absolute energy and for energy in dB. The results were dependant on the type of power used. The average difference between results for the two different kinds of power was 6.78%. With a wave related environmental variable, the ratio was higher for the dB calculation by an average of 4.11% while for wind related variables, the absolute power calculation provided a better ratio by an average of 13.20%. Differences are listed by variable in Table 6 below. The results displayed were calculated taking all 49 fields into consideration.

Table 6 shows the variance reduction ratios for absolute power and power in dB for most significant variables. The results are listed in order of decreasing difference.

Variable Name	Var. Red. Rat. (%) for abs pwr	Var. Red. Rat. (%) for pwr in dB	Diff (%), (dB-abs pwr)
Rat of maxwave2min to tide	9.50	15.93	6.43
Rat of maxwave1min to tide	7.18	12.32	5.14
RMSwave2min	7.22	11.88	4.66
Maxwave2min	7.39	11.6	4.21
Sigwave2min	7.13	11.94	4.81
Sigwave1min	5.38	9.22	3.84
Rat of RMSwave2min to tide	8.41	11.89	3.48
Ratio of sigwave1min to tide	6.09	9.53	3.44
Maxwave1min	5,19	8.61	3.42
Ratio of sigwave2min to tide	7.62	10.71	3.09
RMSwave1min	5.33	9.21	3.08
Rat of RMSwave1min to tide	6.15	9.09	2.94
Wind direction	42.29	31.55	-10.7
Average wind speed	17.84	6.70	-11.1
Top Wind Speed rel. to North	27.17	13.53	-13.6
Top Wind Speed	22.07	7.89	-14.2
Ave. wind speed rel. to North	31.46	15.18	-16.3

To simplify the results presented in this section, all the variance reduction ratio results shown in tables and discussed are based on calculations with power in dB unless otherwise specified. More 1 dimensional variance reduction ratio results are presented in Attachment C.

VARIANCE REDUCTION RATIO RESULTS AND DIFFERENT LENGTHS OF TIME

Variance reduction ratios are calculated by field, tape, set of tapes and overall. There was generally indirect

relationship between the amount of time included in the calculation and the variance reduction results. Table 7 and 8 below summarize the effect of time on wave and wind related variance reduction results.

Table 7 shows a summarizes the effect of time on wave variable variance reduction ratio results, using the ratio of maximum wave (2 minutes) to tide as an example

Time Length	Minimum Variance Reduction Ratio (%)	Maximum Variance Reduction Ratio (%)	Average Variance Reduction Ratio (%)
Field (18min)	0.42%	64.81%	24.76%
Tape (day)	3.55%	25.9%	14.45%
Set	8.98%	18.28%	13.63%
Total	N/A	N/A	15.93%

Table 8 shows the trend of time on wind variable variance reduction ratios using wind direction as an example

Time Length	Minimum Variance Reduction Ratio (%)	Maximum Variance Reduction Ratio (%)	Average Variance Reduction Ratio (%)
Tape (day)	21.86%	66.16%	48.88%
Set	46.57%	52.42%	49.50%
Total	N/A	N/A	31.55%

VARIANCE REDUCTION RATIO RESULTS AND NUMBER OF OUTAGES

Table 9 below shows the average variance reduction ratio for each Tape compared to the number of outages on the –50 dB threshold for that tape. There is a direct relationship between the number of outages and the average normalized mutual information values for that tape.

Table 9 shows the number of outages at the –50 dB threshold and the average normalized mutual information value for each tape

Tape Number	Number of outages at –50dB level	Average Normalized Mutual Information for Tape
226	7	28.22%
227	481	16.50%
228	128	17.33%
240	3	7.10%
241	0	8.80%

2 Dimensional

When the both environmental variables used in a two dimensions calculation are wave variables, the average total variance reduction ratio is 20.15%. If one of the variables is a wind variable, this average variance reduction more than doubles, to 46.61%. In general, if one variable is a wind variable, there is not a significant difference between the variance reductions seen for cases when the second dimension is wave height and those when the second is a wave height ratio. The performance of wind variables in total 2 dimensional variance reduction ratio calculations is shown in Table 10 below. The table shows that wind direction is the variable with the best variance reduction, followed by relative average wind speed and then relative top wind speed.

