Comparison of Hurricane Return Levels Using Historical and Geological Records

JAMES B. ELSNER AND THOMAS H. JAGGER

Department of Geography, The Florida State University, Tallahassee, Florida

KAM-BIU LIU

Department of Oceanography and Coastal Sciences, Louisiana State University, Baton Rouge, Louisiana

(Manuscript received 22 January 2007, in final form 18 May 2007)

ABSTRACT

Hurricane return levels estimated using historical and geological information are quantitatively compared for Lake Shelby, Alabama. The minimum return level of overwash events recorded in sediment cores is estimated using a modern analog (Hurricane Ivan of 2004) to be 54 m s\(^{-1}\) (105 kt) for a return period of 318 yr based on 11 events over 3500 yr. The expected return level of rare hurricanes in the observed records (1851–2005) at this location and for this return period is estimated using a parametric statistical model and a maximum likelihood procedure to be 73 m s\(^{-1}\) (141 kt), with a lower bound on the 95% confidence interval of 64 m s\(^{-1}\) (124 kt). Results are not significantly different if data are taken from the shorter 1880–2005 period. Thus, the estimated sensitivity of Lake Shelby to overwash events is consistent with the historical record given the model. In fact, assuming the past is similar to the present, the sensitivity of the site to overwash events as estimated from the model is likely more accurately set at 64 m s\(^{-1}\).

1. Introduction

Paleotempestology (the study of storms from geological evidence) offers a glimpse at tropical cyclone activity through the ages (Liu 2004; Donnelly and Webb 2004). Coastal wetlands and lakes are episodically subjected to overwash processes during catastrophic hurricane strikes when barrier sand dunes are overtopped by storm surge. The frequency of overwash sand layers in lake and wetland cores provides an estimate of their return period. Records of hurricanes since 1851 are available for studying the historical hurricane climatology, although the records since about 1880 are most reliable. A statistical model of the wind speed maximum from the strongest historical hurricanes provides a way to estimate the expected return level for a given return period. Here we compare results from the two methods for Lake Shelby, Alabama, located along the northern Gulf Coast of the United States.

We begin by describing the geological evidence for hurricane activity using sediment cores taken from the lake. We then describe and model the local record of hurricane activity using historical hurricane data. Because the historical record is limited to the past 155 yr, the model is unstable (parameter estimates vary widely) for storms within the small region defined by storms capable of producing an overwash into Lake Shelby. Therefore, we predict the model parameters at the smaller radial distance of interest using the parameter values estimated from larger radial distances. This procedure provides an estimated return level for a return period corresponding to the frequency of sediment layers in the cores. Standard errors are used to construct a confidence interval on the estimate. The minimum return level from overwash deposits is below (10 m s\(^{-1}\), 16%) the lower bound on the 95% confidence interval of the return level estimated from the historical records.

2. Hurricane-deposited sand layers

Over the past decade or so, several dozen coastal lakes and marshes along the Gulf and Atlantic Coasts have been cored with the hope of better understanding prehistoric hurricane activity. One of the earliest cored sites is Lake Shelby, Alabama (30°15'40"N, 86°36' W).
T bent adjacent $^{14}$C dates. Due to secular changes in ages of the rest are a result of linear interpolation between calendar date (yr BP). By convention, “present” is taken as AD 1950.

We thus arrive at a minimum return level of 54 m s$^{-1}$ for a return period of 318 yr. That is, we expect, on average, a hurricane of at least this intensity to directly strike Lake Shelby once every 318 yr. It is noted that this is a minimum return-level estimate because no historical hurricanes are known to have left a sand layer at this site. In the work that follows we compare this geologically based return level with a statistically based return level using a parametric model and the catalog of known storms since 1851.

### 3. Historical hurricanes

It is assumed that a hurricane of sufficient intensity passing within 45 km of Lake Shelby—a direct hit—will deposit a sand layer at the coring site. For locations on the left-hand side of a hurricane’s track (looking in the direction of motion) a direct hit occurs if the storm passes within a distance equal to the storm’s radius of maximum wind. For locations on the right-hand side, a direct hit occurs when the storm passes to within a distance equal to twice the radius of the maximum wind. On the basis of 59 hurricanes from 1893 through 1979 affecting the U.S. coastline, the mean radius of maximum winds was found to be 47 km (Hsu and Yan 1998). On average, the radius is somewhat smaller for the strongest hurricanes. To the extent that the hurricanes responsible for the sand layers had systematically smaller radii, the choice of 45 km could lead to some error, as addressed later.

