Fitting Statistical Distributions to Data in Hurricane Modeling

Mark E. Johnson
Department of Statistics
University of Central Florida
Orlando, FL 32816-2370
mejohnso@mail.ucf.edu

Charles C. Watson, Jr.
Kinetic Analysis Corporation
1204 E. 49th Street
Savannah, GA 31404-4014
ccwatson@mail.methaz.net

Abstract

Fitting probability distributions to hurricane related data is an essential activity in hurricane planning, designing structures, and catastrophe modeling applications. The recent devastating hurricane seasons and the resultant debates over design criteria and computer models using analyses based on these fits motivate a closer examination of this issue. The primary objective in this paper is to describe the background and applications of historical hurricane data fitting, the operational aspects of which have dictated adjustments to the standard methodology. The emphasis here is on the interaction between data quality and dynamics, the need for rapid but stable assessments of that data, and statistical fitting methodologies. Validation and applications are discussed, along with an analysis of the return periods of damage in the New Orleans area.

1. Introduction

The 2005 Atlantic hurricane season was record breaking in terms of number of storms, extent of damage and impact on the insurance industry. The financial hits taken by insurance and re-insurance companies have generated a flurry of commentary on catastrophe models which are the basis for establishing residential premiums for wind perils from hurricanes. Rising water damage such as storm surge or riverine flooding is not covered by traditional insurance policies. However, since damage by wind or rain intrusion is covered, arguments and legal actions have started over not only the extent of damage but the timing and cause of the damage (if the structure failed by wind before failing from water, the private company pays; if the reverse, Federal flood insurance covers the loss). Since many flooded areas were outside Federal flood zones, which are in theory based on a 100 year return period event, homeowners are left with massive out of pocket losses. The following comments from Matthias Weber are typical:

> "Everyone's a little bit disappointed with the fact that the modeled losses were significantly smaller than the actual losses, and … people are saying the models aren't reliable. … What industry needs to do is demand improvement of computer models." (from Best's Review, 2005).

Of course, when financial losses or loss of life are staggering, design criteria and computer models offer an inviting target for venting. In their December 2005 issue, Best's Review noted that AIR Worldwide Corp., a commercial catastrophe modeler, had cited the quality of exposure data as a more pertinent explanation for the discrepancies between modeled and actual losses. There is clearly intense interest in computer modeling of hurricanes and likewise interest in data sources and ultimately distribution fitting. For further insights and a general discussion of hurricane forecasting and planning issues, see Iman, Johnson and Watson (2006).

Since 1996, the authors have been working jointly on fitting distributions to hurricane phenomena or effects in the context of operational, funded projects. Our initial joint effort (Watson and Johnson, 1997) concerned fitting extreme value distributions to wind and storm surge data at Montego Bay, Jamaica. In this endeavor both statistical expertise and numerical modeling and geophysics were given equal emphasis and an effective team was created. This effort was quickly expanded to the full Atlantic Basin and ultimately worldwide, as described subsequently. Early on we realized that the typical paradigm in statistics of distribution fitting to historical data sets at the analyst's leisure was untenable. Although we continue to observe re-analyses of Fisher's iris data, Brownlee's stack loss data, 1986 baseball salary data and other classical data sets (e.g., Friedman and Meulman, 2004), in the realm hurricane science we are confronted with terabytes of new data sources requiring automated fitting procedures. Not only does every year bring in new volumes of data, but even the historical data sets are not necessarily static. Hurricane Andrew in 1992 has undergone extensive re-analyses, and its maximum wind speed as given in the widely used HURDAT data base (Jarvinen, Neuman, and Davis, 1984, as updated annually by the Tropical Prediction Center, TPC) was increased by 20kts. Hence, fitting methodologies must be robust to such perturbations to the data bases. This environment has a major impact on the focus of our statistical concerns. Some retrospective statistical analyses can be necessitated on a time available basis but for the most part we are confronted with continual pressures to advise decision makers. In our applications, the criterion of goodness-of-fit tends to be trumped by the goodness of our predictions (which are indirectly supported by the distribution fits and other analyses). In spite of the occasional stresses owing to time constraints, it is nevertheless gratifying to see the fruits of these labors having tangible impacts in the field.

This paper first reviews the basics of hurricane modeling necessary to generate data sets for analyses. Note that buoy-based and ground-based anemometers (wind speed gauges) are insufficient to characterize historical events for damage forecasts and estimating insurance premiums. Wind gauges tend to fail during extreme events and their sparse placement precludes detailed analysis even if the gauges survive to record peak winds. Satellite Remote Sensing offers tremendous opportunities, but this data is available only for the last decade. Thus, wind field models are necessary to determine likely wind speeds across an entire region. With annual maximum winds available at any given site, Section 3 considers the fitting of extreme value distributions for massive numbers of individual sites. Graphical summaries and predictive performance indicate that the two parameter Weibull distribution is more than adequate for our purposes. Section 4 documents several examples of data sets used in our work where data preparation issues dominate. Data quality is critical in order to avoid optimal fits to irrelevant or tainted data. Section 5 focuses on basin wide distribution fitting performance, a critical aspect to avoid being duped on the results at one favorite site, while in section 6 we apply our analyses to the debate over the design criteria for the New Orleans flood protection system. We conclude the paper with a call for statisticians to consider operational projects that test their mettle, and for practitioners in other fields to more closely incorporate such expertise as an inherent part of all phases of the analysis.