Table 10 shows the average total 2 dimensional variance reduction ratios for calculations involving wind variables

Wind Variable	Average 2D Variance Reduction Ratio
Wind Direction	58.47%
Average Wind Speed Relative to North	42.78%
Top Wind Speed Relative to North	38.57%

The performance of wave height variables in 2 dimensional variance reduction ratio calculations is shown in Table 11 below. It shows that maximum wave height has the best variance reduction and that all the 2 minute wave averages have better variance reduction than the 1 minute wave averages.

Table 11 shows the average total 2 dimensional variance reduction ratios for calculations involving wave height variables

Wave Height Variable	Average 2D Variance Reduction Ratio
Maxwave2min	38.63%
Sigwave2min	34.42%
Rmswave2min	33.81%
Maxwave1min	32.80%
Sigwave1min	30.25%
Rmswave1min	29.47%

The performance of ratio of wave height to tide variables in 2 dimensional variance reduction ratio calculations is shown in Table 12 below. The table shows that the ratio of maximum wave height 2 minute to water depth is most important, followed by the ratio of maximum wave height 1 minute to water depth.

Table 12 shows the average total 2 dimensional variance reduction ratios for calculations involving ratios of wave height to water depth.

Ratio of wave to tide using wave:	Average 2D Variance Reduction Ratio
Maxwave2min	39.52%
Maxwave1min	35.05%
Rmswave2min	34.74%
Sigwave2min	32.58%
Sigwave1min	30.22%
Rmswave1min	29.03%

The best results for 2 dimensional variance reduction ratio was when the two dimensions were wind direction and the ratio of maximum wave (2 minutes) to water depth. More 2 dimensional variance reduction ratios are shown in Attachment D.

VARIANCE REDUCTION RATIO RESULTS AND DIFFERENT TYPES OF POWER

Comparing dB to actual, when only wave information is involved (i.e. the two dimensions are a wave height and a ratio of wave height to water depth), the dB value is always greater than the absolute power ratio, by an average of 6.63%. When one of the dimensions is a wind variable, the variance reduction ratio is usually greater for absolute power than for power in dB. On 3 occasions, the variance reduction is greater for power in dB. All three of these occasions, the two dimensions were wind direction and a wave height and there was never more than a percent difference between the variance reduction ratio for absolute power and power in dB. The average difference for variance reduction ratios that include a wind variable is absolute power 5.08% greater than power in dB.

VARIANCE REDUCTION RATIO RESULTS AND DIFFERENT LENGTHS OF TIME

The comparison of 2 dimensional results to the time included are given below in Table 13. It shows that sets usually have the best average ratio, followed by tapes and then by the overall value.

Table 13 shows the effect of time included in calculations on the variance reduction ratio results, using ratio of significant wave height (2 min) to tide and average wind speed relative to North as an example

Time Length	Minimum Variance Reduction Ratio (%)	Maximum Variance Reduction Ratio (%)	Average Variance Reduction Ratio (%)
Tape	22.13	63.38	44.038
Set	54.58	60.67	57.62
Total	N/A	N/A	44.91

VARIANCE REDUCTION RATIO RESULTS AND NUMBER OF OUTAGES

The average 2 dimensional variance reduction ratio varies directly with the number of outages on the –50 dB threshold. This data is presented in more detail in Table 14 below.

Table 14 shows the average 2 dimensional variance reduction ratio for each tape and the number of outages for each tape at the –50 dB threshold level

Tape Number	Number of outages at –50dB level	Average Normalized Mutual Information for Tape
226	7	53.42%
227	481	45.84%
228	128	41.60%
240	3	22.85%
241	0	24.12%

3 Dimensional

The smallest 3 dimensional variance reduction ratio occurred when the 3 dimensions were average wave height (1 minute), the ratio of significant wave height (1 minute) to water depth and top wind speed relative to North. This value is 50.57%. The biggest variance reduction ratio is 82.77% and occurs when the three variables are average wave height (2 minutes), ratio of maximum wave height (2 minutes) to water depth and wind direction. The average 3 dimensional variance reduction ratio is 65%, if the wind variable used is wind direction, this average can be increased to 75.5%. The performance of wind variables in total 3 dimensional variance reduction ratio calculations is shown in Table 15 below. The relative importance of wind variables in variance reduction calculations is the same as seen in Table 10 above.

