Tropical Cyclone Intensity Change before U.S. Gulf Coast Landfall

EDWARD N. RAPPAPORT AND JAMES L. FRANKLIN

NOAA/National Weather Service/National Hurricane Center, Miami, Florida

ANDREA B. SCHUMACHER

Colorado State University/CIRA, Fort Collins, Colorado

MARK DEMARIA

NOAA/NESDIS/Center for Satellite Applications and Research, Fort Collins, Colorado

LYNN K. SHAY

RSMAS/MPO, University of Miami, Miami, Florida

ETHAN J. GIBNEY

I. M. Systems Group, NOAA/National Climatic Data Center, Asheville, North Carolina

(Manuscript received 21 October 2009, in final form 28 May 2010)

ABSTRACT

Tropical cyclone intensity change remains a forecasting challenge with important implications for such vulnerable areas as the U.S. coast along the Gulf of Mexico. Analysis of 1979–2008 Gulf tropical cyclones during their final two days before U.S. landfall identifies patterns of behavior that are of interest to operational forecasters and researchers. Tropical storms and depressions strengthened on average by about 7 kt for every 12 h over the Gulf, except for little change during their final 12 h before landfall. Hurricanes underwent a different systematic evolution. In the net, category 1–2 hurricanes strengthened, while category 3–5 hurricanes weakened such that tropical cyclones approach the threshold of major hurricane status by U.S. landfall. This behavior can be partially explained by consideration of the maximum potential intensity modified by the environmental vertical wind shear and hurricane-induced sea surface temperature reduction near the storm center associated with relatively low oceanic heat content levels. Linear least squares regression equations based on initial intensity and time to landfall explain at least half the variance of the hurricane intensity change. Applied retrospectively, these simple equations yield relatively small forecast errors and biases for hurricanes. Characteristics of most of the significant outliers are explained and found to be identifiable a priori for hurricanes, suggesting that forecasters can adjust their forecast procedures accordingly.

1. Introduction

Tropical cyclones[1] can devastate the U.S. Gulf coast (e.g., Rappaport and Fernandez-Partagas 1995; Blake et al. 2007). The infamous Galveston hurricane of 1900 took at least 8000 lives and ranks as the deadliest single-day disaster in United States history. The loss of life (Beven et al. 2008) and the way of life suffered in 2005 from Hurricane Katrina show that the region remains at great risk.

About three tropical cyclones, including one hurricane, make landfall along the U.S. part of the Gulf coast each year on average (e.g., McAdie et al. 2009). Over the past 30 yr the Gulf coast accounted for almost two-thirds (34 of 54) of the hurricane landfalls in the contiguous United States. A ''major'' hurricane (MH), category 3 or higher

[1] In this study, ''tropical cyclone'' represents hurricanes, tropical storms and tropical depressions, as well as subtropical storms and subtropical depressions.

Corresponding author address: Edward N. Rappaport, National Hurricane Center, 11691 SW 17th St., Miami, FL 33029.
E-mail: edward.n.rappaport@noaa.gov

DOI: 10.1175/2010WAF2222369.1

on the Saffir–Simpson hurricane wind scale (SSHWS; Schott et al. 2010; Simpson 1974), strikes the northern Gulf coast almost every other year on average. While major hurricanes constitute only one-quarter of U.S. landfalling hurricanes, they cause around 85% of the damage (Pielke et al. 2008) and most of the fatalities (Blake et al. 2007).

Mitigating hurricane risk requires a more informed public, both well before and upon the final approach of a storm. Accurate operational tropical cyclone forecasts are essential. They are the responsibility of the U.S. National Hurricane Center (NHC) (e.g., Rappaport et al. 2009), a part of the National Weather Service (NWS). NHC's track prediction errors have been cut roughly in half over the past 15 yr, mirroring gains in operational computer model guidance (e.g., Franklin 2009).

Significant improvements in storm intensity forecasts, on the other hand, remain an unmet goal spanning decades (e.g., Hebert 1978, p. 831). The inability to make consistently accurate intensity forecasts has led NHC to list intensity forecasting as its top priority for the research community (JHT 2009) and NOAA has made it a focus of their recently established Hurricane Forecast Improvement Project (HFIP 2009).

NHC's intensity forecast errors in the Atlantic basin currently average about 10 kt[2] for 24-h forecasts and 15 kt for 48-h forecasts (Franklin 2009), a range corresponding to roughly a one category interval on the SSHWS. ''Rapid intensification,'' or RI, when systems intensify by at least 30 kt (about two SSHWS categories) in 24 h, occurs about 6% of the time (Kaplan et al. 2010) and rarely, if ever, is forecast accurately by the NHC.

Kaplan et al. (2010) discuss three influences on tropical cyclone intensity change identified by Marks et al. (1998): inner-core, large-scale atmosphere, and ocean processes. Their Statistical Hurricane Intensity Prediction System (SHIPS) developed for the Atlantic hurricane basin is among the best-performing intensity forecast guidance schemes available to operational forecasters (Franklin 2009).

It is in this era of limited intensity forecast capability that a spate of tropical cyclones occurred recently over the Gulf of Mexico. From 2003–05, for example, 15 tropical cyclones–including eight hurricanes, made landfall on the U.S. Gulf coast. The behavior of these storms reinforced a perception held by NHC hurricane specialists (forecasters) and others (e.g., Vickery and Wadhera 2008) that strong hurricanes, like Katrina and Rita in 2005, often weaken in their final hours prior to landfall along the northern Gulf coast. Both of those hurricanes reached category 5 intensity over the central Gulf before coming ashore at category 3 strength.

This study looks more closely at the intensity change characteristics of Gulf of Mexico tropical cyclones before their U.S. landfall. It begins by examining a basic potential relationship between a storm's ''initial'' intensity at periods of up to 2 days prior to landfall and its landfall intensity. That focus is prompted by a combination of the forecasters' perceptions, the reality that empirical intensity forecast methods remain competitive with more sophisticated approaches, and the observation that initial intensity information [sometimes expressed as a deficit from the maximum potential intensity (MPI), e.g., Emanuel (1988); Holland (1997)] contributes positively on average within the basin-wide framework to SHIPS. We seek to identify such relationships in the Gulf of Mexico region and the underlying causes for these systematic patterns of behavior. The goal of this study is to provide hurricane forecasters with improved objective intensity forecast guidance that can assist them in this important and challenging science and service area.

