

COMPARING HINDCASTS WITH WAVE MEASUREMENTS FROM HURRICANES LILI, IVAN, KATRINA AND RITA

George Z. Forristall

Forristall Ocean Engineering, Inc., 101 Chestnut St., Camden, ME 04343
george@forocean.com

1 INTRODUCTION

The metocean design specifications in the Gulf of Mexico are being thoroughly reevaluated. The reevaluation is motivated by the devastating hurricanes of the last few years. The hurricane hindcasts of Oceanweather, Inc. are the basis for most Gulf of Mexico metocean criteria. For this reason, assessing the accuracy of their hindcasts is an essential part of the ongoing re-evaluation. Hindcasts for Lili, Ivan, Katrina and Rita are compared with every available appropriate set of measurements in this study.

Section 2 describes the sources of our data. The most extensive and best quality data comes from the National Data Buoy Center. They made buoy measurements in the path of all four hurricanes. NRL data gives us some densely spaced measurements directly under the path of Ivan. The industry data give us several sets of continuous observations for Ivan and Rita. Comparisons of the hindcasts to individual data sets are given in Sections 3 - 5. Overall comparisons are given in Section 6. Systematic differences between the hindcasts and measurements can then be separated from issue of sampling variability and random errors.

Estimates of the maximum probable wave height in a storm give an integrated measure of the strength of the storm. Section 7 compares maximum wave height calculations for the hindcasts and measurements. Some of the industry measurements are continuous recordings of wave elevations. Those recordings are compared with short term wave and crest height distributions in Section 8. Conclusions are given in Section 9.

2 DATA SOURCES

Oceanweather's standard, proprietary product for the Gulf of Mexico is GOMOS (Gulf of Mexico Oceanographic Study). GOMOS includes hindcasts of wind strength, wave height and current velocities for all the tropical storms that occurred in the Gulf from 1900 – 2005. Oceanweather made a set of similarly calculated hindcasts for the Minerals Management Service (MMS). This set of hindcasts is for Hurricanes Lili (2002), Ivan (2004), Katrina (2005) and Rita (2005). These hindcasts are in the public domain. We used them for our comparisons with the measurements made during these four hurricanes. The hindcasts are described by Cox et al. (2004), Cardone et al. (2005) and Oceanweather (2006). Tracks of these four hurricanes are shown in Figure 2.1.

Most of the wave measurements that can be used for hindcast verification in the Gulf of Mexico come from buoys operated by the National Data Buoy Center (NDBC). The buoys used in this study are listed in Table 2.1 and their locations are shown in Figure 2.1 as blue crosses. NDBC buoys use a variety of hull sizes and instrument packages. Earle (1996) describes the processing steps used with the different wave systems.

Station	Hull Diameter (meters)	Payload	Wave System	Heave Sensor	Sampling Duration (Minutes)	End Wave Acquisition (minutes after the top of the hour)	Sampling Rate (Hz)
42007	3	DACT	DWA	3/4 g Accelerometer	20	40	2.00
42040	3	DACT	DWA	3/4 g Accelerometer	20	40	2.00
42001	12	ARES	DWPM	Hippy 40, Acceleration	40	44	1.7066
42003	10	ARES	DWPM	3/4 g Accelerometer	40	40	1.7066
42035	3	DACT	DWA	3/4 g Accelerometer	20	40	2.00
42039 (2004)	3	DACT	DWA	3/4 g Accelerometer	20	20	2.00
42039 (2005)	3	ARES	DWPM	3/4 g Accelerometer	40	40	1.7066
42038	3	ARES	(non-directional)	3/4 g Accelerometer	20	50	1.7066

Table 2.1. Measurement systems used aboard NDBC buoys (personal communication Richard Bouchard, NDBC, 2007).

The differences in the payloads and wave systems on the buoys are mainly concerned with directional wave measurements. All of the non-directional measurements that are most important for our comparisons were made with vertically mounted accelerometers. NDBC has devoted considerable effort to the calibration of these systems. Steele et al. (1985) describe the calibration of a buoy equipped with the Hippy sensor. Transfer functions were found that account for the accelerometer, electronics system and hull response. The response amplitude operators (RAO) for the hull were determined by comparing wavestaff measurements or Waverider buoys with the NDBC buoy measurements. Once determined, these transfer functions were assumed to be independent of factors such as the mooring system and water depth. Other sensors and wave systems were calibrated by similar methods.

From May 2004 through May 2005, the Naval Research Laboratory (NRL) conducted an intensive measurement program south of Mobile Bay. This was the Slope to Shelf Energetics and Exchange (SEED) Project (Teague et al., 2007). They deployed 14 current meter moorings in a dense array that spanned water depths from 60 to 1000 m. Six moorings in water depths between 60 and 90 m also had pressure transducers. These particular moorings give us very important records for understanding Ivan. The locations of these six moorings are listed in Table 2.2 and shown as the green trapezoid in Figure 2.1.

Mooring	Latitude	Longitude	Depth (m)
1	29.39	-88.19	60
2	29.43	-88.01	60
3	29.47	-87.84	60
4	29.28	-88.25	88
5	29.34	-88.08	89
6	29.35	-87.89	87

Table 2.2. Locations and depths of the NRL SEED moorings that included pressure measurements.

The SEED pressure transducers were not intended to provide wave data. They operated on a burst sampling schedule of 512 seconds every eight hours, and their depths were too great for ordinary waves to register in bottom pressure. But Hurricane Ivan's very high, long waves passed directly over the SEED array and the pressure transducers recorded them. The first analysis of this aspect of SEED data was reported by Wang et al. (2005).

A few wave sensors on oil industry platforms operated through the hurricanes. They are listed in Table 2.3 and shown as the cyan circles in Figure 2.1.

Site	Operator	Latitude	Longitude	Storm
Marlin	BP	29.107	-87.943	Ivan
Medusa	Murphy	28.392	-89.453	Ivan
Holstein	BP	27.30	-90.55	Rita
Redhawk	Anadarko	27.12	-91.96	Rita
Horn Mtn	BP	28.87	-88.06	Rita

Table 2.3. Locations of industry measurements.

Marlin is a tension leg platform. All of the other platforms are various types of spars, so they all have large columns that diffract the incident wave field. Waves were measured using downlooking wave radars on each of the platforms. Several had two radars. The Marlin and Medusa wave measurements in Hurricane Ivan were discussed by Cooper et al. (2005).

3 NDBC MEASUREMENTS

Buoys 42001 and 42041 were close to and on the right side of Hurricane Lili's track. A comparison of the hindcast for the five closest grid points and the significant wave heights measured at Buoy 42041 is shown in Figure 3.1. The hindcast grid resolution is fine enough that the five hindcasts are nearly identical. The agreement between the hindcasts and measurements is generally very good. But at the peak, the hindcast waves are lower than the measurements. The dashed line in Figure 3.1 and other figures showing time series comparisons is a three hour running average of the measurements. Taking this average eliminates most of the statistical sampling variability discussed in Section 6 but it also eliminates some of the real geophysical variability that is evident in both the measurements and hindcasts.

Figure 3.2 compares the Ivan hindcast at the nearest grid points with the measurements from Buoy 42040. The mooring of the buoy broke at the peak of Ivan (approximately 2100Z on September 15.) The instrument package in the buoy continued to function and its position was monitored by a GPS receiver. According to Richard Bourchard and his colleagues at NDBC (personal communication, 2007), the hydrodynamic transfer function for the buoy should not have been much different after the mooring broke. They think the wave measurements made by the drifting buoy are reliable. Figure 3.3 shows wave spectra before and after the mooring broke along with one of the hindcast spectra. The shapes of the spectra are similar to each other. This agreement supports the conclusion that the buoy's drift did not affect its wave measurements. A Jonswap fit to the measured spectra at the peak of the storm gave $\gamma = 1.54$. The fit to the hindcast gave $\gamma = 1.00$.

The measurements made while the buoy was moored and drifting are distinguished by different colors in Figure 3.2. For times after the mooring broke, the closest grid point to the buoy track in the hindcast was used in the comparison. The hindcast is good for most of the storm, but the hindcast waves were significantly lower than the measurements for three hours at the storm's peak. Two hours after the mooring broke, the hindcast maximum was about 14 m while the measured maximum was 15.96. The measured significant wave height also reached 15.24 m just before the mooring broke.

The highest waves measured in Hurricane Katrina were again at Buoy 42040. Figure 3.4 compares the Katrina hindcast significant wave heights at the four closest grid points with the measured significant wave heights there. Except at the peak of the storm, there is good agreement between the hindcast and

measured. The highest measured significant wave height was 16.91 m. That is the highest significant wave height ever measured by a Gulf of Mexico NDBC buoy. It exactly matches the previous NDBC record made by Buoy 46003 in the northeast Pacific Ocean in January 1991. The maximum significant wave height in the Katrina hindcasts was 13.86 m, 3.05 m less than the measurements.

The hindcast and measured spectra at 42040 in Katrina's peak were again similar. A Jonswap fit to the measured spectra at the peak of the storm gave $\gamma = 2.78$. The fit to the hindcast gave $\gamma = 1.40$. The measured spectra do not give any indication of noise in the measurements.