Table 15 shows the average total 3 dimensional variance reduction ratio for all calculations with each wind variable

Wind Variable	Average 3D Variance Reduction Ratio
Wind Direction	75.55%
Average Wind Speed Relative to North	61.72%
Top Wind Speed Relative to North	59.57%

The performance of wave height variables in 3 dimensional variance reduction ratio calculations is shown in Table 16 below. The relative importance of

wave variables in 3 dimensions is the same as seen for 2 dimensional calculations in Table 11 above.

Table 16 shows the total 3 dimensional variance reduction ratio for all calculations with each wave variable

Wave Height Variable	Average 3D Variance Reduction Ratio
Maxwave2min	69.94%
Sigwave2min	66.97%
Rmswave2min	66.16%
Maxwave1min	65.06%
Sigwave1min	62.90%
Rmswave1min	62.62%

The performance of ratio of wave height to tide variables in 3 dimensional variance reduction ratio calculations is shown in Table 17 below. The relative importance of ratio variables in 3 dimensional calculations is the same as in 2 dimensional calculations (Table 12 above).

Table 17 shows the average total variance reduction ratio for all calculations involving each ratio of wave height to water depth

Ratio of wave to tide using wave:	Average 3D Variance Reduction Ratio
Maxwave2min	72.20%
Maxwave1min	68.70%
Rmswave2min	65.54%
Sigwave2min	63.39%
Sigwave1min	62.52%
Rmswave1min	61.30%

More 3 dimensional variance reduction ratio results are presented in Attachment E.

VARIANCE REDUCTION RATIO RESULTS AND DIFFERENT TYPES OF POWER

Comparing results based on power in dB to those based on absolute power, the absolute power result is always higher, by an average of 1.32%.

VARIANCE REDUCTION RATIO RESULTS AND DIFFERENT LENGTHS OF TIME

In 3 dimensional results, the variance reduction ratios calculated over the two sets are 5.56% higher on average than the variance reduction ratios calculated over all 49 fields.

RESULTS PART 2: Relationships Between Outages and Environment

Result Skepticism

The most recent results in this project begin to find relationships between outages and environmental variables. These calculations were done with 49 data points, one for each 18 minute outage field. This is not sufficient data to complete the variance reduction ratio calculations with any reliability. Outliers give the total data set a very large variance, and when these outliers are placed into sets the sets with outliers have very large variances. The problem is magnified if the outliers fall into sets with a high probability because the set variance is weighted by the set probability as shown in Equation (6) above. Sometimes the outliers cause such large set variances that the variance reduction ratio is negative.

As a partial fix to this problem, outliers were defined as any points more than 2 standard deviations from the data set's mean. The value of outliers was then modified so that all outliers had a value equal to standard deviations from the mean. To an extent, this still accounts for the fact that the data point is very large in relation to the rest of the data without letting it through off data variances too much.

Variance Reductions for Outage Variables

The only environmental variables used in this phase of the analysis were average wave height, significant wave height average wind speed, average wind direction and maximum wave height, all calculated over an 18 minute period. The relative importance of these variables averaged over all thresholds is shown in Table 18 below. The table shows that rms wave height and significant wave height are generally the most important variables, followed by maximum wave height then wind variables.

Table 18 shows the variance reduction for environmental variables, averaged over all threshold levels.

Environmental Variable Name	Average Variance Reduction Ratio
Rmswav	14.95%
Sigwav	12.80%
Maxwav	10.04%
Wdirmean	3.21%
Windavespd	1.22%

The results calculated so far are summarized in Table 19 below. More detailed results are provided in Attachment F. It shows that rms wave height results are always higher than results for significant wave height and that the outage variables with the best variance reduction ratios are the mean outage length and total outage length. The variable with the lowest variance reduction ratio is the individual outage lengths.

Table 19 shows the variance reduction ratios for several outage variables given the environmental variables of significant wave height and average wave height

Outage Variable	Variance Average Reduction Ratio with Sig. Wave Height	Variance Average Reduction Ratio with RMS Wave Height
Mean outage length	27.41%	30.1%
Total outage time	18.29%	22.12%
Mean signal level during outage	9.94%	9.74%
Number of outages	9.52%	12.21%
Individual outage lengths	-0.92%	0.64%

The results show a direct relationship between the severity of the outage and the average variance reduction ratio. This trend is shown in Table 20 below. The table also shows that wave information has a higher average variance reduction ratio than wind information at all threshold levels.