Section 2 describes the database and analysis approach. Section 3 presents our general results and possible connections to underlying physical processes. That discussion continues in section 4 with a focus on storms considered to be outliers. Section 5 covers some potential operational forecast considerations. Section 6 summarizes the findings and looks ahead.

2. Data and analysis method

The study period covers 1979–2008. Landfall refers to the time when the center of the tropical cyclone crosses the coast.[3] We have selected only those tropical cyclones for which NHC issued operational forecasts and that made landfall between the U.S.–Mexico border near Brownsville, Texas, and the southern tip of the Florida Peninsula near Flamingo. The Florida Keys were not considered land for the purposes of designating landfall.

To qualify, a tropical cyclone center must have spent the entirety of the period of interest (from 12 to 48 h, ending at U.S. landfall) over the Gulf of Mexico. The Yucatan Channel bounded that area on the south, and 81°W formed the Gulf boundary on the southeast.

[2] Speeds are provided in knots, as done in forecast operations; 1 kt = 0.51 m s^{-1}.

[3] We included two exceptions documented previously. Hurricane Juan's (1985) center looped just offshore Louisiana, close enough to the coast to bring the maximum winds over land (Blake et al. 2007). The data for that event are included here as the first of its two landfalls. Hurricane Erin (1995) took an oblique path to the Florida Panhandle coast, placing the eyewall over land about 2.5 h before the center crossed the shoreline. Time and intensity of the eyewall arrival were used as the landfall data.

TABLE 1. Intensity data (kt) for U.S. Gulf coast landfalling hurricanes (1979–2008). NA indicates that the system did not continuously have a tropical cyclone center over the Gulf from that time until landfall.

Hurricane	Landfall time and date	Intensity at landfall	Intensity 12 h before landfall	Intensity 24 h before landfall	Intensity 36 h before landfall	Intensity 48 h before landfall
Ike	0700 UTC 13 Sep 2008	95	91	90	85	85
Gustav	1500 UTC 1 Sep 2008	90	95	98	115	NA
Dolly	1820 UTC 23 Jul 2008	75	71	60	46	45
Humberto	0700 UTC 13 Sep 2007	80	47	NA	NA	NA
Wilma	1030 UTC 24 Oct 2005	105	94	85	NA	NA
Rita	0740 UTC 24 Sep 2005	100	109	115	124	151
Katrina	1110 UTC 29 Aug 2005	110	142	142	100	99
Dennis	1930 UTC 10 Jul 2005	105	124	94	76	NA
Cindy	0300 UTC 6 Jul 2005	65	53	33	28	NA
Ivan1	0650 UTC 16 Sep 2004	105	114	119	120	138
Charley	1945 UTC 13 Aug 2004	130	102	NA	NA	NA
Claudette	1530 UTC 15 Jul 2003	80	63	58	55	50
Lili	1300 UTC 3 Oct 2002	80	122	112	92	NA
Irene	2000 UTC 15 Oct 1999	70	65	60	60	NA
Bret	0000 UTC 23 Aug 1999	100	125	120	80	65
Georges	1130 UTC 28 Sep 1998	90	95	95	90	90
Earl	0600 UTC 3 Sep 1998	70	85	50	50	40
Danny	0900 UTC 18 Jul 1997	65	53	35	30	NA
Opal	2200 UTC 4 Oct 1995	100	130	95	78	68
Erin	1600 UTC 3 Aug 1995	75	68	NA	NA	NA
Andrew	0830 UTC 26 Aug 1992	105	120	115	115	NA
Jerry	0030 UTC 16 Oct 1989	75	60	55	55	55
Chantal	1300 UTC 1 Aug 1989	70	66	51	31	21
Florence	0200 UTC 10 Sep 1988	70	58	50	45	45
Bonnie	1000 UTC 26 Jun 1986	75	68	53	43	28
Kate	2230 UTC 21 Nov 1985	85	96	105	103	84
Juan2	1130 UTC 29 Oct 1985	65	75	NA	NA	NA
Juan1	1430 UTC 28 Oct 1985	75	67	57	47	37
Elena	1300 UTC 2 Sep 1985	100	109	106	96	90
Danny	1630 UTC 15 Aug 1985	80	74	58	43	30
Alicia	0700 UTC 18 Aug 1983	100	91	71	61	51
Allen	0600 UTC 10 Aug 1980	100	125	155	130	130
Frederic	0300 UTC 13 Sep 1979	115	115	108	90	75
Bob	1200 UTC 11 Jul 1979	65	65	50	30	20

Using these criteria, 89 U.S. Gulf coast tropical cyclone landfalls occurred during the 30-yr period. They consist of 17 tropical depression, 38 tropical storm, and 34 hurricane (Table 1) landfalls (Fig. 1). This includes two systems, Juan (1985) and Ivan (2004), credited with making landfall twice on the U.S. Gulf coast.