The historical hurricane record spans the last century and a half, but it contains no storm of category 4 or higher intensity passing within this limited distance of the lake. It is therefore necessary to consider the frequency of storms at larger distances. Figure 1 shows the points of maximum intensity for all storms passing within 495 km of Lake Shelby since 1851. The choice of 495 km here is for illustration. The track data are derived from the Hurricane Database (HURDAT; or best track) maintained by the National Hurricane Center (NHC), which consists of the 6-hourly position and intensity estimates of tropical cyclones back to 1851 (Jarvinen et al. 1984). We use spline interpolation (see Jagger and Elsner 2006) to obtain positions and wind speeds at 1-h intervals.

The number of hurricanes, by category, passing within incrementally larger radial distances of the coring site is tabulated (Table 2). The ratio of the number of category 4 and 5 hurricanes to the number of category 0 storms is zero for the shortest distance, but tends to stabilize at around 10%–11% for distances of 270 km or greater. The limited record length implies that return-level estimates for storms passing through a small area will be misleading. Here we demonstrate a

#### Table 1. Paleohurricanes in the sediment cores from Lake Shelby taken in 1990 are shown. “Number” is the overwash sequence event starting from the earliest, “Date” is the radiocarbon ($^{14}$C) date [yr before present (BP)], “Error margin” is the 2 std dev estimate of the $^{14}$C dating, and “Calendar” is the intercept calendar date (yr BP). By convention, “present” is taken as AD 1950.

<table>
<thead>
<tr>
<th>Number</th>
<th>Date (yr BP)</th>
<th>Error margin (yr)</th>
<th>Calendar (yr BP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>770</td>
<td>70</td>
<td>679</td>
</tr>
<tr>
<td>2</td>
<td>1360</td>
<td>80</td>
<td>1288</td>
</tr>
<tr>
<td>3</td>
<td>2190</td>
<td>80</td>
<td>2176</td>
</tr>
<tr>
<td>4</td>
<td>2240</td>
<td>80</td>
<td>2222</td>
</tr>
<tr>
<td>5</td>
<td>2450</td>
<td>80</td>
<td>2481</td>
</tr>
<tr>
<td>6</td>
<td>2650</td>
<td>80</td>
<td>2755</td>
</tr>
<tr>
<td>7</td>
<td>2960</td>
<td>80</td>
<td>3126</td>
</tr>
<tr>
<td>8</td>
<td>3000</td>
<td>90</td>
<td>3179</td>
</tr>
<tr>
<td>9</td>
<td>3110</td>
<td>60</td>
<td>3350</td>
</tr>
<tr>
<td>10</td>
<td>3160</td>
<td>60</td>
<td>3378</td>
</tr>
<tr>
<td>11</td>
<td>3240</td>
<td>80</td>
<td>3466</td>
</tr>
</tbody>
</table>
method that, to some extent, gets around this limitation by borrowing information about the frequency of storms in surrounding areas.

4. Return-period model

There are various empirical and semiparametric approaches for estimating hurricane wind speeds along the coast (Murnane et al. 2000; Powell et al. 2005; Vickery et al. 2006). While useful for estimating the annual probability, they are less so for estimating return levels for return periods that exceed the data record (e.g., 318 yr). Instead, here we employ a fully parametric approach that was developed in Jagger and Elsner (2006). Under the assumption that hurricane wind speeds follow an extreme value distribution and that the data are reliable for estimating the model parameters, the approach is capable of producing useful estimates of return levels for return periods that are much longer than the data record.

The distribution of the maximum wind above a threshold value \( u \) is assumed to follow a generalized Pareto distribution (GPD). The likelihood function is the product of the generalized Pareto probabilities for each wind speed estimate. Following Jagger and Elsner (2006), we model the exceedances \( W - u \) as samples from a family of GPD, so that for an individual hurricane with maximum wind \( W \),

\[
\Pr(W > u | W > u) = \left[1 + \frac{\xi}{\sigma} (v - u)\right]^{-1/\xi}
\]

where \( \sigma > 0 \) and \( \sigma + \xi(v - u) \geq 0 \). For negative values of the shape parameter (\( \xi \)), the GPD family of distributions has an upper limit of \( W_{\text{max}} = u + \sigma/|\xi| \).

The frequency of storms with an intensity of at least \( u \) follows a Poisson distribution with a rate \( \lambda_u \), the threshold-crossing rate. Thus, the number of hurricanes per year with winds exceeding \( v \) is a thinned Poisson process with mean \( \lambda = \lambda_u \Pr(W > v | W > u) \). This is called the peaks-over-threshold (POT) method, and the resulting model is completely characterized for a given threshold \( u \) by \( \sigma, \xi, \) and \( \lambda_u \) (the GPD parameters and the threshold-crossing rate, respectively).