2. Hurricane Modeling Basics

Historical wind speed observations, damage estimates and insurance loss data provide insight into the impacts of wind perils but do not provide direct mechanisms for estimating future losses or future hurricane activity in general. Watson and Johnson (2004) describe hurricane catastrophe modeling as consisting of five basic components:

Wind Field
Damage functions (vulnerability)
Insurance models
Frequency of Events
Input data bases

The hurricane field itself can be mathematically modeled as a function of such characteristics as the minimum central pressure, the far field pressure, the maximum wind velocity (at the eye wall), the radius of maximum winds, forward speed and so forth. More complex wind fields can be devised using additional input parameters. Figure 1 provides a set of four cross-section snapshots of a hurricane. When combined with forward velocity, these functions produce realistic asymmetric wind fields. The simplest function is a solid of rotation (the “Rankin Vortex”). Despite its simplicity, it performs well with some historical storms. The SLOSH wind field is used by the National Weather Service in its storm surge modeling. The AFGWC wind field is used in wave models, while the Standard Project Hurricane (MSPH) is used for a wide variety of design applications. Even models that appear to produce nearly identical winds can produce different damage levels due to the nonlinear nature of damage functions.

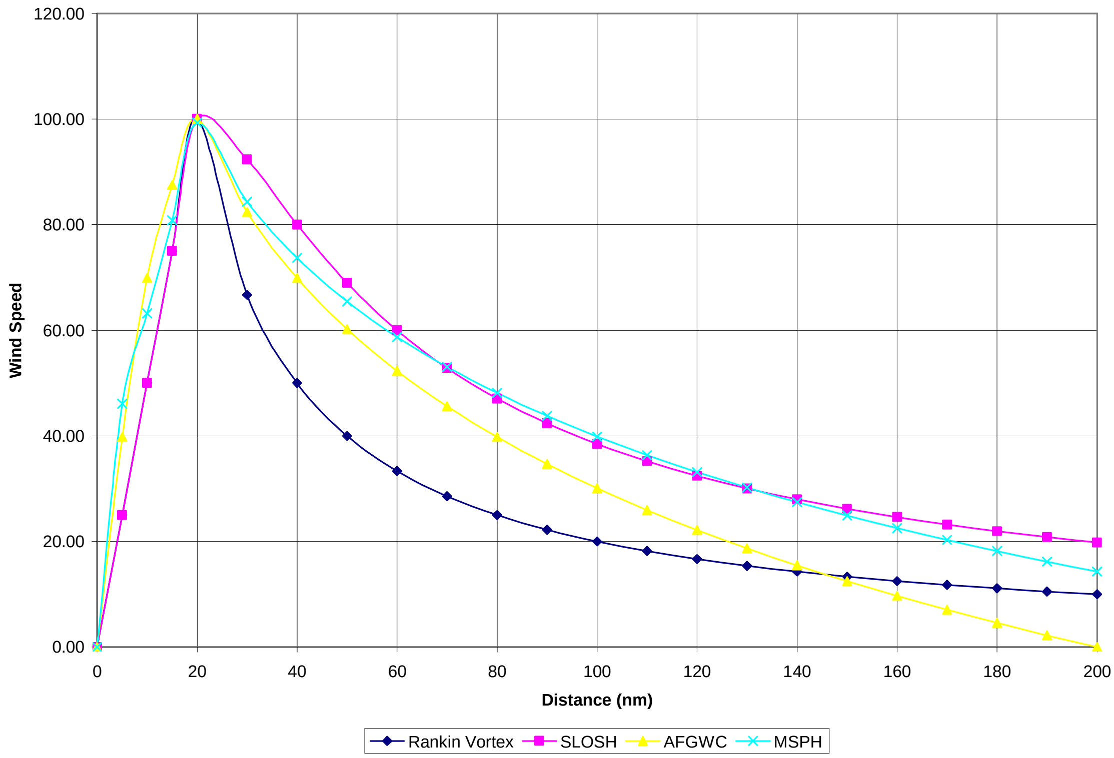

Figure 1: Representative wind profiles for parametric wind models.

Damage functions can be very sophisticated and tend to be the most proprietary aspects of commercial models. A simple but sometimes effective damage function yields no damage for wind speeds less than a nominal wind speed (for example, 30mph), complete damage for wind speeds in excess of 150 mph, and a cubic function in between. The insurance component handles deductibles and converting physical damage to dollar losses (for example, in some jurisdictions 51% damage equates to a total loss due to building code requirements). Table 1 shows the difference in computed damage using the X-Cubed damage function for the wind models used in figure 1 at a distance of 30 nautical miles. As can be seen, small differences in wind speed produce significant differences in damage.

Table 1: Wind (knots) and Damage Percentage for Wood Frame Structure at 30nm

	Wind	Damage
Rankin	66.67	22.26%
SLOSH	92.31	59.09%
AFGWC	82.39	42.02%
MSPH	84.32	45.05%

For the purposes of this paper, the damage and insurance components are not central to the discussion. Moreover, in a recently completed study, Watson and Johnson (2006) refined their earlier assessment of the relative contributions to loss costs to be dominated by meteorology. In particular, we examined nine distinct wind field models, four friction functions and nine damage functions.

Given a hurricane simulation model, the next step in loss estimation is to accumulate damage over a single hurricane (to compare modeled to actual losses) or a collection of hurricanes (to determine long term insurance rates). Incidentally, our approach has been to simulate the complete historical storm set (1346 storms from 1851 through 2005) without any smoothing of tracks or intensities. Our smoothing comes into play at individual sites by recording the annual maximum winds and then fitting an extreme value distribution to the collection of values (Section 3). This approach has been operational in Florida since its initial construction in 1999 to support local mitigation measures for the state emergency management agency. This project was at a high resolution (30 m) which led to consideration of 3 billion sites in the state. Residential loss costs are then computed by overlaying the housing tax parcel data on the basic grid.

The fifth component of hurricane catastrophe modeling involves the supporting data bases. Topography and bathymetry are required for the region of interest. Great strides have been made in this regard. Digital terrain data bases can be constructed from satellite data. The bathymetry which comes into play in wave and storm surge calculations can be gleaned adequately to depths of 50ft. Land use and land cover data are also obtained from remote sensing and then classified to determine the frictional impacts on wind.