We now change from a landfall to a forecasting reference frame. Analyses were conducted at 12, 24, 36, and 48 h prior to landfall. The total number of systems analyzed generally decreases with increasing forecast period from the initial time to landfall because some systems were not continuously tropical cyclones during the final 48 h before coming ashore. In addition, the numbers of cases shift around between hurricanes, tropical storms, and tropical depressions from one initial time to another (e.g., from 36 to 24 h prior to landfall), as systems strengthen or weaken upon their approach to land. For example, 12 h prior to landfall there were forecasts—for at that time—28 hurricanes, 46 tropical storms, and 15 tropical depressions. Still other Gulf of Mexico tropical cyclones do not contribute to this study because they spent less than 12 h over the Gulf before moving ashore, dissipated over the Gulf, or made landfall in Mexico.

a. Tropical cyclone data

We use the intensity and center location information in NHC's ''best track'' database (available online at http://www.nhc.noaa.gov/pastall.shtml#hurdat), which contains the NHC's poststorm analysis estimates. The dataset is based on the NWS's definition of tropical cyclone intensity: the maximum sustained 1-min average, surface (10-m elevation) wind speed of a tropical cyclone (NWS 2009). The database contains 6-hourly representative estimates (at 0600, 1200 UTC, etc.) of the cyclone's center location to the nearest 0.1° and intensity to the nearest 5 kt. NHC began documenting the landfall location and intensity

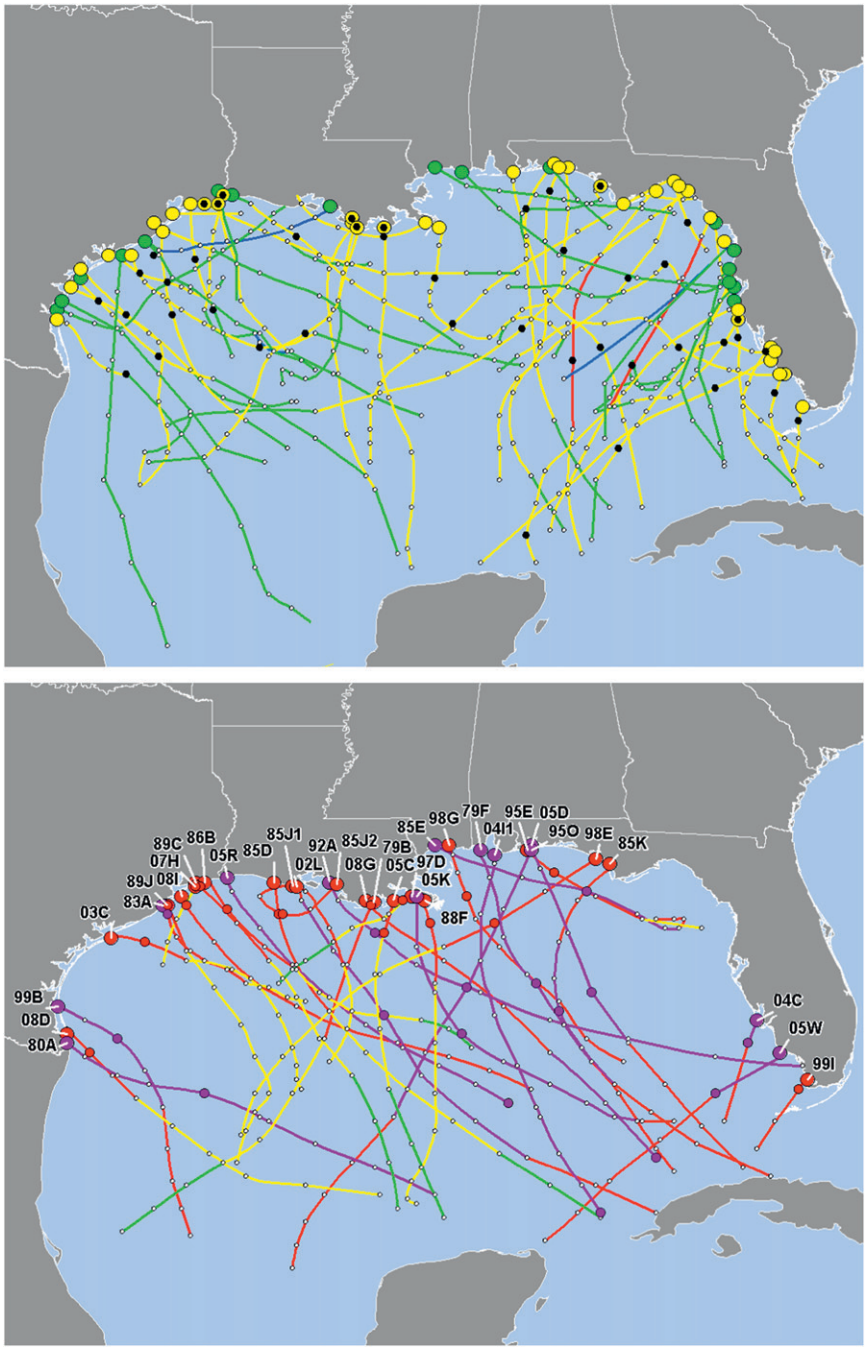

FIG. 1. Tracks of 1979–2008 Gulf coast tropical cyclones during the final ~48 h before U.S. landfall. (top) The landfalling tropical depressions (green) and tropical storms (yellow) and (bottom) the landfalling hurricanes (red), including major hurricanes (purple). The largest dots show landfall locations and intensity stages. The middle-sized dots indicate the locations and stages at maximum intensity. The smallest circles indicate the remaining 6-h positions. The blue line denotes the subtropical phase. Labels in the bottom panel are based on the final two digits of the year and the first letter of the hurricane name.

estimates for tropical storms and hurricanes in the 1980s and for tropical depressions in 1990 in the NHC's Tropical Cyclone Reports.[4] We estimated landfall statistics for the earlier ~15% of the systems back to 1979 from the 6-hourly best-track database and NHC reports on individual systems. After identifying the landfall time and intensity, we stepped back 12, 24, 36, and 48 h to calculate the net change of intensity during those periods. While we could have performed 6-hourly analyses, we chose the above 12-h periods because they match NHC's short-term forecast periods. We linearly interpolated the initial intensity from the 6-hourly data when the anchoring landfall time did not coincide with the 6-hourly best-track times.

b. Ocean heat content (OHC) data

We examined the ocean heat content (OHC) to help determine the oceanic influences on the tropical cyclones. OHC measures the amount of thermal energy in the upper ocean per unit area above an ocean temperature of 26°C. It was calculated using the method described in Shay et al. (2000). Although the OHC predictor is only of secondary importance in the basin-wide SHIPS model, it provides more useful information about the ocean structure (including the mixed layer) than does simply using the sea surface temperature (SST) in the Gulf of Mexico because it locates the energetic mesoscale eddy structure and the Loop Current that have deep, warm structures (e.g., Jaimes and Shay 2009). For example, Mainelli et al. 2008 (hereafter, M08), with a focus on the Gulf of Mexico, showed the importance of OHC on category 5 hurricanes.