The measured mean wave directions at the peak of the spectrum were somewhat erratic. The reason for the erratic directions is that some of the parameters of the directional spectra are set to zero in the files. At first, only the lowest frequencies have zero values, but the zeros creep upward until they are all zero during one hour. These erratic directions cast some doubt on the accuracy of the wave heights. But according to Richard Bouchard at NDBC (personal communication, 2007),

“The heights are OK. The directional wave system on 42040 was an older system that has a known limitation that when the yaw is at periods within the spectrum it throws the directions off. Unfortunately our system uses the zero to indicate the data have been rejected.”

The highest waves measured in Hurricane Rita were at Buoy 42001. The comparison with the hindcasts is shown in Figure 4.5. Overall, the hindcast matches the measurements well. For several hours around the peak of the storm, the hindcast waves are higher than the measurements. Katrina damaged Buoys 42003 and 42007 so there are no measured wave heights for Rita at those sites.

Figure 3.6 is a scatter plot of the simultaneous hindcast vs. measured significant wave heights at the NDBC buoys. All of the data points from the time series are included. The least squares fit through all of the data is

$$H_{hind} = -0.0651 + 0.9903H_{meas} \pm 0.9536 \quad (3.1)$$

That fit is strongly influenced by the smaller wave heights that are not very interesting for engineering studies. A more useful least squares fit for measured wave heights greater than 6 m is

$$H_{hind} = -0.4617 + 0.9187H_{meas} \pm 1.4062 \quad (3.2)$$

Using equation (3.2) gives $H_{hind} = 15.16$ m when $H_{meas} = 16.00$ m for a ratio of 0.9476. For measured waves over 6 m, the bias is -0.22 m, the standard deviation is 1.41 m, and the scatter index is 0.17. The scatter index of 0.1684 is good and similar to what we have come to expect from Oceanweather hindcasts. But the very highest waves over 14 m are not hindcast correctly.

The hindcast errors at high wave heights are even more evident in the quantile-quantile (QQ) plot in Figure 3.7. The data points in that plot diverge strongly from the line of equality for measured wave heights greater than 12 m.

Extreme value analyses with the peaks over threshold method use only the maximum significant wave height in each storm. The least squares fit through the peak to peak comparison gives

$$H_{hind} = 1.5984 + 0.7923H_{meas} \pm 1.9001 \quad (3.3)$$

Using equation (3.3) gives $H_{hind} = 14.27$ when $H_{meas} = 16.00$ for a ratio of 0.8919. The bias is 0.57 m, the standard deviation is 1.90 m and the scatter index is 0.18. But the largest hindcast peaks are clearly lower than the measured peaks.

Averaging the measurements over three hours removes most of the sampling variability, but it also removes real geophysical variability. A least squares fit for waves over 6 m in the measured data after averaging over three hours vs. the (un-averaged) hindcasts gives

$$H_{hind} = 0.2668 + 0.9511H_{meas} \pm 1.2820 \quad (3.4)$$

The fit between the measurements and hindcasts is improved somewhat by the averaging. Figure 3.8 is a QQ plot of the averaged measurements against the hindcasts. The hindcasts are still biased low during the highest waves.

4 NRL MEASUREMENTS IN HURRICANE IVAN

Wang et al. (2005) reported that the maximum significant wave height recorded by the NRL pressure transducers during Hurricane Ivan was 17.9 m. They converted the pressure measurements to surface wave heights using linear wave theory. The conversion was truncated at attenuation factors of 1.5%, which gave cutoff frequencies of 0.14 and 0.12 Hz for water depths of 60 and 90 m. This attenuation factor means that the pressure measurements at the highest frequency are multiplied by a factor of 67 to give wave height. This choice is much more aggressive than the usual procedure of limiting the conversion to amplification factors of 10 or less.

In the electronic supplement to their publication, Wang et al. (2005) show a wave spectrum from NRL Mooring 5. It is very noisy at frequencies above 0.08 Hz. Our calculation of the spectrum at the same mooring and time is shown in Figure 4.1. We calculated the spectrum by taking a Fourier transform of the pressure measurements, amplifying it according to linear wave theory, converting to a power spectrum, and averaging over three adjacent frequency bins. The result is slightly different than that of Wang et al. (2005). The difference is probably attributable to differences in spectral averaging technique. But the results are comparable. The figure also shows the concurrent spectra from the Oceanweather hindcast and nearby NDBC Buoy 44040. As shown in Figure 2.1, Buoy 44040 is very close to the NRL moorings. Mooring 5 was on the right side of the storm track and Buoy 44040 was on the left side, so we expect that conditions at Mooring 5 should be somewhat more severe. Because the spectrum from Mooring 5 has only six degrees of freedom, the statistical variability of individual spectral estimates is rather large. The differences at the peak of the spectra are not surprising. But the very large spectral ordinates in the NRL spectra above 0.10 Hz disagree not only with these buoy and hindcast spectra but with the usual shape of such spectra from observations of previous hurricanes.

Spectra from the other five NRL moorings that had pressure sensors also have much more energy at high frequencies than appears in the hindcast or NDBC buoy spectra. We believe that the high frequency energy in the NRL spectra is the result of noise amplification in the pressure measurements rather than the true surface wave spectra. Selecting the best cutoff frequency for the conversion of pressure measurements to wave heights is always difficult. There is a necessary choice between the risk of amplifying noise and the loss of information and energy at high frequencies. For these measurements in relatively deep water with short wave records, the problem is particularly acute. Our strategy was to look for guidance from the buoy and hindcast spectra. Forristall (1981) showed that at frequencies above about 1.5 times the spectral peak, hurricane spectra typically decay as αf^{-4} . The buoy spectra generally agree with this observation. For each NRL spectrum, we therefore picked a frequency for which a

patched f^{-4} tail best agreed with the concurrent buoy spectrum. The results are shown as the dashed blue line in Figure 4.1. The frequency above which the patch was applied usually had an amplification factor of 8 – 10.

Significant wave heights for the NRL measurements were calculated as four times the square root of the area under the spectra, including the patch at high frequencies. Calculated this way, the significant wave height at NRL Mooring 5 at 2200Z on September 15 is 14.87 m. The highest significant wave height in the NRL mooring data is 15.48 m at Mooring 3 at 0000Z on September 16. The scatter index between hindcasts and NRL measurements processed in this way is 0.1656. It is about the same as for the NDBC buoy measurements despite the fact that the shorter time series of the NRL measurements have larger sampling variability.

5 INDUSTRY MEASUREMENTS

All of the industry wave measurements were made using downward look wave radars mounted on a TLP or Spar. Waves the large columns that form these platforms are modified by diffraction. Figure 5.1 shows the measurements made at Medusa by the wave radars mounted on the southeast and northwest sides of the deck of the truss spar. The two time series of measurements differ considerably. According to the hindcasts, waves approached from the east at 1800Z on September 15 and backed to approaching from the northeast at 0000Z on September 16. Diffraction effects from the large column of the spar may have been large, but by midnight the two radars should have had about the same exposure. It is hard to see how diffraction can explain the difference between the measurements. Because it is not clear which of the measurements is most accurate, both of them are included in our statistics.

Rita's waves were continuously recorded on the north and south sides of the Redhawk cell spar. Figure 5.s compares the hindcast wave heights at the nearest grid point with the two radar measurements. For most of the storm, the south radar recorded slightly higher waves than the north radar. But at 1300Z on September 23 the significant wave height measured by the south radar was much higher than that measured by the north radar. An inspection of the raw time series showed that the signal from the south radar during that hour was extremely noisy. Less severe noise was found in other hours. Therefore, only the measurements from the north radar were used in the statistical comparisons.

A least squares between all of the good industry significant wave heights and the hindcasts is

$$H_{hind} = -0.3587 + 1.0114H_{meas} \pm 1.0770 \quad (5.1)$$

The bias between the hindcast and measured wave heights is 0.0396. The scatter index is good at 0.1227.

6 STATISTICAL COMPARISONS

Figure 6.1 is a combined scatter plot of all of the data from the NDBC buoys, NRL pressure gauges, and industry measurements. Hindcasts significant wave heights are compared to measurements at the same time. The least squares fit through all points with measured significant wave heights greater than 6 m is

$$H_{hind} = 0.0909 + 0.9764H_{meas} \pm 1.2879 \quad (6.1)$$

The bias between the hindcasts and measurements greater than 6 m is -0.1128. The scatter index is 0.1493. Figure 6.2 is a QQ plot of the data. The agreement between the distributions of the hindcast wave height and the measured wave heights is excellent.

A least squares fit through the maximum hindcast and measured significant wave height in each data set gives.

$$H_{hind} = 0.8510 + 0.8810H_{meas} \pm 1.6659 \quad (6.2)$$

The bias between the hindcasts and measurements is 0.4512 m. The scatter index is 0.1549. That is about the same as the scatter index for all of the data.

All of the comparisons are affected by the measurement's sampling variability. Sampling variability is an important consideration in calculating spectra. But it also has an important effect on the estimation of the significant wave height. This is because the significant wave height is calculated from the integral of the spectrum.