3. Distributional Fitting at individual sites

Johnson and Watson (1999) have found that the two parameter Weibull distribution is quite suitable for fitting annual maximal winds at individual sites. This distribution has density function:

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right], \qquad x \ge 0.$$

where $\alpha > 0$ is the shape parameter and $\beta > 0$ is the scale parameter. Early on we considered other extreme value or reliability-type distributions (for example, lognormal and inverse Gaussian). The lognormal distribution provided upper tail behavior that was meteorologically unrealistic (the tails are too heavy suggesting 200mph wind speeds throughout the Atlantic over a 100 year time frame). The inverse Gaussian was slightly worse in prediction than the Weibull and was not as convenient for generating prediction limits, to be described shortly. Its prediction performance was also substandard compared to the Weibull.

Upon selecting the Weibull distribution, we recognized that we would be fitting it throughout multiple hurricane basins at an enormous number of sites. Hence, it was critical to develop an estimation strategy that was virtually foolproof. We found that an initial detailed grid of values for alpha (beta can be obtained in closed form once alpha was estimated) was much safer than using fast Newton-Raphson iterative scheme but which occasionally is an unreliable numerical method (diverges).

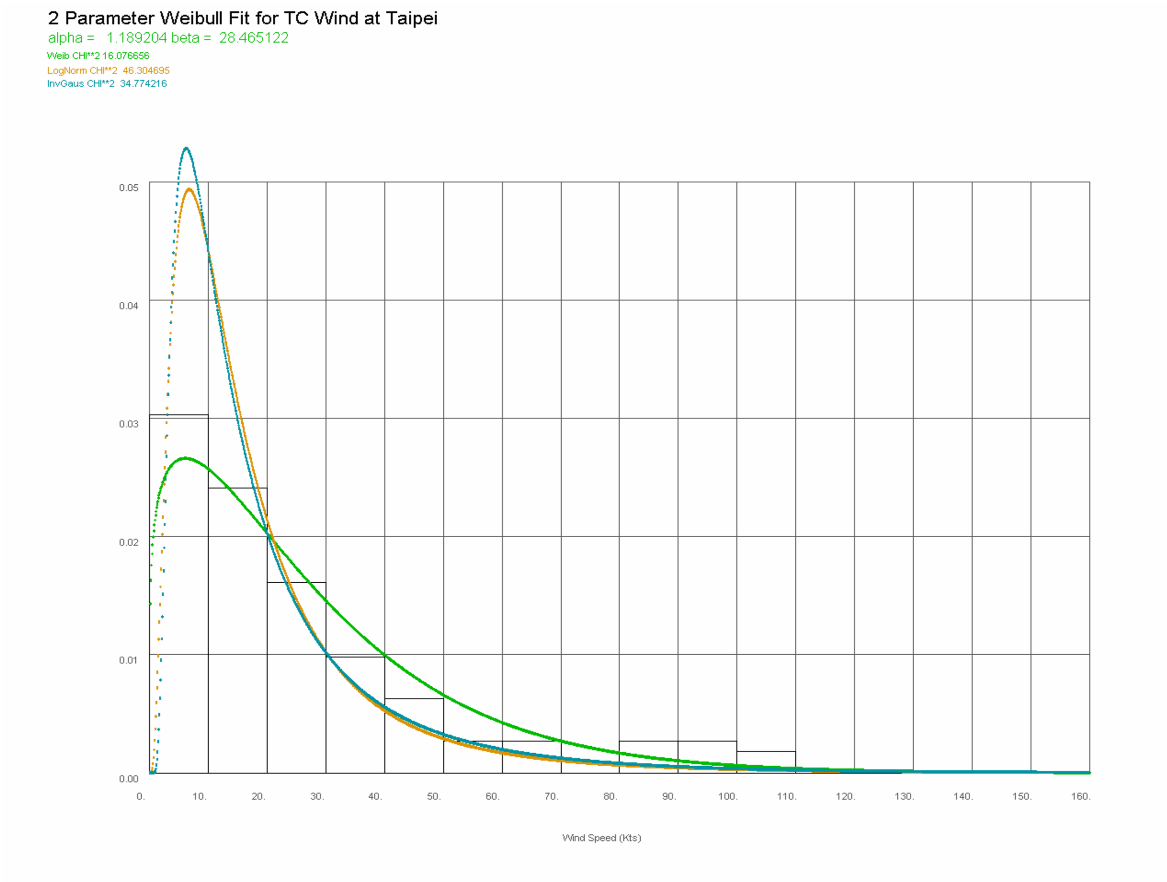

Figure 2: Weibull Fit for Tropical Cyclone Wind at Taipei, Taiwan

Figure 2 shows the distributional fits at Taipei, Taiwan. Viewed graphically, the Weibull fit is a clear winner at this site, and the χ^2 goodness of fit statistic is superior to the statistic for both the Log Normal and Inverse Gaussian fits. We caution against drawing too many conclusions for the analysis at a single site. For any given site, sometimes the Weibull would prove superior, and sometimes perhaps the Log Normal. However, when thousands of sites were considered the Weibull was superior in predictive ability, despite the Log Normal having fairly comparable χ^2 values. In fairness, the χ^2 is an omnibus test and our major concern is with the tail behavior (extreme events) for which there are not many events to assist in the distribution discrimination. We found that the Log Normal distribution was much more sensitive to the extreme event at a given location, also causing a high bias. This indicates that it is important not to lose sight of the overall objective in the quest for optimization of a narrow statistical objective.

Return periods are well understood by the emergency management and government agencies that have funded our projects. To illustrate, the hundred year wind is roughly the maximum wind that we expect to see in the next hundred years. Following Simiu and Scanlan (2003), more formally, the hundred year wind is the wind speed such that there is a 1 in 100 chance of exceeding it in the next year (the Atlantic hurricane season is officially June 1 through November 30, although hurricanes have occurred in every calendar month over the past 150-plus years). Return periods matter since structure engineers and architects typically design buildings and infrastructure against various extreme events, such as 100 year wind storms.