Table 20 shows the average variance reduction ratios at different threshold levels for wave variables, wind variables and an average of all environmental variables.

Outage Level (dB)	Average of Variance Reduction for Wave Variables	Average of Variance Reduction for Wind Variables	Average Variance Reduction for Wind Variables
-35	9.78%	4.57%	7.70%
-40	10.76%	-0.20%	6.38%
-45	8.28%	1.52%	5.58%
-50	8.65%	2.19%	6.07%
-55	16.48%	4.18%	11.56%
-60	21.62%	1.05%	13.39%

The information from Table 20 above is also displayed in Figure 6 below.

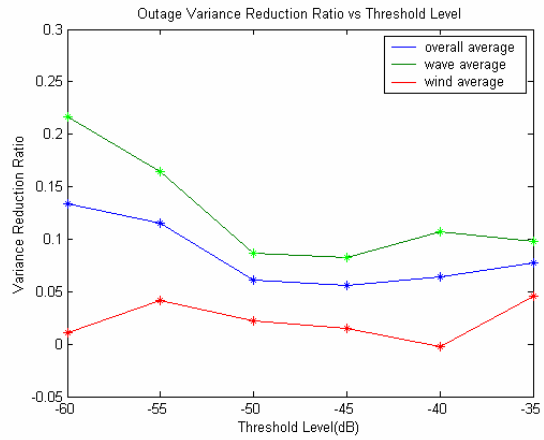

Figure 6: The relative importance of wind variables (bottom line) wave variables (top line) and their average (middle line) in variance reduction ratio results.

DISCUSSION

Relationships Between Energy level and Environment

Applying Results to Oceanography for main three variables

The best three variables to use in variance reduction ratio calculations are wave height, ratio of wave height to water depth and wind direction or relative wind speed. These three variables make sense with oceanographic concepts. If the wave height is very small, then there will not be any breaking waves so there will not be any bubbles in the water and without bubbles nothing will absorb the communications signal. There is a threshold seen in figures such that as long as the wave height stays below a certain level the signal will not be lost. Above this threshold, the signal is sometimes lost.

The ratio of wave height to water depth is important because waves only break when they get big in relation to the water depth. If the water is deep enough, the waves will not break no matter how big they are. As shown in Figure 5, the ratio also has a threshold below which the signal will not dissipate because the waves are not breaking and there are no bubbles in the water. Above this threshold breaking waves sometimes cause outages.

The long shore component of the wind is likely important because it sets up a long shore current that pushes bubbles up and down the coast, a process is called advection. If the bubbles are moving parallel to the coast and communications are happening

perpendicular to the shore, the current is likely to push bubbles into the communications path. In contrast, if there is not a long shore wind the bubbles will not move along the shore and only very localized breaking waves will affect the communications channel. This situation is depicted in Figure 7 below. The wave near the top of the figure will create the displayed bubble field when it breaks. These bubbles will not come between the transmitter and receiver unless there is a long shore wind to push the bubbles south into the communications path.

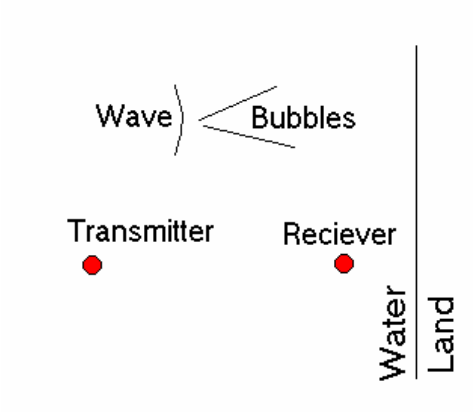

Figure 7 shows the importance of considering long shore current when explaining signal outages

Comparing Mutual Information Results and Variance Reduction Ratio Results

There are a lot of similarities between 1 dimensional normalized mutual information results (Tables 2 and 3) and one dimensional variance reduction ratios (Table 5), especially when looking at wave results. The results for wind have some similarities but at least one significant difference.