Sampling variability is different from measurement error. The significant wave height is a random variable. Consider a storm in which the environmental conditions are constant for many hours. The significant wave height calculated from different time intervals in that storm will still be a random variable, just as the individual wave heights are a random variable. In a hurricane, the conditions vary rapidly. But if the same storm were repeated many times, the significant wave height in a given time interval would be different for each realization of the storm. Hindcasts cannot predict this variability. They are intended to find the ensemble average of many repeated storms. In comparing hindcasts and measurements, we must always remember that the measurements have sampling variability and the hindcasts do not.

The distribution function for significant wave height depends on the record length and the shape of the spectrum. The spectra in the high sea states during hurricanes can be fit using a Jonswap spectrum with a peak enhancement factor of about 1.5. For spectra with that shape, the coefficient of variation (COV) of the significant wave height is approximated well by

$$COV = \frac{0.509}{\sqrt{f_p T}} \quad (6.3)$$

where f_p is the frequency at the peak of the spectrum and T is the length of the record. The main factor determining the COV is the record length. Those buoys with 40 minute records have a COV of 3% – 4% while those with 20 minute records have a COV of 5% – 6%. The 512 second records from the NRL pressure measurements have a COV of 8% - 9%. The industry measurements were from fairly long records and their COV is smaller. The square of the scatter index between the hindcasts and measurements is equal to the square of the COV plus the square of the error between the hindcasts and measurements. For a COV of 5% and a scatter index of 15%, the predicted error is about 14%, so the sampling variability does not materially affect the error calculation in equation (6.1).

Sampling variability is more important when comparing peak significant wave heights in a storm. Forristall et al. (1996) showed that sampling variability produces a bias in the peak measured significant wave height. If there are several samples of H_S at times near the peak of the storm, the maximum of these samples is very likely larger than the expected value of H_S at the peak. The magnitude of the bias depends on the spectral shape, the sampling interval, and the variation of H_S near the peak of the storm.

Following Forristall et al. (1996), we fit the variation of H_s as a parabola and estimated the bias through Monte Carlo simulations. Some of the storms such as Hurricane Lili have very narrow peaks so the bias is very small. The bias can even be negative because non-continuous sampling may miss the true peak of the storm. In general, the expected bias in the peak values of H_s is about 4%. The observed bias between the hindcast and measured peaks was 0.4512 m. The mean measured peak was about 11 m, which gives a predicted bias of 0.44 m. The bias in the maxima can be explained by sampling variability. This indicates that the hindcast peaks are actually unbiased.

7 ESTIMATES OF MAXIMUM WAVE HEIGHTS

As discussed in the previous section, sampling variability in the measurements causes some difficulty in comparing maximum hindcast and measured significant wave heights. It is possible to avoid this problem by calculating an integrated measure of the strength of the storms. This can be done in a way that also has engineering significance by calculating the most probable maximum wave height in the storms. The best way to do this calculation is to use the Borgman integral. The idea is to integrate a short term distribution of wave heights over the history of the significant wave heights in the storm.

Because we are interested in a comparison between hindcast and measured wave heights, the exact form of the short term distribution is actually not important. Nevertheless, it is good to use a realistic one. We used the empirical distribution suggested by Forristall (1978) that has been shown to agree with many observations. It is given by

$$P(h) = \exp\{-2.262(h/H_s)^{2.126}\} \quad (7.1)$$

The probability distribution for the maximum wave height in each measured and hindcast storm history was then calculated by taking the Borgman integral over equation (7.1). Figure 7.1 is a scatter plot of the estimated maxima from the hindcasts vs. the estimated maximum wave heights from the measured time series. A least squares linear fit to the scatter plot gives

$$H_{hind} = 1.3157 + 0.9030H_{meas} \pm 2.5300 \quad (7.5)$$

The bias between hindcast and measured values is -0.4075. As expected, integrating over the storm histories eliminates the positive bias found in the comparison of maximum significant wave heights, but it introduces a similar negative bias. The scatter index between the maximum wave height estimates from hindcasts and measurements is 0.1424, about the same as for the significant wave heights.

8 SHORT TERM STATISTICS

The continuous measurements at some of the industry sites make additional verifications of short term wave and crest height distributions possible. For example, measurements at the Marlin TLP were made at a 4 Hz sampling rate. The wave height distribution from the southwest radar for 0300 to 2030 local time is given in Figure 8.1. The crest height distribution for those times is given in Figure 8.2. These figures exclude the data early and late in the storm when there were some noise spikes in the data. Individual waves were calculated using the zero down-crossing method. The wave and crest heights in each 30 minute segment were normalized by the significant wave height at that time.

The sample distributions of wave height are compared to the Rayleigh distribution and the Forristall (1978) empirical distribution in Figure 8.1. The wave heights from 0300 to 2030 local time agree well with the empirical distribution. The sample distributions for crest height are compared to the Rayleigh distribution and a distribution derived from second order simulations of directionally spread waves (Forristall, 2000) in Figure 8.2. The measured wave crests agree well with the second order distribution.

The most interesting feature of these comparisons is that the presence of a TLP or spar near the measurement location appears to have no effect on the short term distributions. Van Iperen et al. (2004) saw similar results in model experiments on a concrete gravity structure with large columns. The structure did not seem to affect the wave and crest statistics except in very extreme waves when wave breaking appeared to truncate the distributions. The agreement means that the standard wave and crest height distributions can be used for spars and tension leg platforms.

9 CONCLUSIONS AND RECOMMENDATIONS

For years, Oceanweather hindcasts have been the standard tool for estimating design wave heights in the Gulf of Mexico. This detailed comparison with measurements made during the four recent devastating storms give error statistics similar to those found in previous comparisons of Oceanweather hindcasts with measurements. Reliable extreme values can be developed from the hindcasts. But the under-prediction of the two highest peaks in the measurements indicates that more research is needed on wave generation in the most extreme conditions.

Statistical comparisons for all of the measured significant wave heights (Figures 6.1 and 6.2) show that the hindcast bias is negligible. The least squares fit between hindcast and measured wave heights is close to equality. The scatter index of 15% is about what we have come to expect from high quality hurricane hindcasts. The storm maxima statistics shown in Figures 6.3 and 6.4 are slightly biased, but the bias is attributable to measurement sampling variability.

Estimates of maximum wave heights from time series of hindcast and measured significant wave heights agree well. Integrating over the storms to find the estimated maxima eliminated the bias due to significant wave height sampling variability. Comparing estimated maximum wave heights is a good method for comparing hindcasts and measurements because the estimated maximum is often used in design studies.

Standard short term wave and crest height distributions agree well with measurements on industry platforms. We recommend the continued use of these distributions.

Good as they are, the overall statistics may not give the best picture of the skill of the hindcasts in the most extreme waves. Considering only the NDBC measurements, there are several measurements and no hindcasts above 14 m (Figure 3.6). These data points are from Buoy 42040 in Hurricanes Ivan and Katrina. The NDBC wave experts find no reason to doubt the measurements. The large errors are at the peaks of the storms. This finding is consistent with that of Cardone et al. (1996) in hindcasts of two extreme extra-tropical storms. They speculated that the cause of the errors lay in dynamic fetch associated with intense surface wind maxima or jet streaks that were not modeled with sufficient detail. Fine scale features of hurricane wind fields may cause similar problems. A renewed effort to understand wave generation in the most extreme conditions is recommended.

The hindcasts agree with the measurements for all except a few points on the tail of the distribution. Fits to the Oceanweather hindcast should give accurate estimates of extreme conditions. Design wave heights

are calculated by fitting extreme value distributions to 50 or 100 years of hindcast data. The fits are made data from all strong storms and the design conditions are found by extrapolating the fitted distribution. Our comparisons show the highest few points in the extreme value plot will probably be too low. We recommend that extreme value analyses based on the GOMOS hindcasts should concentrate on fitting the body of the distribution without using outliers on the tails. This recommendation agrees with the best practice that extreme value estimates should not emphasize tail fits. Based on the evidence presented here, fits using hindcast values from 6-8 m to 12-13 m may give the best results.

Finally, there are some conclusions and recommendations concerning wave measurements.

The high quality data from the NDBC buoys has been carefully calibrated. The systems on them have been perfected over many years. The data would be even more valuable if continuous wave records were recorded. Transmitting continuous records in real time would be difficult, but it should be possible to store the data on the buoys for later analysis.

The NRL pressure measurements are valuable because they give closely spaced data directly under the path of Hurricane Ivan. They were not intended to make wave measurements so they had short record lengths. It is difficult to convert deep water pressure measurements to wave elevations. The difference between our interpretation and the interpretation in the NRL publications suggests significant uncertainty in the conversion.

The particular value of the industry measurements is that many of them are continuous records. That made it possible to study the short term distributions of wave and crest heights. All the industry measurements were made from structures with large members. The wave field in the vicinity of those structures must be distorted by diffraction effects. This distortion is illustrated by the fact that measurements on opposite sides of the structures differ considerably. It would be valuable to conduct diffraction calculations or tank tests to understand how the measurements near the structures differ from the waves in the free field.

10 ACKNOWLEDGEMENTS

This study was funded by the American Petroleum Institute. We appreciate their support. We thank Bill Teague and D.W. Wang for providing the raw NRL pressure measurements under a cooperative research agreement. Richard Bouchard and his colleagues at NDBC were very helpful in explaining the details of their buoy measurements. We thank Anadarko, BP and Murphy for supplying the industry wave measurements.