Daily OHC analyses estimated from satellite altimetry are available back to 1995. We used a hurricane season climatology, as described by Mainelli-Huber (2000), instead of the annual climatology discussed in Shay et al. (2000).

c. Other atmospheric and oceanic variables

In addition to the OHC data described above, variables from the storm environment are examined to help explain the behavior of the tropical cyclone intensity changes. These variables are obtained from the developmental database for the SHIPS model (DeMaria et al. 2005). Atmospheric variables include the magnitude of the 850–200-hPa shear, the 200-hPa temperature, the 200-hPa divergence, and the relative humidity (RH) in the 850–700-hPa layer. The first three variables are the most important atmospheric predictors in SHIPS. The RH variable is of lesser importance for the basin-wide sample, but was included because of the possible interactions with dry environmental air as storms approach the Gulf coast. The oceanic variables include the SST and OHC at the storm center. All atmospheric variables were obtained from the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS). The SST is obtained from the weekly Reynolds SST analyses and the OHC values are taken from the fields referenced in section 2b.

The SHIPS data are available back to 1982. All available cases for the same forecast sample described in section 2a were obtained. As will be described later, the primary interest is in cases at hurricane intensity at the advisory time. With this restriction, the SHIPS sample includes 60 cases with at least a 12-h forecast, decreasing to 8 cases with a 48-h forecast.

In real time the SHIPS atmospheric parameters are determined along the forecast track from GFS forecast fields. For the developmental data, the variables are derived from GFS analyses along the observed track. Since our primary interest is in understanding the differences in intensity changes, the developmental data are used here. This is actually a necessity because the GFS forecast fields are not available back to 1982. For the forecast verification described in section 5, however, the SHIPS model was run with the NHC official track rather than the best track, for consistency with the verification of the official forecast and other guidance models.

The SST is not used directly in the SHIPS model. Instead, the SST is used to estimate the MPI from an empirical relationship developed by DeMaria and Kaplan (1994) with a small correction to account for the storm translational speed. The MPI was determined by finding the maximum observed intensity for each value of SST from a long-term sample of Atlantic tropical cyclones.

The SHIPS model uses a linear regression method and has been available in real time since 1990. The Logistic Growth Equation Model (LGEM) was developed to relax some of the linear constraints of SHIPS and has been available in real time since 2006. As part of the LGEM development, a modified MPI relationship accounting for environmental vertical shear was derived (DeMaria 2009). The modified MPI is also a function of a vertical instability parameter. However, that parameter is not available back to 1982 due to inconsistencies in some of the thermodynamic variables from the operational and reanalysis GFS fields used to develop SHIPS. When the vertical instability parameter is above the basin-wide mean (which is nearly always the case in the Gulf of Mexico), the modified MPI depends mostly on the shear. In this study the modified MPI is calculated assuming a constant value of the vertical instability that is one standard deviation above the

[4] Tropical Cyclone Reports, known as Preliminary Reports prior to 2000, are available online at http://www.nhc.noaa.gov/pastall.shtml.