Wave Results

Both variance reduction ratios and normalized mutual information results show that a ratio of wave height to water depth typically has better variance reduction than the wave height alone. This trend is more pronounced in variance reduction ratios, but more consistently seen in mutual information results. For example, with variance reduction ratios, the ratio of maximum wave (2 minutes) to water depth is 4.33% better at reducing variance than maximum wave (2 minutes) by itself. With mutual information results, the difference between these variables is only 0.26%. However, when comparing results for the ratio of significant wave height (2 minutes) to tide and significant wave height (2 minutes) by itself, mutual information results are equal, but the variance reduction ratio for significant wave height (2 minutes) by itself is 1.23% higher than the one for ratio of significant wave height to water depth. These results seen with significant wave height in 2 minutes go against the general trend.

Results show a ratio of wave height to water depth is better at reducing variance in energy received than the wave height itself. This makes sense since the ratio of wave height to water depth is actually a combination of the variables tide and wave height. Essentially, creating this ratio is making one variable with some information from both tide and wave height. One would expect the new variable to have a variance reduction between that shown by the initial two variables and this is what the results show. Since tide has a higher mutual information value than wave height (Table 2) including some of this information in with the wave height data will increases the mutual information above what wave height alone presented. As shown in Tables 2 and 3, the ratios do not have as much mutual information as tide, but they have more than wave height alone.

Another similarity shows both 1 dimensional variance reduction and 1-dimensional mutual information values improve when waves are averaged over a longer time period. Without fail wave heights averaged over 2 minutes have better results than those averaged over 1 minute and ratios of wave height to water depth where the wave height is a 2 minute average have better results than their 1 minute counterparts. This trend likely occurs because the bubbles from a breaking wave are likely to stay in the water column for about 2 minutes. Very short time wave averages were some of the first variables to be discarded (Table 1) because the wave conditions in the last 10 seconds to not have much impact on the bubbles in the water column. 2 minutes was the longest average studied in this project, it is possible that longer averages will perform better than the 2 minute averages did.

Wind Results

There are fewer similarities between variance reduction ratios and normalized mutual information results when studying wind results. Both methods of analysis show that relative wind speed is a more important variable than wind speed alone. The difference between relative wind speed and wind speed is more severe for conditional variance results. Both methods show that the average wind speed relative to North is more important than top wind speed relative to North. In variance reduction ratio results, top wind speed performs better than average wind speed, but this result is reversed for mutual information results. No further study has been done on the average and top wind speed variables to determine which is actually more important.

Relative wind speeds likely have higher mutual information values and higher variance reduction ratios for the same reason that ratio of wave height to water depth has better results than wave height alone. Creating a relative wind speed creates a variable that takes both wind direction and wind speed into account. Since wind direction has more information and more variance reduction than wind speed alone, including some of the wind direction information with wind speed creates a variable with more information about the environment than wind information itself. As shown in Tables 2 and 5, relative wind speed end up with a result between the result for wind direction and the result for wind speed.

The biggest difference between the two methods of analysis is the wind direction result. The variance reduction result shows this as the single most important variable with an average of 31.55% variance reduction by itself (Table 5), while the mutual information results show wind direction is relatively unimportant, with only 7.62% mutual information (Table 3).

Variance Reduction Ratio Results and Different Types of Power

The difference between absolute power and power in dB is that when absolute power is less than 1, small differences are amplified by converting the power to dB. An absolute power of 1 maps to 0 dB, and every time absolute power gets a power of 10 smaller, the new value maps to 10 dB below the initial dB result. Every time absolute power decreases to half the initial value, the new value maps to 3dB below the initial value. Table 20 below summarizes this explanation:

Table 20 shows the relationship between absolute power and power in dB

Absolute Power	Power in dB
1	0
0.5	-3.0103
0.25	-6.0206
0.1	-10.0000
0.05	-13.0103
0.025	-16.0206
0.01	-20.0000
0.001	-30.0000
1*10^-4	-40.0000
1*10^-5	-50.0000
1*10^-6	-60.0000

The fact that 1 dimensional variance reduction ratios for wave information are better when power is in dB (Table 6) probably means small changes in the waves make big changes in the amount of energy received. On the other hand, the fact that wind variables have much better 1 dimensional variance reduction ratios when power is absolute probably indicates that small changes in the wind lead to small changes in the energy received.

By the time the results are taking 3 dimensions into account, there is not much difference (1.32% average) between the variance reduction ratios for power in dB and those for absolute power. Since absolute power variance reduction ratios are all slightly larger, we maybe able to say that in general as the environment changes slightly, the energy received changes slightly. However, these results may also mean that the severe energy fluctuations due to wave height and wave ratio changes are balanced out by the more subtle energy variations associated with changes in wind information when all three are taken into account.