Upon fitting a Weibull distribution via maximum likelihood to a site's annual maximal winds, various return period winds can be readily extracted. The 90th percentile represents the 10 year event, the 96th percentile the 25 year event and the 99th percentile corresponds to the commonly used 100 year return period. The 50th percentile is our best guess of next year's expected maximal wind. These percentiles extracted from the fitted distribution do not take into account the uncertainty in parameter estimation. The percentiles offer our best estimate of the return periods but our decision maker clients wanted to build in further mechanisms for accounting for the uncertainty owing to estimation. This issue could be avoided by merely using the 155 years worth of annual maxima at each site and thus rely totally on the historical storm record. There are however some areas that have not experienced any direct hits of hurricanes yet it is meteorologically possible or even probable. Given the relatively small swath of the most intense winds (generally only 30 to 60 miles wide), of over such a short period it is possible for a site to have fallen in between major damage swaths. Some proprietary models address these spatial gaps by fitting distributions to the historical storm tracks and intensities and then simulate 100,000 years or more of events to fill out the regions. This approach is computationally onerous in order to cover those areas that have not experienced the rare category 5 storms. Our approach is to capture the parameter estimation uncertainty by sampling from the asymptotic distributions of alpha and beta (easily and adequately modeled using the Fisher Information matrix and the normal asymptotic distribution), generating 10,000 realizations of fitted Weibull distributions and then determining percentiles from these samples. For example, the 10,000 upper 90th percentiles at a given site are computed and sorted. We use the 9500th ordered value as an upper 95% prediction limit for the 10 year return period.

Table 2 : Prediction Limits for Taipei, Taiwan

YEAR	50%	75%	90%	95%	99%
10YR:	58.2	61.2	63.9	66.0	70.4
25YR:	77.1	81.6	86.7	90.5	104.5
50YR:	90.6	97.0	105.1	111.3	130.3
100YR:	103.2	112.9	124.1	133.2	157.7

Table 2 shows the prediction limits computed from the Weibull distribution shown in Figure 2. This information allows the planner to build in an appropriate level of protection against uncertainty in the parameter estimates. In our work in the Caribbean and else where, some standard guidance has emerged. For easily replaced infrastructure such as coastal recreational facilities, the 25 year 50% value is used. For single family homes, which have typical design (and mortgage!) lifetimes of 30 to 50 years, the 50 year 50% or 75% limit can be used, while the 100 year is typically used for larger multi-family use. Facilities important for economic recovery or shelters often use the 100 year 75% and 90% values. Critical lifeline infrastructure (hospitals, etc.) should use the 100 year 95% prediction limit.

4. Data Quality and Preparation Issues

The data fitting efforts described in the earlier sections are predicated upon some key data sets that continue to evolve (and hopefully improve!) over time. Of particular interest is the data set known in the field as HURDAT (Jarvinen et al, 1984). One of the author's (MEJ) initial exposure to the data base was in 1996 while on sabbatical at the National Hurricane Center in Coral Gables, Florida. The associate director at the time, Jerry Jarrell provided a copy of the data set on a 3.5" floppy which turned out to be very difficult to decipher. Eventually it was discovered that the data was stored in a Wordperfect File! Although it was helpful to have the data set, it was astonishing that this wealth of information resided in such a peculiar format. At that time, there were also found a number of obvious typos, incorrect entries, or coded entries that were inconsistent. For example, a missing central pressure value was sometimes given as 999 which also is a legitimate central pressure!

Fortunately, since 1996, the Hurricane Research Division has made great strides to revise and to clean the data base as well as to extend the time back to 1851 (Landsea et al 2004a). Certain decades of data have also been re-visited and adjustments made based on consensus of a review committee of experts. This effort is not without some controversy, as the group increased the maximum wind speed of Andrew (1992) by 20 knots while not changing the corresponding minimum central pressure value, causing considerable confusion in the catastrophe modeling community. It should also be noted that when we began our work in 1996, there were some groups that advocated the use of only hurricane data obtained after the advent of routine aircraft reconnaissance in 1948. Other applications included only data since 1900. For both of these limited data sets and especially for the 50 years set since 1948, the estimation of extreme percentiles corresponding to the 100 year return period, for example, are greatly inhibited. Our own examination of the data prior to 1948 did not demonstrate strong inter-decadal trends and we have opted to use all of the HURDAT data that is available. We were one of the first groups to advocate the 1886-present HURDAT data set, and when the 1850-present time frame became available, upon further examination again opted to use the full data set. Table 3 provides a summary on the return periods for three south Florida Cities determined from various subsets of HURDAT.

Table 3: 100 Year MLE Wind Speed (knots) from various subsets of HURDAT

	Miami	Tampa	Ft. Meyers
1851-2003	104	88	108
1871-2003	108	85	107
1886-2003	107	86	104
1900-2003	118	82	93
1950-2003	92	66	92

Not surprisingly, the full HURDAT data set provides the most stable estimates of wind return periods. In particular, the inclusion of results from 1886 down to 1851 provides tighter bands on the prediction intervals without a notable overall shift. The table also demonstrates that our values for southwest Florida tend to be larger than those observed by groups employing data for the 1900-present data. There was a cluster of intense hurricane activity in the late 19th century which meteorologists point to in the context of intra-decadal variability. With consideration of the full data set, hurricane Charley of 2004 (which struck Punta Gorda near Sarasota, Florida) does not appear quite so unusual compared to the historical record.