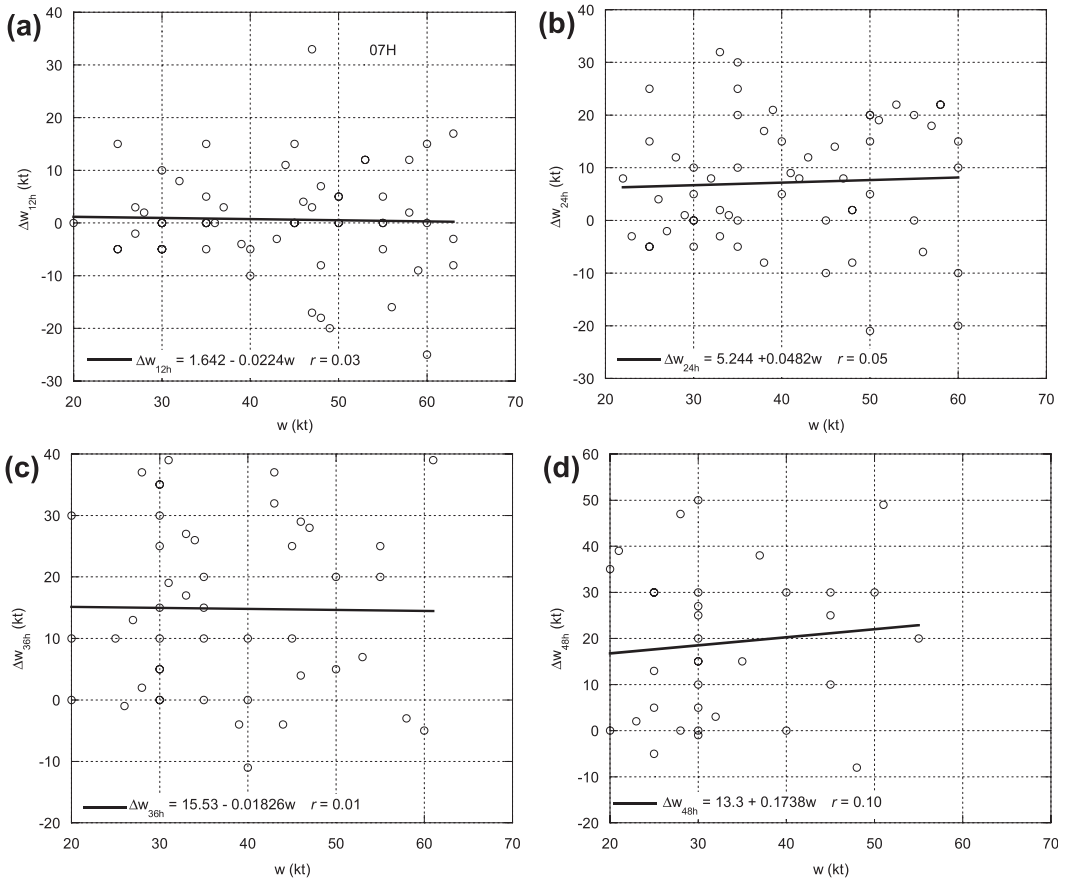

FIG. 2. Intensity change Δw (kt) of 1979–2008 U.S. Gulf coast landfalling tropical cyclones as a function of initial intensity w (kt) as a tropical depression or tropical storm (a) 12, (b) 24, (c) 36, and (d) 48 h prior to landfall. A circle can represent multiple cyclones. The lines are least square fits based on the regression equation at the bottom of each panel. The associated correlation coefficient is given by r. The labeled point in (a) is discussed in the text.

Atlantic sample mean from the 2001–08 cases for which operational GFS analyses are available.

3. Generalized patterns of behavior and their interpretation

To simplify interpretation, and in consideration of preliminary analyses, we binned the tropical cyclones into two groups based on their intensity at the forecast start time. One group contains tropical depressions and tropical storms, and the other comprises hurricanes.

a. Tropical storms and tropical depressions

Tropical storms and depressions took varied tracks to land, with most attaining their peak intensity just before or upon reaching the coast (Fig. 1a).

Figure 2 shows intensity change as a function of the initial intensity of the tropical depressions and tropical storms for the four time periods. It also provides a linear least squares regression line and equation for each forecast period and the associated correlation coefficient r. The panels in Fig. 2 indicate a large historical spread of possible intensity changes, as much as 50–70 kt, for each forecast period. The fit lines in all four panels are essentially flat and, as quantified by the near-zero correlation for each line, indicate that initial intensity has little predictive value for intensity change for these systems. On the other hand, the flat fit lines in Fig. 2 shift upward with increasing period to landfall. There is no net change in intensity for depressions or storms in their final 12 h before landfall. For longer periods over the Gulf, depressions and tropical storms strengthen at an average rate that is independent of the initial intensity, about 7 kt for every additional 12 h over the water, reaching about +20 kt by 48 h.

b. Hurricanes

Most of the hurricane landfalls during the period occurred along the central northern Gulf coast, between Galveston, Texas, and Apalachicola, Florida (Fig. 1b).

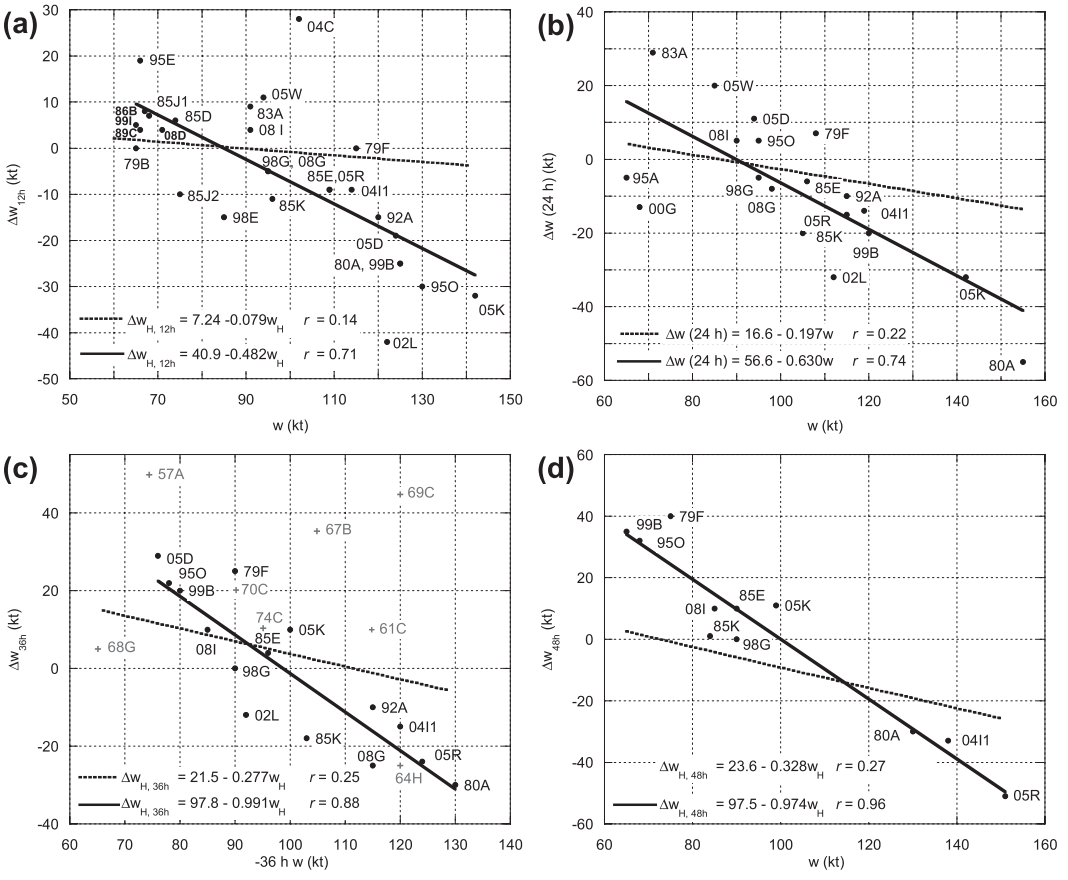

FIG. 3. As in Fig. 2, but with solid lines for the initial intensity of Gulf hurricanes (H) and dashed lines for the remainder of Atlantic basin hurricanes of 1979–2008. The labels are based on the final two digits of the year and the first letter of the hurricane's name. A fourth digit indicates a hurricane's first (1) or second (2) landfall. The gray labels in (c) denote 1955–78 Gulf hurricanes.