Because the number of storms exceeding any wind speed \( v \) is a Poisson process, the return period for any \( v \) has an exponential distribution, with mean \( r(v) = 1/\lambda_v \). By substituting for \( \lambda_v \) in terms of both \( \lambda_u \) and the GPD parameters, and then solving for \( v \) as a function of \( r \), we can find the corresponding return level for a given return period as

\[
r_l(r) = u + \frac{\sigma}{\xi} \left[ (r \lambda_u)^{\xi} - 1 \right].
\]

Table 2. The number of tropical storms and hurricanes passing within radial distances of the Lake Shelby coring site. Column headings are distances (km). Category (Cat) 0 refers to storms with maximum sustained winds of 17–32 m s\(^{-1}\); Cat 1, 33–42 m s\(^{-1}\); Cat 2, 43–49 m s\(^{-1}\); Cat 3, 50–58 m s\(^{-1}\); Cat 4, 59–69 m s\(^{-1}\); and Cat 5 > 69 m s\(^{-1}\).
Our model estimates 56 m s\(^{-1}\) based on data over the period of 1886–1987. The threshold value (dotted line) is set at 33 m s\(^{-1}\) within 139 km of San Juan, Puerto Rico (18.2°N, 66.1°W), based on data over the period of 1886–1987. The points are the empirically estimated return levels for each hurricane occurring within the region.

We estimate statistical models for increasing distances between 45 and 585 km at increments of 45 km. The POT parameter values are plotted as a function of radial distance from the site in Fig. 3. We predict the values of \(\sigma\), \(\xi\), and \(\lambda\) at a radial distance of 45 km using separate bivariate linear regressions. For \(\sigma\) and \(\xi\), only distances greater than 180 km are used in the regression models. Return-level curves using the extrapolated POT parameters are compared with the raw POT parameters in Fig. 4. The extrapolation of the parameters produces return-level models for shorter distances that are consistent with return-level models for longer distances. The extrapolated values for \(\sigma\), \(\xi\), and \(\lambda\) at a radial distance of 45 km are 17.6 m s\(^{-1}\), -0.240, and 0.080 storms per year, respectively.

These results are based on using hurricanes in HURDAT over the period of 1851–2005. Because the near-coastal hurricane records are most reliable after about 1880 (Landsea et al. 2004), we rerun the analysis using only hurricanes in the period of 1880–2005, and find that the extrapolated values for \(\sigma\), \(\xi\), and \(\lambda\) at a radial distance of 45 km are 18.1 m s\(^{-1}\) (2.8% increase), -0.230 (4.1% decrease), and 0.077 storms per year (3.8% decrease), respectively. This results in a 2% difference in return-level estimates, as explained next.

5. Comparisons

Under the assumption of uniformity of geological and meteorological conditions across the Gulf Coast since the late Holocene, a quantitative comparison can now be made between the return level obtained using the modern analog approach and a return level from a
statistical model estimated from historical data. From Eq. (2) the extrapolated parameter values for a radial distance of 45 km provide an expected return level of 72.7 m s$^{-1}$ for a return period of 318 yr. This value is 74.1 m s$^{-1}$ (2% higher) using the parameters estimated from the shorter 1880–2005 period. In either case the estimated return-level intensity is above the minimum threshold (69 m s$^{-1}$) for a category 5 hurricane, suggesting that the sediment cores found in Lake Shelby are indeed from catastrophic events, and are possibly stronger than the minimum value estimated using the analog approach.

A confidence interval (CI) about the mean return level can be obtained by scaling the CI obtained from the model using data from a larger radial distance (e.g., 270 km). Because the return period from the core site is 318 yr, we divide 318 by 6 (270/45) to get 53 yr. The CI on the return levels for a return period of 53 yr using a distance of 270 km serves as an approximate CI on the return level for the return level using a radial distance of 45 km. In this way we obtain a 95% CI of between 63.8 and 77.5 m s$^{-1}$ for the return level of a 318-yr storm event affecting Lake Shelby. This CI includes storms of category 4 intensity. We note that using a smaller radius (22.5 km, because storm surge tends to occur only on the right-hand side of the landfall location) results in a return level of 66.5 m s$^{-1}$, which, as expected, is less than the return level computed using the larger radius. This value is within the 95% confidence interval based on the 45-km radius.

Another source of uncertainty is the return-period estimate of 318 yr from the core site, which is based on 11 being the best estimate for the number of overwash events affecting Lake Shelby during the past 3500 yr. Considering the uncertainty on that estimate to be in the range of 10–14 events from different coring interpretations, the return period ranges from 350 to 250 yr. This sets an uncertainty bound on the expected return level from the model of between 70.7 and 73.5 m s$^{-1}$.

Other sources of uncertainty related to using a single
coring site and changes in sedimentation rates over time would widen that uncertainty bound.