11 REFERENCES

- Cardone, V. J., R. E. Jensen, D. T. Resio, V. R. Swail and A. T. Cox. (1996), Evaluation of contemporary ocean wave models in rare extreme events: Halloween storm of October, 1991; Storm of the century of March, 1993. *J. of Atmos. and Ocean. Tech.*, **13**, 198-230.
- Cardone, V.J., A.T. Cox, K.A. Lisaeter and D.S. Szabo (2004), Hindcast of winds, waves and currents in northern Gulf of Mexico in Hurricane Lili, *Proc. Offshore Technology Conference*, Houston, TX.

- Cooper, C., J. Stear, J. Heideman, M. Santala, G. Forristall, D. Driver and P. Fourchy (2005), Implications of Hurricane Ivan on deepwater Gulf of Mexico metocean design criteria, *Proc. Offshore Technology Conference*, **OTC 17740**, Houston, TX.
- Cox, A.T., V. Cardone, F. Counillon, and C. Szabo (2005). Hindcast study of winds, waves and currents in northern Gulf of Mexico in Hurricane Ivan, *Proc Offshore Tech. Conf.*, **OTC 17736**, Houston Texas,
- Earle, M.D. (1996), Nondirectional and directional wave data analysis procedures, NDBC Technical Document 96-01, National Data Buoy Center, Stennis Space Center, MS.
- Forristall, G.Z. (1978), On the statistical distribution of wave heights in a storm, *J. Geophys. Res.*, **83**, 2353-2358.
- Forristall, G.Z. (1981), Measurements of a saturated range in ocean wave spectra., *J. Geophys. Res.*, **86**, 8075 – 8084.
- Forristall, G.Z. (2000), Wave crest distributions: Observations and second order theory, *J. Phys. Oceanogr.*, **30**, 1931-1943.
- Forristall, G.Z., J.C. Heideman, I.M. Leggett, B. Roskam, and L. Vanderschuren (1996), Effect of sampling variability on hindcast and measured wave heights, *J. Waterway, Port, Coastal and Ocean Eng.*, **122**, 216-225.
- Oceanweather (2006), Hindcast data on winds, waves and currents in northern Gulf of Mexico in Hurricanes Katrina and Rita, consultant report to Minerals Management Service, Herndon, VA.
- Steele, K.E., J. C.-K. Lau, and Y.-H. L. Hsu (1985), Theory and application of calibration techniques for an NDBC directional wave measurements buoy, *IEEE J. Oceanic Eng.*, **OE-10**, 382-396.
- Teague, W.J., E. Jarosz, D.W. Wang, and D.A. Mitchell (2007), Observed oceanic response over the upper continental slope and outer shelf during Hurricane Ivan, *J. Phys. Oceanogr.*, in press.
- Van Iperen, E.J., Forristall, G.Z., Battjes, J.A., and Pinkster, J.A. (2004), Amplification of waves by a concrete gravity sub-structure: linear diffraction analysis and estimating the extreme crest height, *Proc. Proceedings of OMAE 2004:23rd International Conference on Offshore Mechanics and Arctic Engineering*, **OMAE2004-51022**, Vancouver.
- Wang, D.W., D.A. Mitchell, W.J. Teague, E. Jarosz and M.S. Hulbert (2006), Extreme waves under Hurricane Ivan, *Science*, **309**, 896.

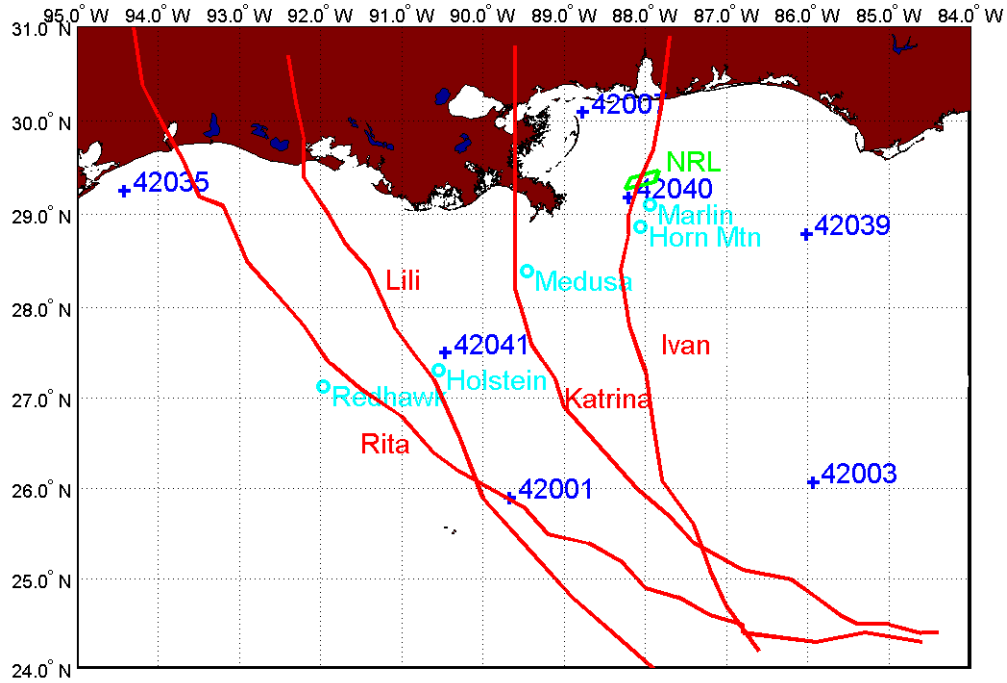


Figure 2.1. Storm tracks and locations of measurement

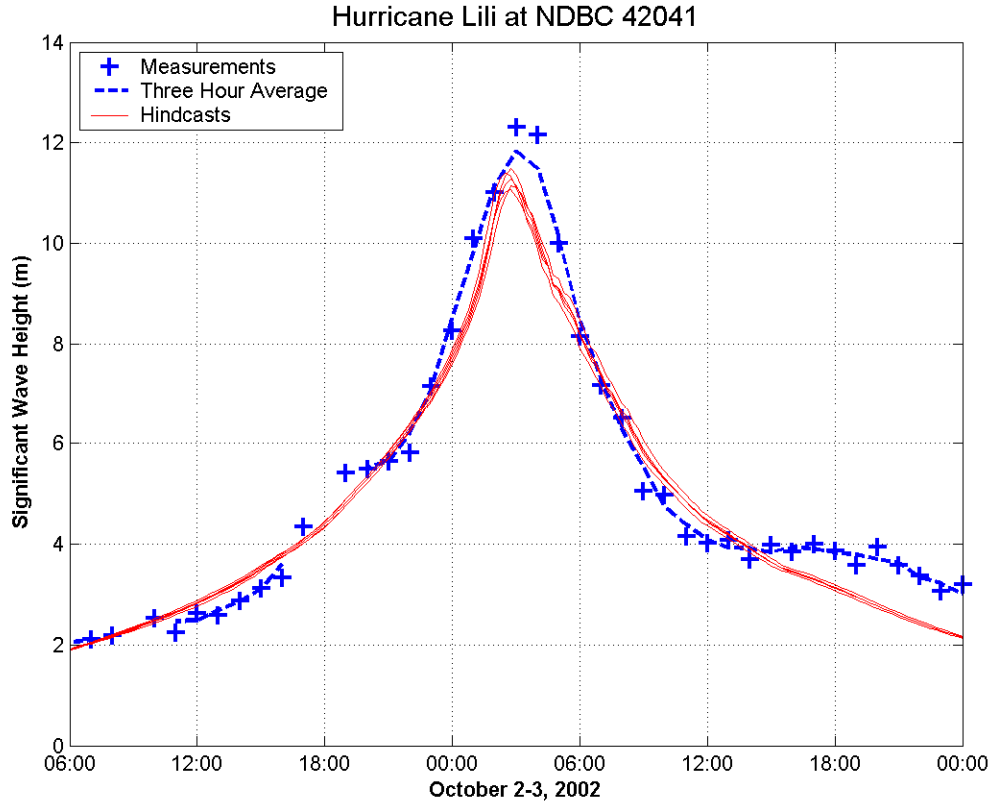


Figure 3.1. Measured and hindcast significant wave heights at Buoy 42041 in Lili.