Variance Reduction Ratio Results and Different Lengths of Time

Common sense says variance reduction ratio results should vary indirectly with the amount of time considered in the calculations because uncertainty in the environment will only increase with time. Some examples of long-term uncertainty might be bottom profile, current, or temperature. These things do not change quickly but they may change from hour to hour and certainly from day to day. So, as the results expand to include more time, these unaccounted for variables may change and impact the results in ways that do not show up in the result on shorter time intervals.

However, no results exactly fit the expected trend. Average variance reduction ratios should start highest with 18 minute fields, decrease for day long tapes, decrease again for multiple day sets and reach a minimum value when all 5 days are included in total variance reduction calculations. 1 dimensional wind results have their best variance reduction ratio at the set level instead of the tape level. 1 dimensional wave variables have a higher total variance reduction than the average seen in the sets or the tapes. 2 dimensional results are highest for sets, and even the average variance reduction across all tapes is less than that for the total. 3 dimensional results seem to fit the expected trend, the total average variance reduction is lower than the set average.

These results are hard to explain. Perhaps there are not many unaccounted for environmental factors important enough to impact results when they change. If nothing is unaccounted for then having data spread over a longer time period may give a better picture of the overall environment and as a result, variance can be reduced more given more time and more data. Since the

slope of the beach changes between the two sets and this variable in known to be unaccounted for, the fact that variance reduction ratios do typically fall when moving from the set level to a total result is encouraging.

A mathematical explanation for unexpectedly high ratios may come from looking at the bins defined for the environmental variables. If the bins are defined too precisely there will be a very small number of samples in each bin and the variance will be artificially lowered. Imagine the most extreme case where each bin only has one sample, there would be no variance in each bin's environment samples. Regardless of the actual relationship between that environmental variable and the energy received, the sum of weighted set variances will be zero and the variance reduction ratio will be 100% (Equation 6).

To add some numbers to this discussion, at a 1 dimensional level, energy values were sorted into 15 to 20 bins based on the value of their environmental variable. This means at a 2 dimensional level, the energy values are being sorted into 225 to 400 bins and at a 3 dimensional level there could be anywhere from 3375 to 8000 bins that the energy values could fall into. The energy values are linked to about 180 thousand pressure samples. Each pressure sample has 20 or 50 energy samples associated with it, dependent on the sampling rate used for pressure data. Environmental variables tend to form a bell curve when they are placed in bins, so that sets with moderate values are much more common that sets with high or low environmental variable values. In the real world, this means moderate wind speeds are more common than extreme winds or no winds and that moderate wave heights are more common than very small or very large wave heights. Therefore, the bins most likely to be empty or nearly empty are the bins with extreme values for one or more of the environmental variables under consideration. The fact that energy was sampled more quickly than pressure data helps out because even if a bin only has one associated environmental sample, there are at least 20 energy values in the bins to calculate the variance of.

If a study of bin sizes can show that the bins are all adequately sized, the next thing to check is the independence of environmental samples in the bins. The environmental variables studied do not tend to vary quickly, so several consecutive environmental samples are likely to end up in the same bin. If a bin has 10 samples included, but they all came from one 2 second period of time, it's probably not valid to treat the energy values in the bin as if they were independent. It would be nice to have data points from different hours and different days in each bin to get more accurate results.

One thing that did show up as expected when adding more time was that outliers disappeared as more time was added. In other words, in general the 18 minute field with the highest variance reduction ratio had a higher variance reduction ratio than the tape with the best result, and the best tape result was higher than the best set result. The same thing happened with the lowest variance reductions, the lowest field was lower than the lowest tape, which was much lower than the lowest set result.

Variance Reduction Ratio Results and Number of Outages

Variance reduction ratios and mutual information values both increased for tapes with more severe outages. This makes sense because a variable that has a constant value contains no information and has no variance. When a signal is very steady over a period of time, it approaches this hypothetical constant valued signal. Conversely, the more the signal varies, the more information it contains and the more variance it has. It is much easier to find mutual information with, or reduce the variance on, a signal that has information and variance to begin with. This explains the direct relationship seen between number of outages and average mutual information values (Table 4), and the similar relationship between outages and variance reduction ratios (Tables 9 and 14).