U.S. Gulf coast hurricane activity outside that area was limited for several hundred miles, increasing again over far southern Texas and the southwestern Florida coast.

Figure 1b shows two primary modes of motion for the landfalling hurricanes. Most moved toward the northwest or west-northwest, whereas a smaller group had mainly northeastward headings. The average forward speed for hurricanes was about 10 kt, exceeding the average of about 8 kt for weaker tropical cyclones. About half of the hurricanes, 16 of 33, were major hurricanes for part of their time in the Gulf.

Hurricanes approaching the U.S. Gulf coast behaved differently from tropical storms and tropical depressions. The amount and the sign of the intensity change in hurricanes are strong functions of initial intensity (Fig. 3), while neither relationship exists for tropical storms and depressions (Fig. 2). On average, the weakest hurricanes strengthen the most and the strongest hurricanes weaken the most at all lead times. The transition from one end of the initial intensity scale (65 kt, category 1) to the other (~160 kt, category 5) is nearly linear and passes through zero change between 85 and 100 kt in each of the periods (Figs. 3a–d). That is, for tropical cyclones in the Gulf, on average, all roads lead to near the 95–96-kt category 2–3 threshold of major hurricane status.

Figure 3a shows an r of 0.71 for the line at 12 h. The correlation is noteworthy given r was near zero for tropical storms and depressions. If a relationship exists between a storm's initial intensity and the amount of intensity change before landfall, then one might expect it to be strongest for intensity change at the shortest forecast period. For hurricanes, remarkably, *r increases* with time to landfall, reaching 0.96 by 48 h. That is, over the past 30 yr, a Gulf hurricane's intensity 2 days before U.S. landfall was a predictor of high accuracy for the amount of intensity change by landfall (and, therefore, the landfall intensity), explaining more than 90% of the variance.

We assessed the statistical significance of the lines in Fig. 3 in two ways. We evaluated the statistical significance of the regression lines using the standard t test on the null

hypothesis of zero slope. That null hypothesis was rejected at the 99% level at all time periods. We also compared the Gulf landfall cases to the remaining thousands of forecasts made for the 1979–2008 Atlantic basin hurricanes that stayed over water during the selected forecast periods. The regression lines for them are dashed in Fig. 3. The lines for the basin-wide cases, like the Gulf landfall subset, indicate positive intensity change for low-end hurricanes and negative intensity change for the strongest hurricanes. We applied the Chow (1960) test on the null hypothesis that the regression coefficients for the Gulf-landfall and non-Gulf-landfall cases were the same. To perform this test, we examined the entire sample of overwater 12-, 24-, 36-, and 48-h best-track intensity changes for all tropical cyclones to determine the serial correlation between such intensity changes computed from the 6-hourly sequential best-track points, using a methodology adapted from Siegel (1956) (e.g., Aberson and DeMaria 1994). This gave the effective sample size. Following Laurmann and Gates (1977), the effective sample size was then used in place of the actual sample size in the calculation of the Chow statistic and the degrees of freedom of the F distribution. The statistic indicates the null hypothesis can be rejected at the 99% level for the 12-h intensity changes, but it cannot at the 90% threshold at 24–48 h. We attribute the latter result to a combination of the small number of Gulf landfall cases (even for a 30-yr sample), the similar—but less extreme—slope of the basin-wide regression lines, and the large scatter (low r) about the basin-wide lines.

We also performed an analysis of the central pressure because it is sometimes considered a proxy for tropical cyclone intensity and can be measured independently from wind speed. Central pressure changes (not shown) were consistent with, and correlations were similar to, the findings for wind speed for the respective tropical depression–tropical storm and hurricane subgroups.

It is interesting that the relationships and large correlation coefficients discussed above hold even though many of the tropical cyclones underwent eyewall replacement cycles (e.g., Willoughby et al. 1982) and/or RI periods. Regarding the latter, however, we found no hurricane (or tropical storm) RI events that began in the northeastern Gulf north of 25(0.1)°N, east of 90°W. A review of NHC's 6-hourly best-track database for earlier years indicates RI cases beginning there are very rare. Eloise (1975) and the first storm of 1945 could be the only two cyclones in the past 100 yr to qualify.[5]

[5] As noted earlier, landfall intensities are not available in general prior to 1982. It is possible one or more storms reached an intensity (e.g., at landfall or between 6 h best track points) higher than indicated in the 6-hourly data, qualifying them as undergoing RI.

c. *Associated physical processes*

This section begins a discussion of patterns of behavior in the intensity change identified above in terms of internal and environmental factors. In section 4 we continue our look at the processes but focus there on their relevance to storms that are the outliers in Figs. 2 and 3.

As noted in the introduction, many environmental factors can influence Atlantic basin tropical cyclone intensity. Among these for the Gulf are interactions with the Loop Current and associated upper-ocean (Gulf) eddies, airmass characteristics of the flow approaching the cyclone from elevated terrain to the southwest through northwest, interaction with midlatitude baroclinic systems, modifications to internal storm structure occurring during the tropical cyclone's passage over land (Cuba, Yucatan Peninsula, or Florida), and changes in the low-level wind field associated with frictional inflow from coastal areas.

While each of these influences can be important independently or in combination for any event, given the infinite variety of possible configurations and influences, it is hard to make the case that the atmospheric environment alone works in such a systematic way as to generate the signal seen in the general patterns of behavior of the Gulf tropical cyclones seen in Figs. 2 and 3. This assumption is confirmed below, when the quantitative analysis of the SHIPS variables is described. As indicated in this paper's referenced studies on OHC, we reason that the ocean, through its heat and moisture fluxes, often helps govern the behavior of Gulf tropical cyclones in an important way. Figure 4 shows where the tropical cyclones reached their maximum intensities plotted on the average OHC values for June–November, as calculated from available 1995–2008 operational OHC analyses. (If the best track contained multiple occurrences of the highest wind speed, then the first occasion among them with the lowest central pressure at those times was plotted.) Figure 4 shows most of the MH reached their peak intensity over the central Gulf, in the area of highest seasonally averaged OHC. Closer inspection of contemporaneous OHC data (e.g., Fig. 2 in Shay et al. 2000) for the strongest hurricanes since 1995 hints at a tendency for most of these storms to reach or retain their maximum wind speed, or minimum central pressure, ~6–18 h after passing the local OHC maximum. Using autocorrelation analyses, increases in hurricane intensity (surface pressure decreases) suggest a lag of about 12–15 h upon encountering these deep warm heat reservoirs such as warm-core eddies or the Loop Current itself (Mainelli-Huber 2000; Shay and Uhlhorn 2008).