HURDAT has several advantages: it is widely available, is in an easily decoded format, and is routinely updated. Unfortunately, it has a number of aspects that make it problematic. On aspect of HURDAT which demonstrates fairly convincingly the lack of statistical expertise input to its development is the rounding or truncation (it is unclear which technique is used) of positions to the nearest 0.1 degree. Such precision was common in the 1960s when 16 byte integer arithmetic was more common and every digit was an effort to retain. Since 0.1 degree is about 6 nautical miles, this rounding introduces undo error and an additional source of variation. The counterpart group of HRD in the Pacific reports locations to the nearest 0.01 degree which is an improvement. The rounding phenomenon is not confined to spatial location. Maximum wind speeds in HURDAT are rounded (or truncated) to the nearest 5 knots on the pretext that wind speeds are impossible to measure very accurately. However, in the spirit of measurement systems analysis source of variation studies, rounding introduces yet another source of variation to the data set. Additionally, with damage proportional to the square or cube of wind speed, the rounding procedure can adversely propagate into the damage calculation. A further complication is that missing values are sometimes estimated from simple relationships rather than flagged as missing. Thus, central pressure may be estimated from winds and vice versa without the data point being flagged as being computed rather than observed. This precludes the use of advanced data mining techniques that could shed light on the problem areas and opportunities rather than simply obfuscating them.

Although the HURDAT data set has a relatively long record (157 years at present), it does not include vital additional variables characterizing hurricane behavior such as the radius of maximum winds, the far field pressure, and radius to far field pressure. The far field pressure is known to be influential on the sensitivity and uncertainty of catastrophe models (Iman, Johnson and Watson, 2005a, 2005b). One data set available to assess far field pressure involves comprehensive archive (CARQ) records from the Automated Tropical Cyclone Forecast (ATCF) system data base (Sampson and Schrader, 2000). The ATCF system is used for operational forecasting at the National Hurricane Center, the Joint Typhoon Warning Center, and other agencies. Our initial consideration of these records discovered anomalies with specific entries and some obvious errors. For example, there were a few examples where the radius to maximum winds was greater than the radius to far field pressure or there was a negative pressure gradient. Although this data set represents a superb set of detailed data on recent events (1990 to present), the data set is not a panacea. Our work has been completed and submitted to a meteorological journal, as the substance is of passing interest to statisticians and the statistical content is quite elementary. The implications here are that data quality is essential for providing suitable inputs to catastrophe models. The ATCF CARQ records suffer from some similar problems as HURDAT in terms of structure. Watson and Johnson (2006) note the preponderance of even pressure values (1010 and 1012 rather than 1009 and 1011 mb). This is due to the tendency of the human analysts to use even numbers (a technique taught in most university level weather analysis classes), and offers an additional source of variation to the data that can further propagate depending on the analysis. It is vital for the statistician to understand and to be involved in the underlying decision making process when data bases are designed. For another example, in the ATCF system positions in the real time position estimate file are recorded to the nearest 0.01 degree, yet as noted earlier are truncated to the nearest 0.1 degree in the "best fit" and HURDAT files. A meteorologist would say "we can't reliability measure storm positions to the nearest 0.1 degree some times, and the older data is worse than that, so why bother to record it with greater accuracy?" The statistician would advocate retaining the accuracy to avoid additional uncertainty as well as to allow the potential for improvement in measurement accuracy.

5. Predictive Ability and Performance

In light of the data quality issues outlined above, and the simplicity of the statistical approach, one might well ask how well the ultimate predictions made by the approach outlined in section 3 perform. Here we considered two issues: stability and accuracy. Stability is important as one would hope that as new data is added to the process, the original estimates would not change substantially. Imagine designing a structure for what was thought to be a 100 year event, and then discovering the next year that the design storm actually occurs every 30 years (on average). As noted earlier the HURDAT data set has undergone many changes over time. At the onset of the Montego Bay study, the data set extended from 1886 to 1996. In addition to annual updates, in 2001 the data set was extended to 1871, while the current version runs from 1851 through 2005 (Landsea et al., 2004a). Not only have additional historical years been added, changes have been made to the track and intensity of storms already in the data set, most famously the 2004 "upgrade" of Hurricane Andrew from Category 4 to Category 5 (Landsea et al., 2004b). Given the dynamic nature of the data base, techniques used to compute return periods for lifeline infrastructure must be resilient and insensitive to such changes while maintaining accuracy. Table 4 shows our original 1996 10, 25, 50 and 100 year Maximum Likelihood Estimates for Montego Bay, Jamaica, along with the estimates using the update 1871-2002 and 1851 to 2005 data sets, as well as the observed 0.9, 0.96, 0.98, and 0.99 percentile values from the full 1851-2005 data set. In addition to data base quality improvements, the 1851-2005 analysis used a multi-model ensemble approach to making the wind estimates (Watson and Johnson, 2004).

Table 4. Stability of Weibull Approach to data set evolution

Return Period	1886 to 1996	1871-2005	1851-2005	Observed
10	38.9	36.4	34.4	35.4
25	66.6	66.8	64.2	68.4
50	89.5	89.2	88.0	89.6
100	99.9	99.7	93.6	94.1

In light of the 5 knot rounding/truncation issues discussed in the previous section, the stability of the values is remarkable. The close agreement between the forecast and observed proportions is initially comforting, but an analysis for a single site is not diagnostic and could in fact be misleading: the close agreement might be evidence of over fitting.

A better performance metric is to compare the predictive ability of the technique over many sites using automated techniques. Using the prediction limit technique described in section 2 and the longer data sets created by TPC, we have an opportunity to construct a test on a basin wide basis. First we created a 6 nautical mile grid over the Atlantic Basic from 15N to 45N, and 60W through 100W (135,000 points) to span the area impacted by hurricanes.. Next we segmented the historical data set into two parts. The period from 1851 through 1955 was used to fit the Weibull distribution at each site. Fifty year wind speeds were computed at 50%, 75%, 90%, and 95% prediction limits for the 27,106 land cells in this grid that experienced at least one hurricane force wind event between 1851 and 1955. Finally, we compared the number of sites that exceeded the various prediction limits during the fifty year period 1956-2005, the second segmentation of HURDAT. Thus, the data set used for fitting and the data set used for verification were independent; it was as if we were in 1955 trying to forecast the next 50 years. The results of the analysis are presented in Table 6. The agreement is very satisfying and confirms that the use of the Weibull distribution is reasonable. Any other distributional fit deemed “superior” would be going after the difference column which is nearly negligible from a practical standpoint.