Other Variance Reduction Ratio Trends

Looking at 1 dimensional wind variables, the wind speeds most likely to destroy a communications were those at moderate speeds which seems strange because stronger winds would create bigger waves, bigger waves are more likely to break and leave bubbles in the water column that destroy signals. It's possible that this happens because the higher wind speeds are only brief gusts. Gusts of wind do not significantly alter wave heights and so do not effect the bubbles in the water.

Variance reduction ratios always increase with higher dimensions. This occurs because more information can never increase the variance in data. If the information added in a second or third dimension is worthless, the variance reduction will be the same as it was before that dimension was added, but it will never increase. For an example, this means the 3 dimensional variance reduction for wind direction, significant wave height (2 minutes) and ratio of maximum wave height (2 minutes) must be greater than the 1 dimensional variance reduction for any one of these three variables

and must also be more than the 2 dimensional variance reduction for any combination of 2 of these 3 variables.

2 dimensional variance reductions with two wave variables are less than those with a wind and a wave variable. This may be for either of 2 reasons. First and most likely, the wind variables may truly be more important than the wave variables. The 1 dimensional variance reduction results (Table 2) support this conclusion because all three wind variables used in the 2 dimensional calculations had very high 1 dimensional variance reduction ratio results. The second possible solution is that since the wave height and wave height ratio are both based on wave heights and came from the original set of pressure data they are not completely independent of each other. In contrast, wind data is completely independent from pressure data since it came from a separate source and has no reliance on the wave conditions. Since the wind variables are bringing in more independent information, this may explain why the results are better when wind information is one of the variables than when both are wave variables. However, if this were the best explanation, it would seem that the 2 dimensional results with the lowest variance reduction ratios should be those where the wave height is the same as the wave height used in the ratio. However, these are not the lowest results. The lowest results tend to come when the two wave variables are based on different 1 minute averages.

Relationships Between Outages and the Environment

As stated in results above, the results calculated so far for relationships between signal outages and the environment are not entirely dependable. However, so far it seems that rms wave height and significant wave height are the best outage predictors. As the severity of the outage increases, the importance of maximum wave height in 18 minutes increases. This could be because unusually large waves are the ones that create unusually severe outages. Small outages such as those seen at the –35 dB are the easiest ones to predict with wind variables, but even at this level, the wind variables do not predict as well as wave variables. Wind is known to effect small-scale surface waves. It is possible that these waves put some bubbles in the water, not enough to destroy the signal completely, but enough to get it below the –35dB threshold. These slight outages being a result of wind would also help support the idea presented above that wind variables have better variance reduction ratios when power is absolute power because the wind causes small energy changes.

It makes sense that larger outages have higher variance reduction ratios. Outages are an exception to the rule, usually signals should get through the water without problems. Large outages are a rare exception and such rare occurrences require special circumstances to create and almost every time these special circumstances appear, on outage results, which explains the high variance reduction ratios for severe outages.

The variance reduction ratio results would be more reliable if they were based on more data points. They also might be improved if the definition of an outage is modified. Right now any time the signal strength falls below a specified threshold this is called an outage. However, especially at the –35 dB threshold there are a lot of times when the signal will only fall below the threshold for 0.01 seconds, and in reality this may not disturb Naval operations enough to call it an outage. Using this definition of an outage there are some 18 minute time periods with more than 2000 outages on the –35 dB level. Perhaps the definition of an outage should be expanded to include a minimum time below the threshold level or a minimum difference between the threshold and the mean signal level during the outage.

RECOMMENDATIONS FOR FUTURE DIRECTION

From here, the first step is put a bit more effort into reality checking the 3 dimensional energy variance reduction results and the 1 dimensional outage variance reduction results. The 3 dimensional results may be based on data spread over too many bins to support a realistic variance. Even if bins have a good number of points in them, they may not include independent data points if all the points fall too close together in time. The outage results of greatest concern are those showing a negative variance reduction. In the real world variance reductions should not be negative. A negative result would mean knowing more information increased the uncertainty in the energy received. More information can never increase the uncertainty above its original state. These results are likely a consequence of too many outliers increasing the set variance to an unrealistically high level.

It may be worth looking at 3 dimensional variance reductions for sensors other than number 6 which has been used in all analysis so far.