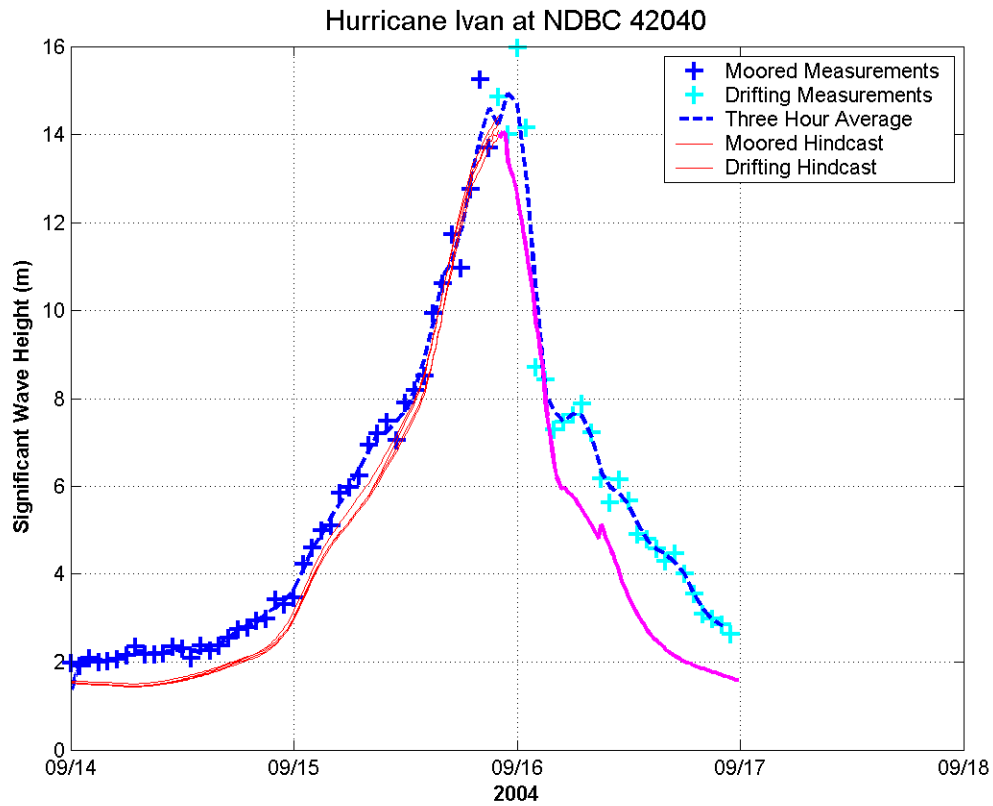


Figure 3.2. Measured and hindcast significant wave heights at Buoy 42040 in Ivan.

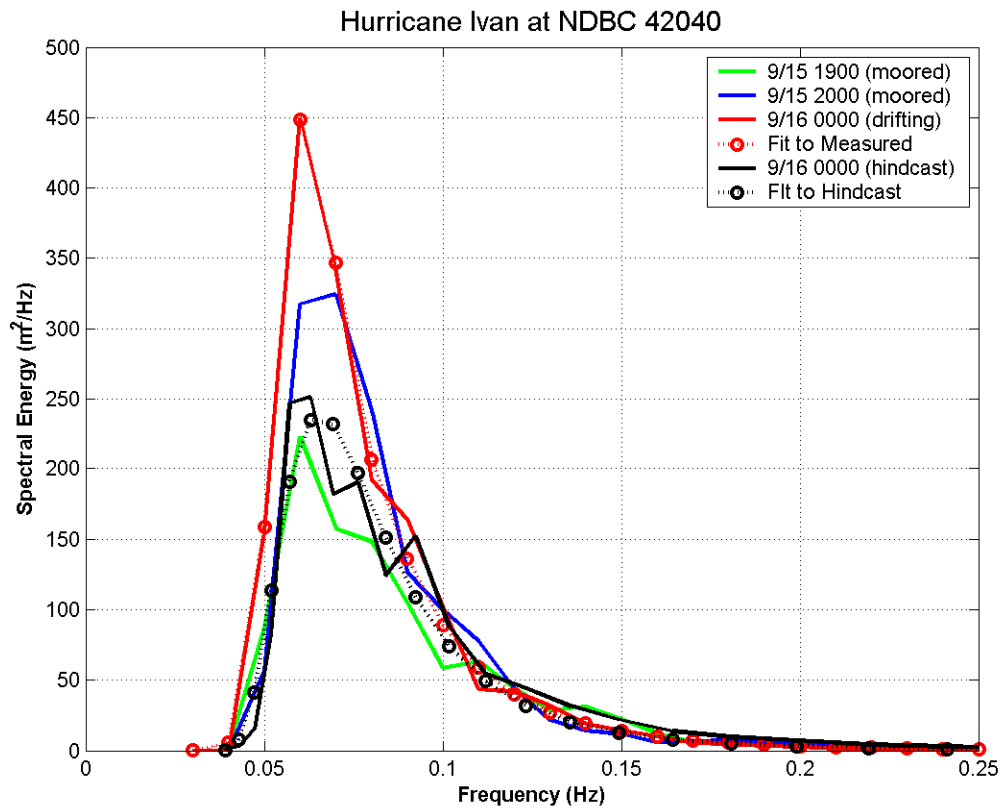


Figure 3.3. Measured and hindcast spectra near the peak of Hurricane Ivan at Buoy 42040. The Jonswap fit for the measurements has $\gamma = 1.54$. The fit for the hindcast has $\gamma = 1.00$.

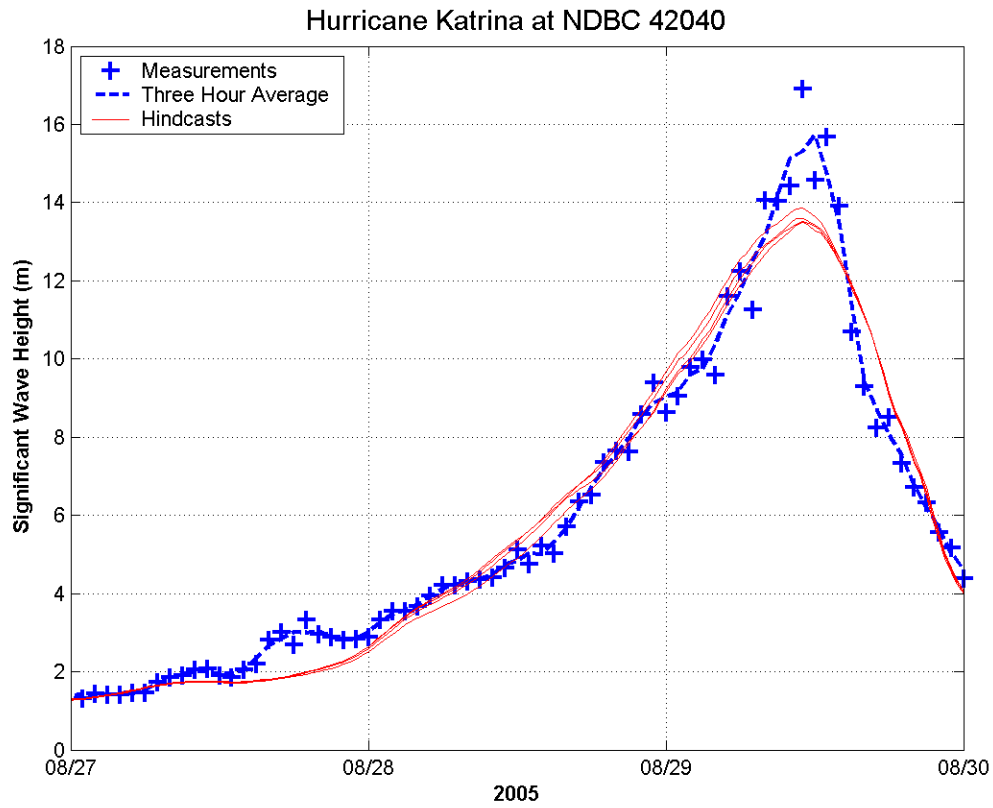


Figure 3.4. Measured and hindcast significant wave heights at Buoy 42040 in Katrina.

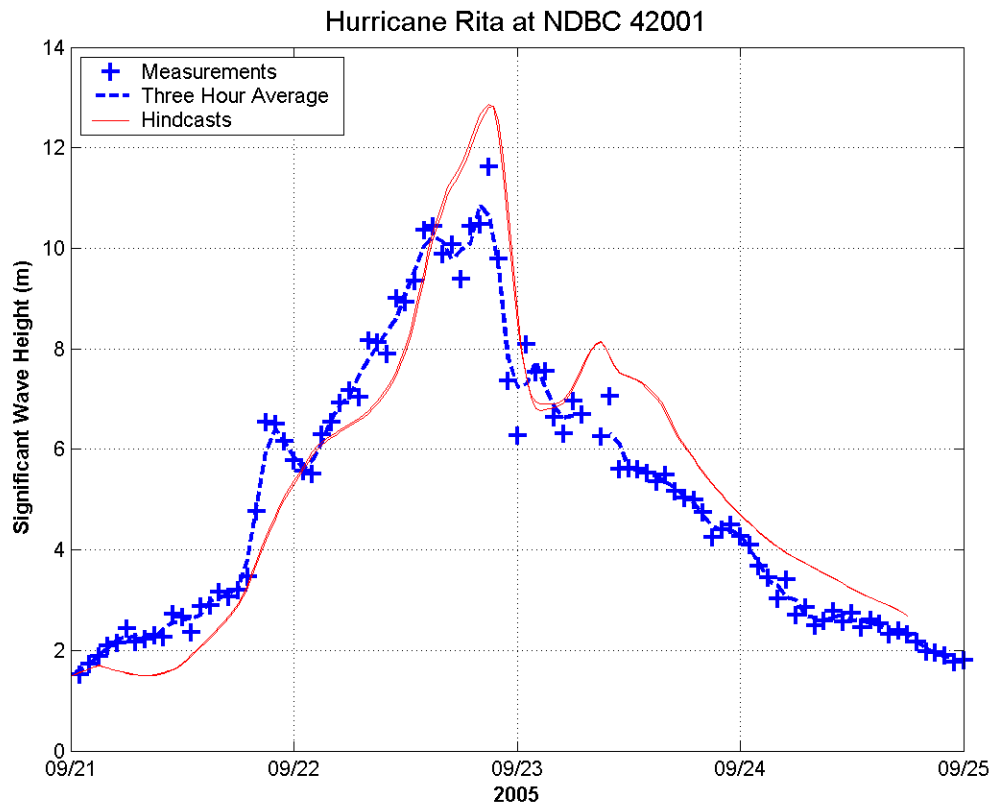


Figure 3.5. Measured and hindcast significant wave heights at Buoy 42001 in Rita.