Table 6. Cross-validation study of 50 year prediction limits

Percentage of Sites within 50 Year Prediction Limit		
Predicted	Observed	Difference
50%	49.52%	0.48%
75%	74.44%	0.56%
90%	90.91%	0.91%
95%	95.95%	0.95%

There are a number of issues which are raised by Table 6. The difference column although small in practical magnitude, is “large” if viewed in the context of a binomial reference distribution. A difference of 0.0095 for the 95% prediction limit is statistically significant, if not practically significant for a large sample size of 27, 108. The binomial model is at best a rough guide, however, since it is doubtful that adjacent sites are independent. One strong but relatively rare event can impact a large number of sites.

The difference between observed and predicted evidenced in table 6 could be attributable to a number of sources. An interesting possibility is that we are detecting natural or anthropogenic climate change signals. It is well established that hurricane activity is influenced by global climate patterns such as the El Nino/Southern Oscillation (ENSO) (Watson, 2002b). The ENSO is classified into three modes: warm phase (El Nino), normal, or cold phase (La Nina) according to temperatures in the eastern South Pacific. Historically, normal years constitute about 40% of years, while the remaining 60% are evenly split between El Nino and La Nina years. By segmenting the overall (1851-2005) data set according to the state of the ENSO, we can assess the impact of ENSO on hurricane frequencies on a site by site basis. This site by site aspect is important as we have found that while a warm phase ENSO tends to suppress hurricane formation basin wide, it actually increases the probability for hurricane force winds at some sites. In other words, even though there are fewer storms, they tend to traverse different areas than in other years. Table 7 shows the impact of ENSO on Montego Bay, Jamaica. Montego Bay is an astonishing 7.5 times more likely to experience a “100 year event” during a cold phase La Nina year than the overall historical average (computed as 1.0). In contrast, during a warm phase of El Nino year, Montego Bay is only 1/5 as likely to experience a “100 year event”. Compare the Montego Bay results with the results for Miami, Florida, where the risk is more consistent, but biased somewhat against cold phase (La Nina) years. Since it appears that 2006 will be a cold phase year, Jamaica should take special care in their hurricane preparations this year.

Table 7: Excerpt of output of the TAOS Statistical Analysis Program for Montego Bay

```
100.year 50% value for all years:   82.7 kts
  probability of exceeding in normal year: 0.00638   0.6
  probability of exceeding in elnino year: 0.00165   0.2
  probability of exceeding in lanina year: 0.07516   7.5
```

Table 8: Excerpt of output of the TSAP program for Miami, Florida

```
100.year 50% value for all years:  104.1 kts
  probability of exceeding in normal year: 0.01222   1.2
  probability of exceeding in elnino year: 0.01154   1.2
  probability of exceeding in lanina year: 0.00899   0.9
```

We first applied this type of analysis to the storage of an oil platform in transit through the Caribbean during the 1998 hurricane season. The client, being self insured for the full value of the billion dollar asset, obviously wanted to find a safe harbor to sit out the season. (Incidentally, for the site we recommended, the platform did not experience tropical storm winds or higher.) Starting in 2005, our research team has begun applying this approach to forecast annual probabilities of hurricane force winds on a city by city basis. We are conducting further research in this area with respect to other climate signals such as the Multi-Decadal Oscillation, the North Atlantic Oscillation, as well as for potential anthropogenic climate change signals. The human climate change debate aside, understanding that annual probabilities are not identical from year to year but are predictable given various natural climate signals (available in the Northern Hemisphere spring) allows such variability to be incorporated in the short term planning process.

Given the demonstrated predictive ability from the basin wide tests, we are reasonably confident that our return period predictions are valid. Next we apply the analysis to the recent devastation of the city of New Orleans.

6. Case Study: the New Orleans levees: Wrath of Mother Nature or Ordinary Extreme Event?

The collapse of the New Orleans levees drew considerable attention to return periods through the simple concept of Saffir Simpson categories, given that higher category storms occur less often. Initially, the Army Corps of Engineers declared that the levees were "overtopped" since they were designed for a Category 3 storm, but Katrina was a Category 5 event. Subsequently, with the realization that Katrina was not a category 5 at landfall the defense was modified to "not a category 4." As it became clear that in fact the levees collapsed due to internal failures, the story changed again (Marshall, 2005, Warrick and Hsu, 2005). In terms of the rebuilding of New Orleans, decision makers should consider a return period concept on a phenomena by phenomena basis to assess the likelihood of a recurrence rather than the overly simplistic Saffir/Simpson categories, because damaging phenomena produced by a hurricane (wind, waves, and storm surges) interact in highly nonlinear and spatially diverse ways.

How strong an event was Katrina, and how often should the New Orleans area experience impacts similar to those experienced in late August 2005? The answer to this question will likely receive continued scrutiny, especially if the answer and the outcome become even more politicized. Hurricanes are classified according to peak wind speed but as is readily seen from the wind profiles shown in Figure 1 the phenomenon is not one-dimensional. The damaging wind swath for most storms is actually quite narrow, especially when compared to the overall size of the storm. Figure 3 shows Hurricane Katrina the day before landfall, with the approximate region of damaging winds noted. A satellite depiction of a storm can be highly misleading as the outer cloud shield extends two to three times the radius of the damaging winds. Even within the damaging wind zone winds are highly variable. While the peak winds at landfall in Mississippi were in the 120 to 130 mile per hour range, winds in New Orleans were far less. The wind speed at NASA's Michoud assembly plant, located east of the city and thus closer to the storm track, was only 96 mph. The NOAA buoy over Lake Pontchartain, north of the city, recorded a peak wind of 78mph. (Knabb et al, 2005) A detailed inspection of damage as seen on high resolution aerial photography conducted by one of the authors (Watson) showed minimal direct wind damage to most single family structures, indicating winds in the 70 to 80mph regime. Most damage was due to either flooding or falling trees. The dramatic destruction to mid and high rises in the downtown area, as well as damage to the Super Dome, appeared to be due to a combination of debris and poor construction.