Category 1 and 2 hurricanes (as well as most tropical storms; cf. Fig. 1a) attained their maximum intensities near the shore. The only two Gulf hurricanes over the

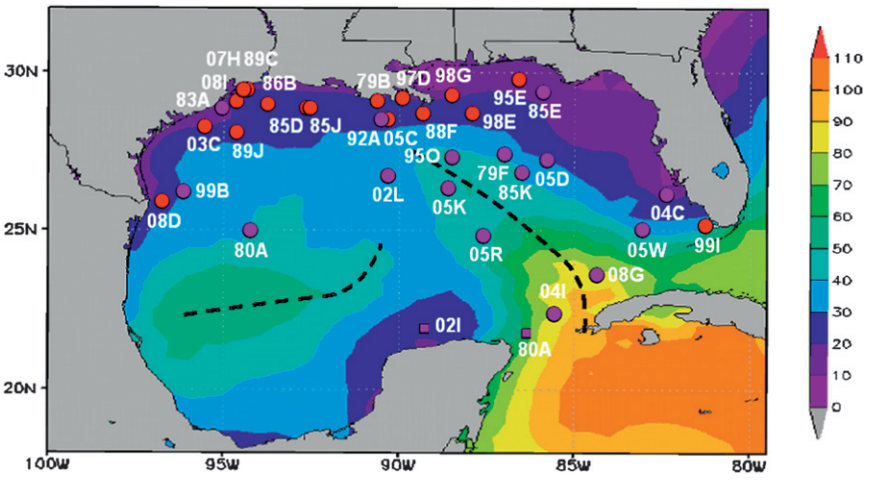

FIG. 4. Locations and magnitudes of maximum intensities reached by 1979–2008 hurricanes making landfall on the U.S. gulf coast, superposed on a plot of the average June–November OHC (kJ cm $^{-2}$) values calculated from archived 1995–2008 operational OHC analyses. The purple dots indicate major hurricanes and the red dots show other hurricanes. The squares near the Yucatan Peninsula indicate points where Hurricanes Allen (1980; 80A) and Isidore (2002, 02I) reached their absolute maximum intensities over the Gulf more than 48 h from U.S. landfall. The labels show the final two digits of the year and the first letter of the hurricane name, with a trailing 1 for the cyclone's first of multiple landfalls. The dashed lines show the primary axes of OHC maxima in the Gulf of Mexico.

30-yr period to weaken back to tropical storm status before U.S. landfall were Allison (1995) and Gordon (2000). Both moved ashore in the Florida Big Bend area north of Tampa. The above observations are consistent with previous research. In particular, M08 analyzed several category 5 hurricanes—including some in the database of this study. They concluded that "the upper ocean thermal structure is fundamental to accurately forecasting intensity changes of tropical cyclones." M08 observed that the strongest hurricanes had passed over relatively high OHC regions, like the Loop Current or a warm core eddy that it sheds. From a slightly different perspective, these deep heat reservoirs are nearly equivalent to keeping SST constant as in uncoupled atmospheric models since these features are more resistant to cooling by hurricane-induced turbulent mixing, making them important to the evolution of strong hurricanes (Shay and Uhlhorn 2008). From a forecasting perspective, the incorporation of OHC in SHIPS improved the intensity forecast guidance basin wide by as much as 5% on average at one forecast time (84 h). The largest improvement for an individual case was 20%, during Hurricane Ivan (M08).

For any given forecast period, a tropical cyclone contributes to this study through only one of the data subsets, either to tropical storms and depressions or to hurricanes. This led us to compare the statistics derived for these temporally independent subsets near the tropical storm–hurricane threshold (60–65 kt) (cf. Figs. 2a, 3a and 2b, 3b, etc.). The strengthening seen at 65 kt is about 5–15 kt greater than indicated for 60-kt storms. This might be a measure of the noise in the sample and/or analysis technique. Alternatively, it might point to a fundamental difference in the intensity change responses of tropical storms and low-end hurricanes to environmental conditions. The 64-kt operational threshold between tropical storms and hurricanes evolved from surface weather and sea conditions described in the development of the Beaufort scale two centuries ago rather than any then-known physical differences in cyclones near that intensity. Yet, the latter might also be the case, as suggested by Kaplan et al. (2010). For example, low-end hurricanes usually have at least rudimentary eyewall structures and secondary circulations, and, with inertial stability varying with vortex strength (e.g., Nolan et al. 2007), would be expected to respond differently to their environments than tropical storms.

The SHIPS model database is used to further explain why the MH tended to weaken before landfall but the nonmajor hurricanes (NMHs) tended to intensify, so that both evolved toward maximum winds around 90–100 kt at landfall. As described in section 2c, there were 60 cases in the SHIPS database from 1982 to 2008 that were of hurricane intensity over the Gulf at the beginning of a forecast period. These were divided into cases that were initially MHs (28 cases with a mean intensity of 118 kt) and initially NMHs (32 cases with a mean intensity of 80 kt). The SHIPS variables are examined for each 6-h period from 0 to 36 h up to the time just

before landfall. The sample size after 36 h was too small for a meaningful comparison of the MH and NMH cases.

A comparison of the atmospheric variables between the two groups showed that there were no significant differences in the 200-hPa temperature and divergence from 0 to 36 h. The average vertical shear, however, was lower and the low-level RH was higher along the storm tracks at all time periods from 0 to 36 h for the MH cases. This result indicates that the atmospheric variables were actually more favorable for the MH cases, and does not explain why those storms tended to weaken while the NMHs continued to intensify.