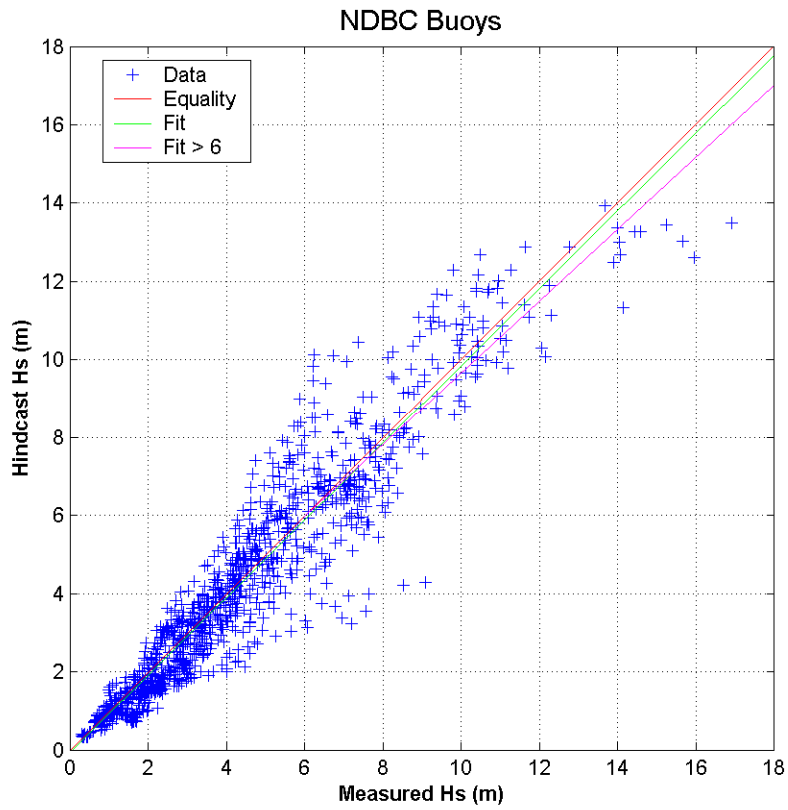


Figure 3.6. Scatter plot of measured vs. hindcast significant wave heights at NDBC buoys.

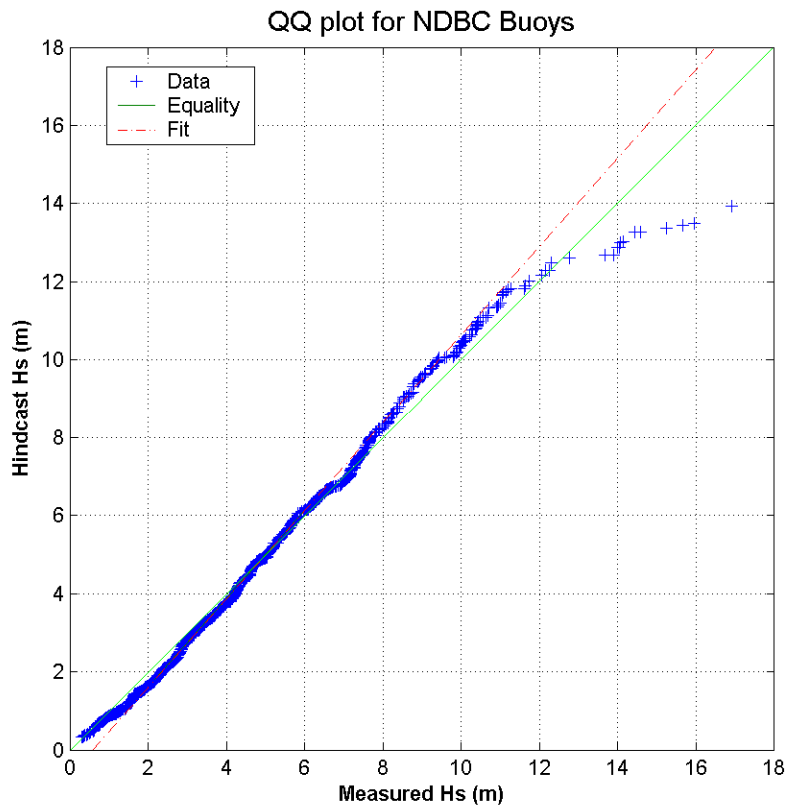


Figure 3.7. QQ plot of measured vs. hindcast significant wave heights at NDBC buoys.

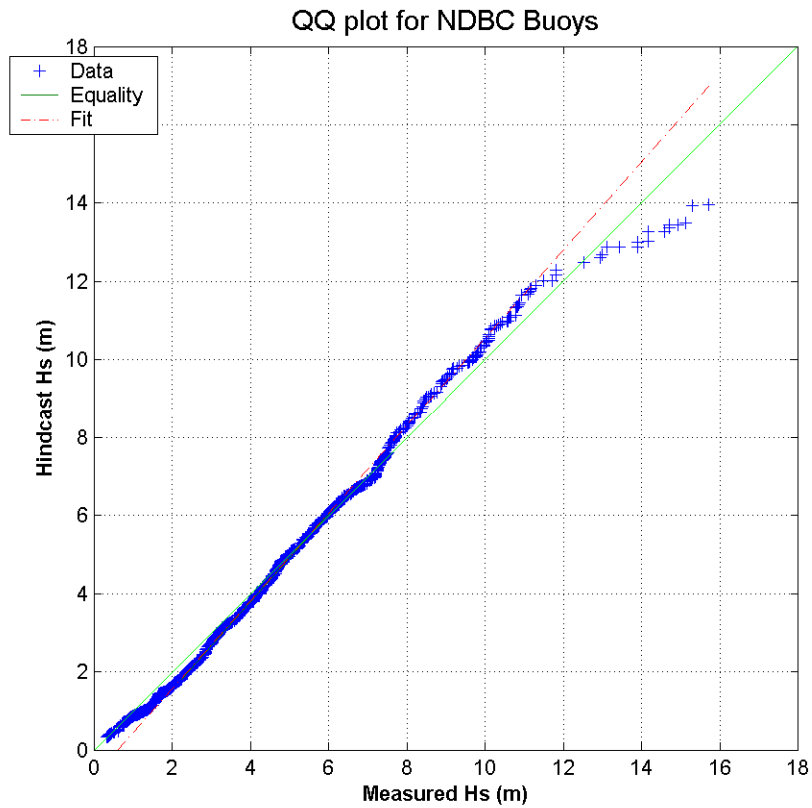


Figure 3.25. QQ plot of 3 hour averaged measurements vs. hindcast significant wave heights at NDBC buoys.

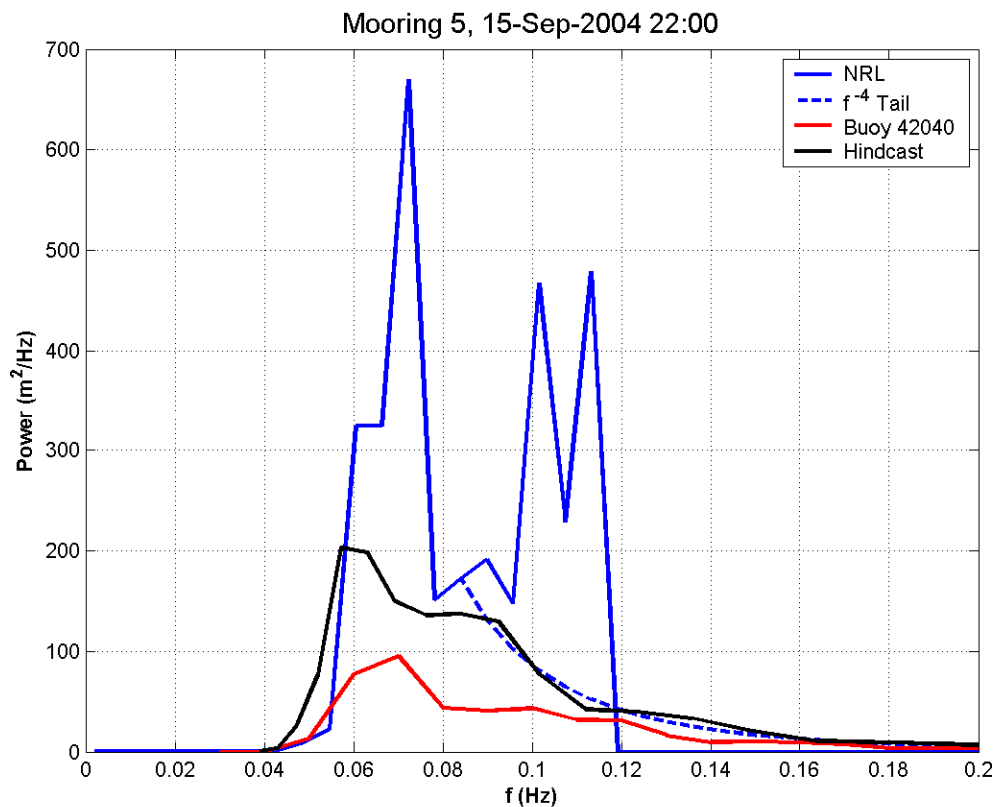


Figure 4.1. Wave spectra at NRL Mooring 5 at 2200Z on 15 Sep 2004.