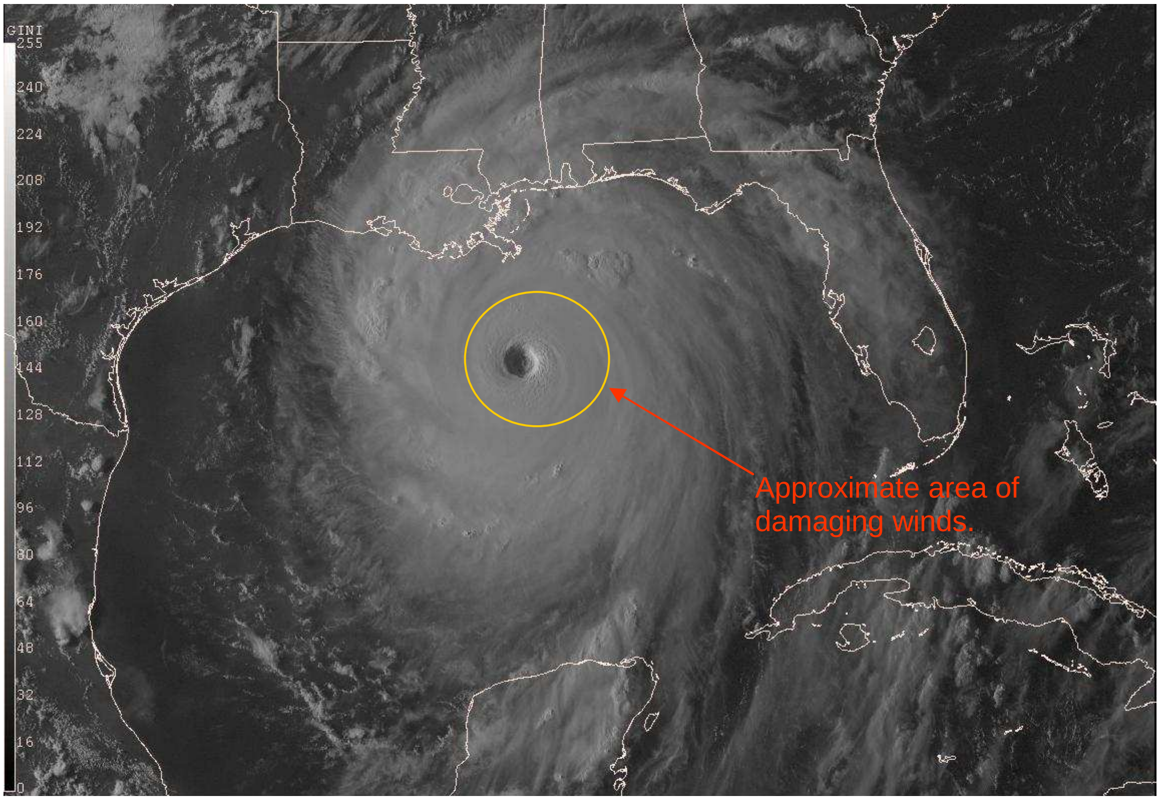

Figure 3, Hurricane Katrina, 28 August 2005, 2240UT

The question of storm surge and the failure of the New Orleans flood protection system is complex from a return period standpoint. Engineering studies are currently underway, and there are both political and scientific disputes in progress over both design criteria and quality of construction (van Heerden, 2005). Our analyses indicate that for the New Orleans area, Katrina's impacts were well within what should have been anticipated. The peak 80mph winds experienced in New Orleans have a return period of approximately 85 years, using the methodology described earlier. As reported in an investigation by PBS's NOVA program, the levee system failed at a peak surge of 10.5 feet, below the rated value of 14 feet msl (although the ongoing subsidence of the New Orleans Area reduced the margin by perhaps as much as 2 feet). Using our methodology, a 10.5 ft surge at the 17th Street Canal is a 79 year event. It should be noted that the Corps did not use return periods in their design, instead designing to a category 3 storm (111 to 130mph winds, surges of 9 to 12 feet). New Orleans presents a unique challenge in that because of Lake Pontchartain, wind and surge return periods are more disconnected than is evident in other areas. The 100 year wind for the city is approximately 115mph, a category 3 storm, but the 100 year storm surge is nearly 20 feet (19.6 ft). Thus, leaving aside issues of construction quality, not taking into account the complexity of the area and assuming storm surge return periods tracked wind return periods may have contributed to an under design of the flood protection system.

7. Conclusions

The science of statistics has direct implications for both lives and property. Data used in these analyses occasionally suffer from the oft quoted computer science adage, “Garbage In, Garbage Out”. This is nowhere truer than in the field of tropical cyclone planning and response, where decisions involving trillions of dollars of property and millions of lives depend on statistical projections derived from dynamic data bases of often suspect accuracy. We have argued indirectly in this paper that the choice of data to be fit is a critical and possibly overlooked concern. Moreover, even with a relevant data set being identified, further data preparation and quality control are necessary to avoid radically different results caused by subsequent re-fitting of the evolving data sets. Techniques that may work well for the leisurely analysis of data sets that have been carefully screened and vetted may not be appropriate. Thus, statisticians must avoid overly theoretical approaches and concentrate on solving the problem of interest to the client, even if this means not using the most glamorous approach. Our primary example of distribution fits involved the two-parameter Weibull distribution, which was shown to perform admirably for the purpose of predicting return periods for perils both at the site and basin level. Finally, convincing specialists in other fields to include statistical considerations early in the process is essential. Their involvement in the analysis process from data collection to presentation can result in an improved product for critical decision making purposes.