It is noted that Hurricane Ivan struck as a strong category 3 event within 45 km of Lake Shelby in 2004. Sediment cores taken from the lake following this event contain no evidence of overwash. However, in nearby Little Lake, adjacent to Lake Shelby but smaller in size and closer to the sand dunes, there is a distinct sand layer from Ivan. In fact, short cores taken from Little Lake indicate seven events over the past 1200 yr (Liu et al. 2003). Using the Lake Shelby model, we find an expected return level of 67 m s\(^{-1}\) with a 95% CI of between 59.4 and 71.9 m s\(^{-1}\).

6. Limitations

The analog method applied to the geological record of overwash events produces a return-level sensitivity that is consistent with the historical record given the extreme value model. In fact, the historical record suggests that a hurricane of at least 64 m s\(^{-1}\) is needed to produce a sand layer in Lake Shelby. Ivan was 54 m s\(^{-1}\) at landfall, and therefore would not have been expected to produce an overwash deposit. Given the model and the historical data, statistical uncertainty bounds can be placed on the estimated return level for a specified return period, more work is needed to account for the uncertainty in specifying this return period.

Because the analog method is based on a single modern event, the uncertainty about site sensitivity is large. It is likely that, on average, the sand layers in the core were caused by events that were stronger than 54 m s\(^{-1}\). On the other hand, there are known intensity biases in the best-track data for storms prior to aircraft reconnaissance (Jarvinen et al. 1984) and longer sampling intervals that effectively smooth the storm’s peak intensity. Thus, the return level estimated from the model using the historical record might be somewhat underestimated.

The comparison between historical and geological records of past extreme hurricanes depends on the assumption of uniformity. Comparisons are less meaningful under conditions in which the sensitivity of the site to overwash events changes over time (Donnelly et al. 2004; Liu 2004). Moreover, the comparisons are void under conditions in which the hurricane frequency is dramatically different today than it was during the middle or late Holocene. From sampling theory of extreme events, if there was significant variation in either the hurricane climatology or the sensitivity of the site to overwash events over the ages, the time between successive overwash events would not follow an exponential distribution because of this extra variation. The Kolmogorov–Smirnov goodness-of-fit test is applied to the interarrival times of the events in Lake Shelby. A \(p\) value of 0.978 on the test statistic indicates insufficient evidence to reject the null hypothesis of an exponential distribution.

We note that there is a tendency for hurricanes to weaken prior to landfall along the Gulf Coast (Knabb et al. 2005). An analog method that relies on the storm intensity at the point of landfall may underestimate the surge event. Hurricane Katrina in 2005 is an example. While Katrina made landfall near Buras, Louisiana, with maximum winds of 57 m s\(^{-1}\), winds a day earlier peaked at 77 m s\(^{-1}\) and the catastrophic storm surge was indicative of Katrina’s large size and perhaps earlier intensity. At landfall, Katrina had hurricane-force winds that extended a radial distance of 138 km from the center of circulation. Because the statistical return-period model uses wind values within a radial distance of the coring site, most of the maximum winds are higher than the landfall wind (see Fig. 1).

7. Summary

By counting hurricane-induced sand layers in coastal sediment cores, paleotempestology can provide an estimate of the frequency of extreme events in the past, and thus their return periods. But what are the intensities of these prehistoric hurricanes? The geological record alone provides no direct answer to this question, which is vital to an accurate assessment of the hurricane risk. Under certain circumstances numerical surge models can be used in conjunction with geomorphic evidence of past inundation to help answer the intensity question (Nott 2003). Instead, here we develop a statistical procedure for calibrating the geological record to the historical record at any location where historical records are available.

The method is based on “borrowing information.” Extreme hurricane winds are estimated locally by parametrically modeling the historical hurricane data (at all intensities) taken from larger regions. Limiting values for the model parameters at the location of interest are then available through interpolation. This procedure confirms that Lake Shelby is not likely to get an overwash sand deposit unless storm intensity at landfall exceeds 54 m s\(^{-1}\), as inferred from Hurricane Ivan in 2004 (under the assumptions of uniformity over time and a reliable geological record), and suggests that the sensitivity is probably closer to 73 m s\(^{-1}\).

The methodology might be most useful in the cases where no modern analog is available and where the interarrival times of overwash events do not show extra variation. The statistical model can be improved by in-
cluding estimates of uncertainty on the older historical records. This can be done by using a Bayesian approach as illustrated in Jagger and Elsner (2006). In fact, Bayesian techniques are a natural way to incorporate geological information into a parametric model to refine site-specific return-period estimates. Moreover, because the present method eliminates the need for a catalog of synthetic storms, it represents an advance that could change the way the next generation of hurricane risk models is built.

Acknowledgments. Support for this study was provided by the National Science Foundation (ATM-0435628 and BCS-0213980). Critical comments by the reviewers on earlier drafts are gratefully acknowledged. The views expressed within are those of the authors and do not reflect those of the funding agency.

REFERENCES