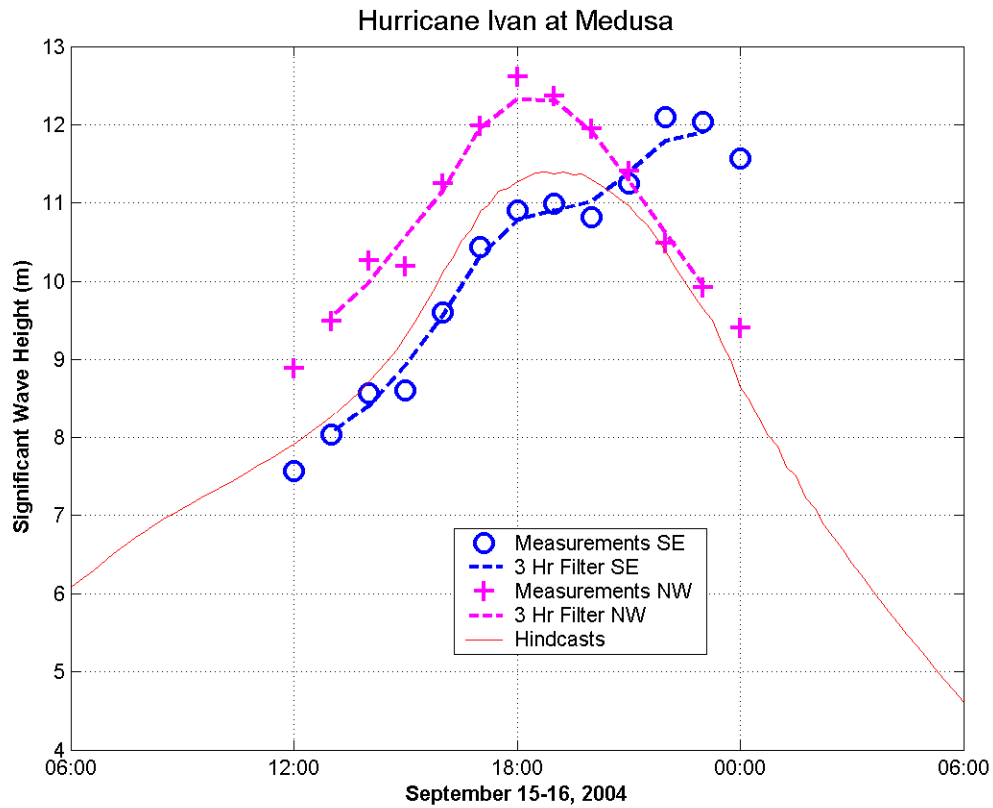


Figure 5.1. Measured and hindcast waves in Hurricane Ivan at Medusa.

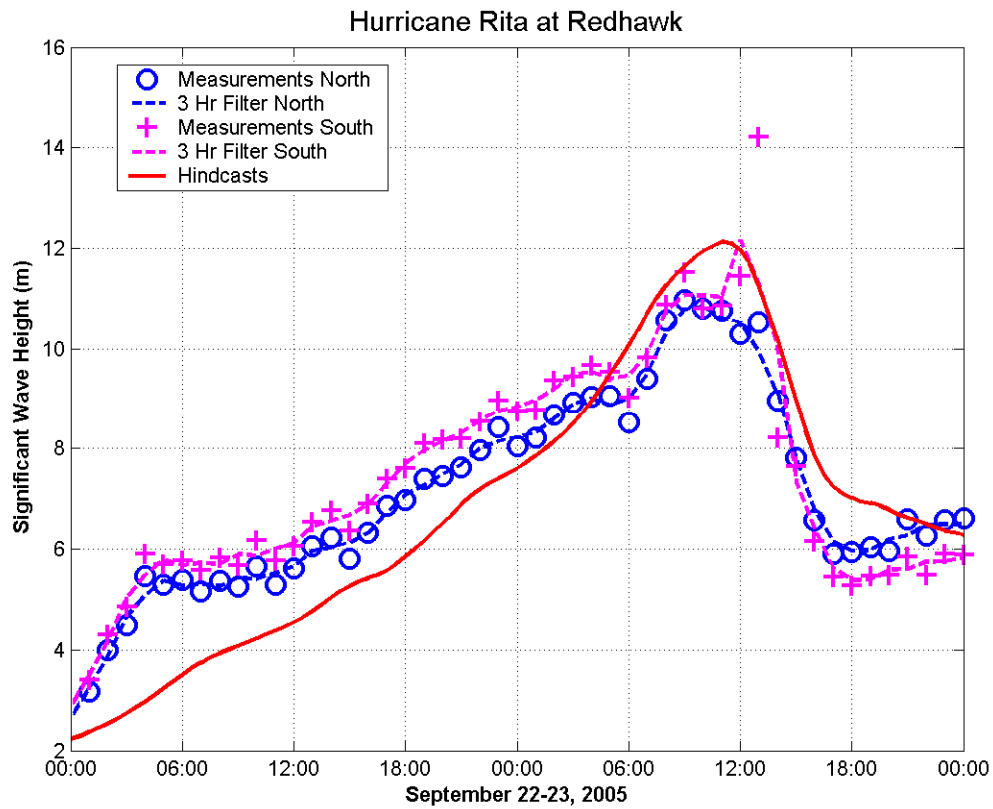


Figure 5.2. Measured and hindcast waves in Hurricane Rita at Redhawk.

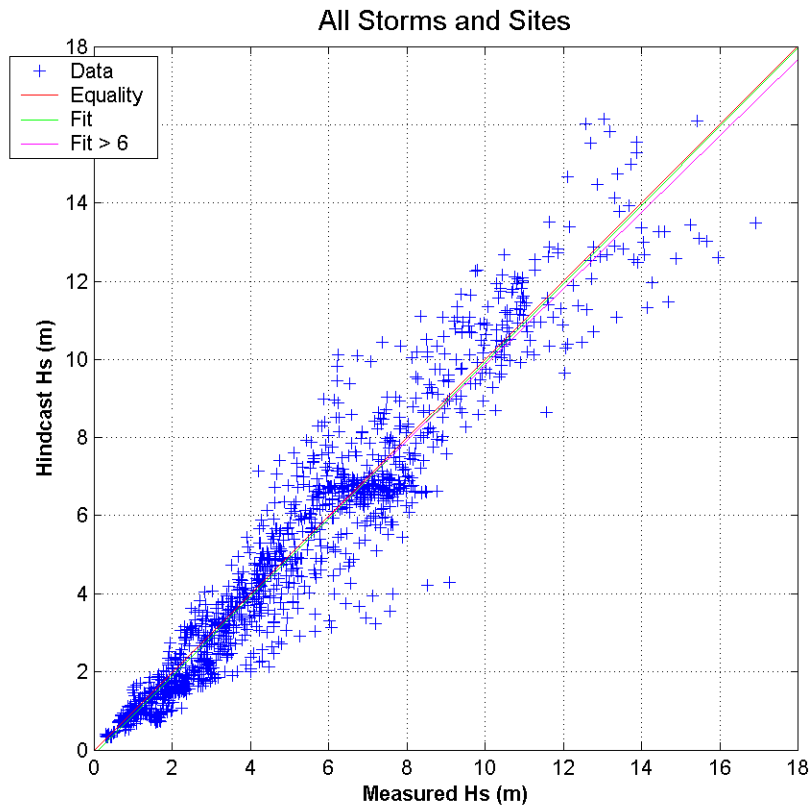


Figure 6.1. Scatter plot of all measurements vs. hindcasts.