References

Coastal Engineering Research Center, 1984: *Shore Protection Manual*, **1**, 3-81, US Army Corps of Engineers, Washington, D.C.

Dunion, J., C. Landsea, S. Houston, and M. Powell, 2003: A Reanalysis of the Surface Winds for Hurricane Donna of 1960, *Monthly Weather Review*, **131**, No. 9, 1992–2011.

Florida Commission on Hurricane Loss Prediction Methodology, 2005: *Report of Activities as of November 1, 2005*, FL State Board of Administration, Tallahassee, FL.

Friedman, J. H. and J. Meulman, J., 2004: Clustering objects on subsets of attributes (with discussion), *JRSS, Series B*, **66**, 815-849.

Hansen, M., R. DeFries, J.R.G. Townshend, and R. Sohlberg, 2000: Global land cover classification at 1km resolution using a decision tree classified, International *Journal of Remote Sensing*, **21**, 1331-1365.

Holland, G., 1980: An analytic model of the wind and pressure profiles in hurricanes, *Mon. Wea. Rev.*, **108**, 1212-1218.

Iman, R. L., M. E. Johnson and C. Watson, Jr., 2005a: Sensitivity Analysis for Computer Model Projections of Hurricane Losses, *Risk Analysis*, **25**, 1277-1297.

Iman, R. L., M. E. Johnson and C. Watson, Jr., 2005b: Uncertainty Analysis for Computer Model Projections of Hurricane Losses, *Risk Analysis*, **25**, 1299-1312.

Iman, R. L., M. E. Johnson and C. Watson, Jr., 2006: Statistical Aspects of Forecasting and Planning for Hurricanes, *The American Statistician*, **60**, 105-121.

Jarvinen, B. R., C. Newman, and M. Davis, 1984: A tropical cyclone data tape for the North Atlantic Basin, 1886-1983: Contents, limitations, and uses. NOAA Technical Memorandum NWS NHC 22, Coral Gables, FL, 21 pp.

Jelesnianski, C., J. Chen, and W. Shaffer, 1992: *SLOSH: Sea, Lake, and Overland Surges from Hurricane Phenomena*, NOAA Technical Report NWS 48, Silver Spring, MD: National Weather Service.

Johnson, M. E., 1997: Caribbean Storm Surge Return Periods, Organization of American States Caribbean Disaster Mitigation Project Workshop, Kingston, Jamaica, October 31, 1997.

Johnson, M. E. and C. Watson, Jr., 1999: Hurricane Return Period Estimation, *10th Symposium on Global Change Studies*, Dallas, TX, 478-479.

Landsea, C., C. Anderson, N. Charles, G. Clark, J. Dunion, J. Partagas, P., Hungerford, C. Neumann, and M. Zimmer, 2004a: The Atlantic hurricane database re-analysis project: Documentation for the 1851-1910 alterations and additions to the HURDAT database. *Hurricanes and Typhoons: Past, Present and Future*, R. J. Murnane and K.-B. Liu, Editors, Columbia University press (in press).

Landsea, C. , J. Franklin, C. McAdie, J. Beven II, J. Gross, R. Pasch, E. Rappaport, J. Dunion, and P. Dodge, 2004b: A re-analysis of Hurricane Andrew's (1992) intensity. *Bulletin of the American Meteorological Society,* **85**, 1699-1712.

Marshall, Bob, 2005: "17th Street Canal levee was doomed; Report blames corps", The Times-Picayune (New Orleans), November 30, 2005.

Miller, B., 1962: Characteristics of hurricanes, *Science*, **157**, 1389-1399.

Sampson, Charles R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (Version 3.2), *Bulletin of the American Meteorological Society*, **81**, 1231-1240.

Schwerdt, R., F. Ho, and R. Watkins, 1979: *Meteorological Criteria for Standard Project Hurricane and Probable Maximum Hurricane Wind Fields, Gulf and East Coasts of the United States,* NOAA Technical Report NWS 23. Silver Spring, MD: National Weather Service.

Richard D. Knabb, Jamie R. Rhome, and Daniel P. Brown, 2005: *Tropical Cyclone Report: Hurricane Katrina, 23-30 August 2005,* National Hurricane Center, 20 December 2005. 42pp.

van Heerdan, Ivor, 2005: “Failure of Levee Systems Surrounding Greater New Orleans During Hurricane Katrina, Preliminary Assessment”, testimony before the US Senate Committee on Homeland Security and Governmental Affairs, November 2nd, 2005.

Watson, C., Jr., 1995: The Arbiter Of Storms: a high resolution, GIS based storm hazard model, *Natl. Wea. Dig.*, **20**, 2-9.

Watson, C., Jr., 2002a: Using integrated multi-hazard numerical models in coastal storm hazard planning, *Solutions for Coastal Disasters ‘02 Conference Proceedings,* American Society of Civil Engineers, Reston, VA.

Watson, C., Jr., 2002b: Implications of climate change for modeling coastal hazards, *Solutions for Coastal Disaster ‘02 Conference Proceedings,* American Society of Civil Engineers, Reston, VA.

Watson, C. Jr, and M. Johnson: 2003: *An assessment of computer based estimates of hurricane loss costs in North Carolina*. Report for the North Carolina Department of Insurance, Raleigh, N.C.

Watson, C. Jr, and M. Johnson: 2004: Hurricane Loss Estimation Models, Opportunities for Improving the State of the Art, *Bulletin of the American Meteorological Society,* **85**, 1713-1726.

Watson, C. Jr, M. Johnson, and M. Simons, 2004: Insurance Rate Filings and Hurricane Loss Estimation Models, *J. Insurance Research*, **22** (3), 39-64.

Warrick, Joby, and Spenser Hsu, 2005: “Levee’s Construction Faulted in New Orleans Flood Inquiry”, *The Washington Post,* November 3, 2005, Page A03.