Other variables worth looking at in the relationships between energy and environment are wind on averages shorter averages than the current 5 minute averages, wave information on longer averages than 2 minutes, and information about local currents. Knowing currents would help prove or disprove the theory that wind is

important because it sets up a long shore current and promotes advection of bubbles in the water.

Since tide has a higher average 1 dimensional mutual information than any wave height variable (Table 2) it may be worth running some 3 dimensional calculations with tide as one of the 3 variables, probably using it to replace either the wave height or the wave height to water depth ratio.

There are a lot of variables worth looking at for relationships between outages and environment. To date, no work has been done with ratios of wave height to water depth or with relative wind speeds. More data points will add dependability to the results, and redefining what an outage is will probably help too.

CONCLUSIONS

In conclusion, the strong relationships between energy level and environment show that it is possible to create the model we are working on. The most significant environmental factors seem to be wave height, the ratio of wave height to water depth and wind direction. When wave height data accounts for two minutes of data, the relationships are stronger than when wave information is a shorter average. These results all make sense based on knowledge of oceanographic concepts.

The results from this project so far may be universally applicable to sand bars, because the slope of the sand bar off Scripps Pier changed drastically between data sets one and two and the total results calculated for both time periods look pretty good, averaging 65% variance reduction based on 3 environmental variables. This average can be increased to 75.5% if the wind variable used is wind direction.

The results for relationships between outages and the environment so far show that average and significant wave height are the most important variables in predicting outages. The easiest thing to predict is the mean outage length. Other things that may be predictable are the average signal strength and the number of outages in an 18 minute period. Results show that wind variables are more related to the smaller outages seen at the –35 dB level and that maximum wave in 18 minutes can help predict the largest outages seen at –60dB. Larger outages have higher variance reduction ratios on average.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the contributions of Dr. Preisig, CAPT Hartnett, Dr. Gross, Dr. Wolcin and LT Czerwonka for their aide in this project. Mrs. Mary Crevier also helped make it possible by assuring the logistics worked out.

ATTACHMENTS:

A: Tape summary for environmental variables
B: Tape Summary for 1 dimensional mutual information values
C: 1 dimensional variance reduction ratios
D: 2 dimensional variance reduction ratios
E: 3 dimensional variance reduction ratios
F: Summary of outage variance reduction ratios

REFERENCES

1) Douglas, Scott L., “Influence of wind on breaking waves,” *Journal of Waterway, Port, Coastal and Ocean Engineering*, vol 116, pp. 651-663, Nov-Dec 1990.
2) Wang, Liang, “Approach to calculating methods of wave breaking depth and height,” *China Ocean Engineering*, v 7, pp. 85-98,1993.
3) Kweon, Hyuck-Min, “Parametric model for random wave deformation by breaking on arbitrary beach profiles,” *Proceedings of the Coastal Engineering Conference*, v 1, Part 1, Sep 2-6 1996.
4) Beji, Serdar, “Spectral modelling of nonlinear wave shoaling and breaking over arbitrary depths,” *Coastal Dynamics - Proceedings of the International Conference*, pp 285-294, Jun 1997.
5) Su, M.-Y., Teague, W.J.; Burge, R.E.; and Wesson, J.C., “Variations of breaking wave statistics across the nearshore zone,” *Proceedings of the International Symposium on Ocean Wave Measurement and Analysis*, v 1, pp 510-523, Nov 3-7 1997.
6) Yu, Ke and Svendsen, Ib A., “Breaking waves in surfzones,” *Coastal Dynamics - Proceedings of the International Conference*, pp 329-340, Sep 1995.
7) Durand, N., Allsop, N.W.H., “Effects of steep bed slopes on depth-limited wave breaking,” *Proceedings of the International Symposium on Ocean Wave Measurement and Analysis*, v 2, pp 1400-1413, Nov 1997.
8) Su, Ming-Yang, Wesson, J.C., Teague, W.J., and Burge, R.E. “Breaking wave statistics in nearshore and surf zones,” *Proceedings of the International OTRC Symposium*, p 297-304, Apr-May 1998.
9) Mizuguchi, M., “Breaking of irregular waves on a slope,” *Proceedings of the Coastal Engineering Conference*, v 1, pp. 488-501, Oct 1992.
10) Battjes, J.A., Beji, S., “Breaking waves propagating over a shoal,” *Proceedings of the Coastal Engineering Conference*, v 1, pp. 42-50, Oct 4-9 1992.