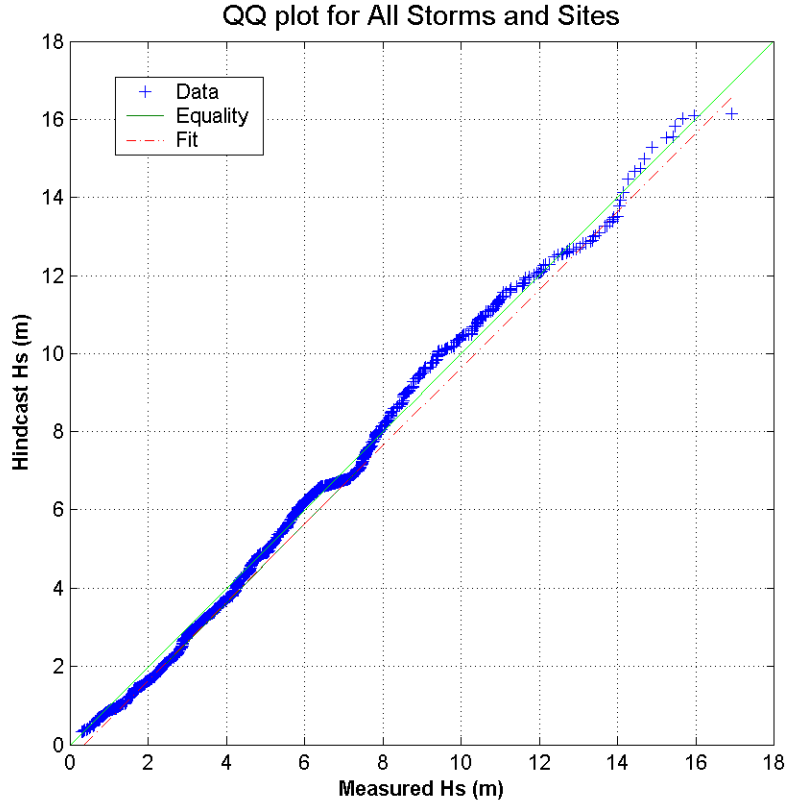


Figure 6.2. QQ plot of all measurements vs. hindcasts.

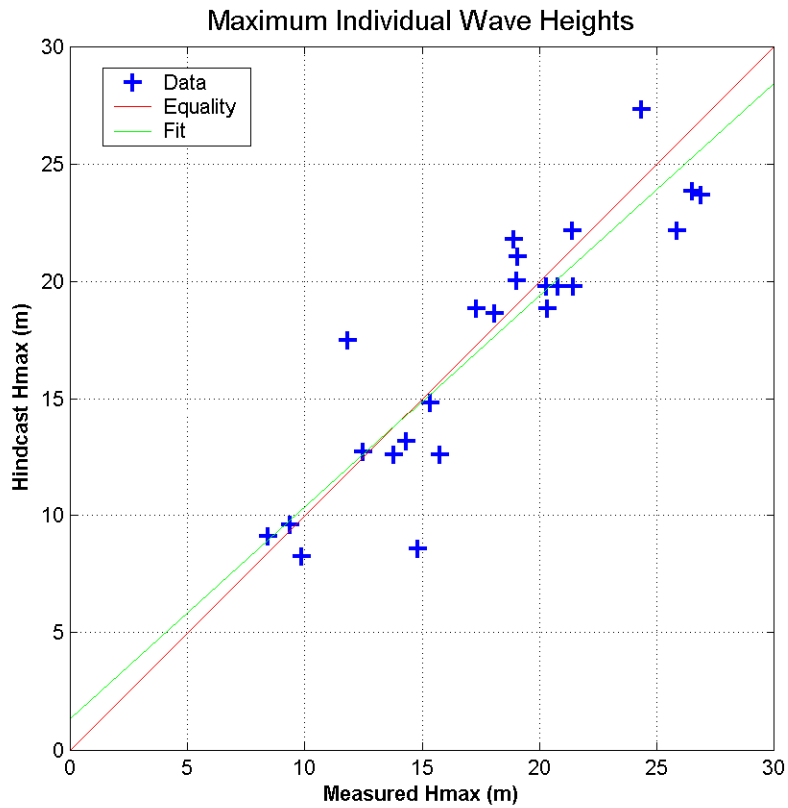


Figure 7.1. Scatter plot of estimated maximum wave heights from measurements and hindcasts.

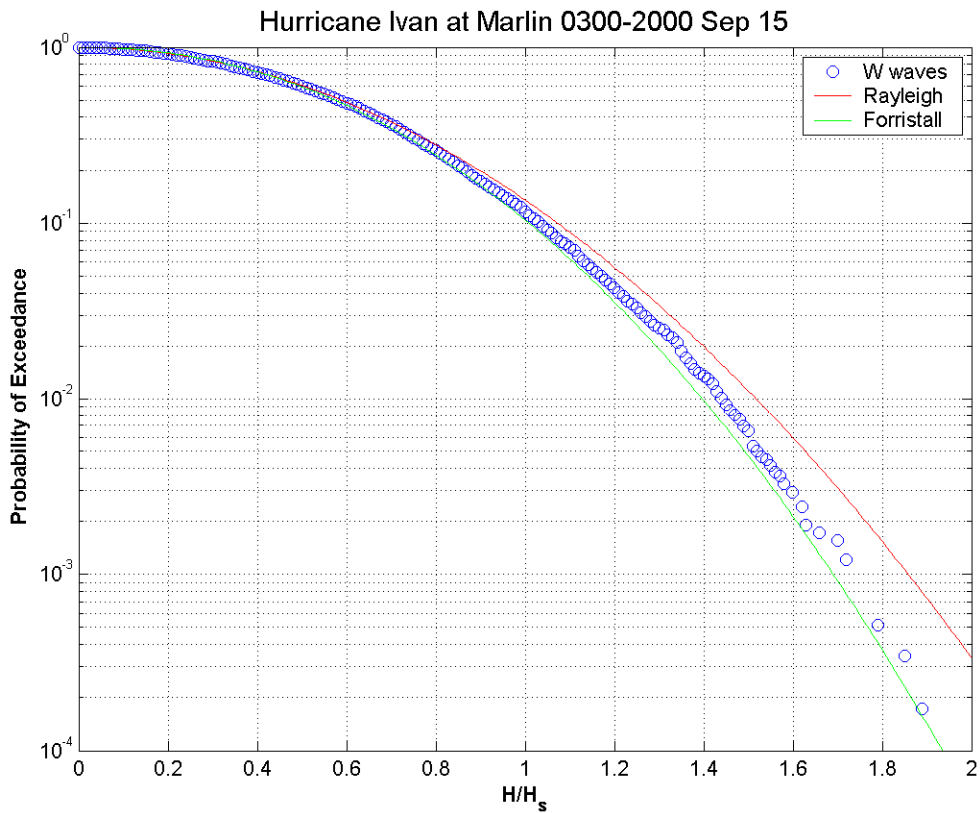


Figure 8.1. Wave height distribution at Marlin.

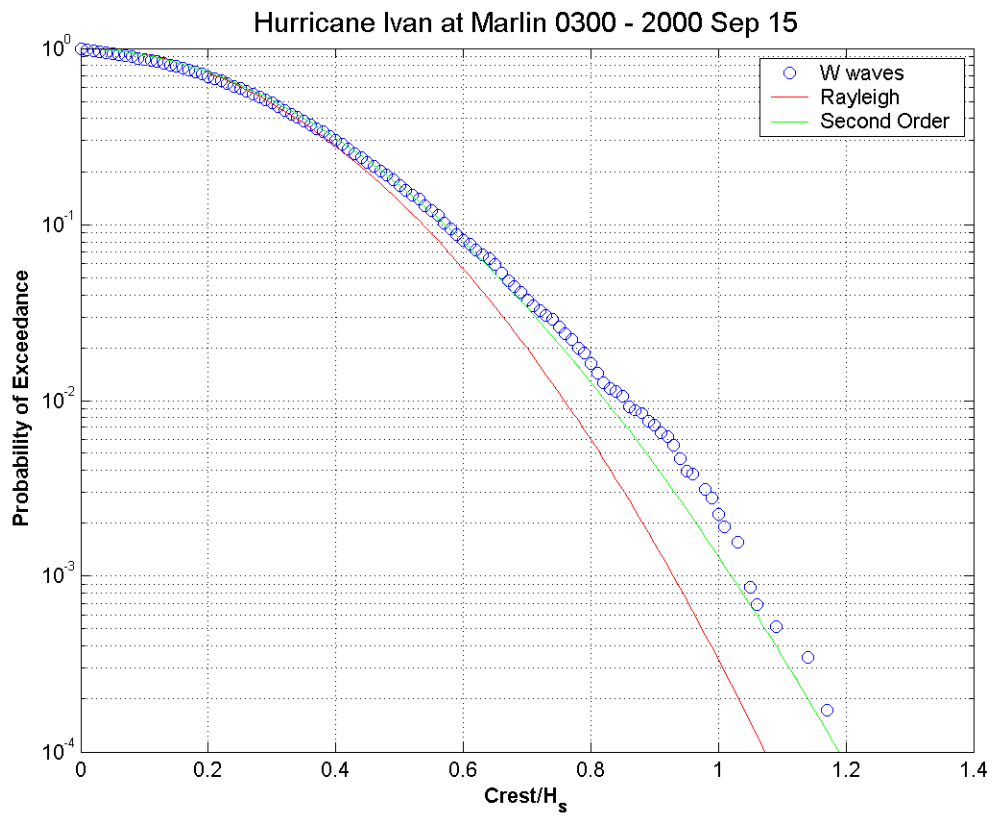


Figure 8.2. Crest height distribution at Marlin.