The next step was to calculate the MPI and modified MPI along the storm track for the MH and NMH cases; the results are shown in Fig. 5. If the average MPI along the storm track is greater than the average initial intensity for each subgroup (MH or NHM), that indicates there is room for further intensification. The average MPI in Fig. 5 was a little higher for the MH cases because the SSTs were a little higher, but for both the MH and NMH cases, the average MPI was high enough to support a category 5 storm up to landfall. This MPI represents the maximum intensity for a given SST under the most ideal atmospheric and subsurface oceanic conditions. Figure 5 shows the average shear-modified MPI values, which are considerably lower than the original MPIs for both the MH and NMH samples. As described above, the average initial intensity of the MH sample is 118 kt. Thus, those cases are already close to their modified MPI values so there is little room for further intensification. In contrast, the average initial intensity of the NMH sample is 80 kt, so there is still room for some strengthening.

Because the modified MPI is about the same as the initial intensity for the MH cases, it still does not explain why these storms weakened before landfall. The SST values used in the MPI calculation are from the prestorm environment. As described above, however, the SST cools underneath the storm due to upwelling and mixing. Although the SST behind the storm can be up to 5°C lower by 1–2 days after the storm passes, the SST directly below the storm's center usually only falls by about 1°C, according to Cione and Uhlhorn (2003). Their study also suggests that for a deep mixed layer the magnitude of the inner-core SST cooling is generally less for fast-moving and relatively weak storms. In the region of the Loop Current (the high OHC areas in Fig. 4), the SST reduction is probably very small (e.g., Shay and Uhlhorn 2008). Along the northern Gulf coast, however, where the OHC is much lower, the SST drop under the storm is probably closer to 1°C. Figure 5 shows the shear-modified MPI, assuming a 1°C cooling. For the MH cases, the sea surface cooling lowers the modified MPI to values between 100 and 110 kt. For the NMH the cooling lowers the modified MPI to about 90 kt. This result indicates that when the MHs move over the low OHC values of the northern Gulf, the SST reduction they induce lowers the modified MPI to a value below the average initial intensity. In contrast, even with the SST reduction, the NMH cases still have a little room to intensify. The MPI modified by wind shear and SST help explain why the MHs (NMHs) tend to weaken (strengthen) toward the category 2–3 threshold.

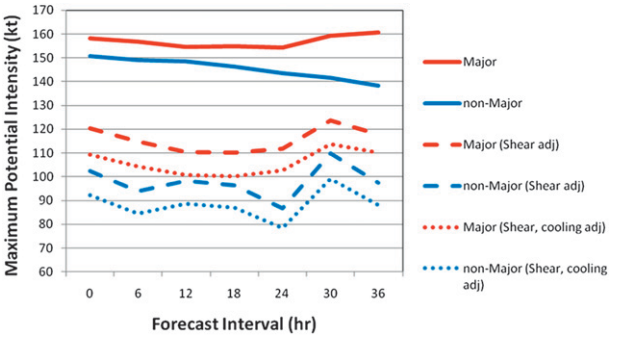

FIG. 5. The unmodified MPI (solid), shear-modified MPI (dashed), and shear-modified MPI with 1°C of SST cooling (dotted) obtained from the SHIPS formulation for U.S. landfalling Gulf hurricanes.

4. Special cases and associated processes

Section 3 presented results and discussion from the perspective of broad averages. The dataset, however, also contains a number of distinct outliers. Beyond keeping the correlation coefficients from being even higher than noted, the outliers serve as important exceptions to relying solely on the formulas shown in the panels of Fig. 3 for operational forecast applications. Some of the cyclones with the greatest departures from average are discussed in this section.

a. Hurricanes Charley (2004), Humberto (2007), and Juan (1985)

Charley (2004) (Franklin et al. 2006) contributes to the database at only one time, 12 h. It merits special attention, however, because its 130-kt intensity at landfall exceeds all others in this study and because of its very large departure from the mean intensity change statistics for 12 h (see 04C in Fig. 3a). During its final 12 h over the Gulf, Charley intensified by about 30 kt, an amount that would have met the criteria for RI even if it had taken 24 h to occur.

Charley's surface circulation was very small. Based mainly on reconnaissance aircraft data, the NHC estimated the distance from the center to the eyewall as only 2.5 n mi (1 n mi = 1.85 km). Hurricane-force winds extended outward in the direction of motion just 15–20 n mi,

and tropical-storm-force winds extended outward from the center only 40–45 n mi. In a hurricane of average size and intensity, hurricane-force winds occur outward to about 40 n mi with tropical-storm-force winds found out to about 110 n mi (Knaff et al. 2007). Radii in the largest hurricanes can be 5–10 times larger than those found in Charley.

Forecasters have suggested from their experience that small tropical cyclones can strengthen or weaken quickly relative to larger systems. While the authors are not aware of an observational study focused on this relationship, several papers have touched on the issue. DeMaria (1996) indicated that small, low-latitude weak storms are more sensitive to the wind shear than their counterparts. Kaplan and DeMaria (2003), referencing DeMaria and Kaplan (1994), found the most rapidly intensifying systems are ''smaller than average, were in an environment with low vertical shear and weak upper-level forcing, and were further from their empirically derived maximum potential intensity'' (e.g., were relatively weak). Conversely, 850-hPa positive vorticity contributes to forecast intensity increases in SHIPS, and others (e.g., Hebert 1980, p. 985) have suggested that larger storm circulations presage strengthening.

SHIPS diagnosed less than 5 kt of shear and an SST of about 30°C as Charley neared the southwestern Florida Peninsula. Other SHIPS contributors were near their norms. Charley then presented a mixed signal within the context of those studies, being exceptionally small, in a very light shear environment, over seasonal maximum water temperatures, but it was already strong (~100 kt) when its fast strengthening occurred.

Charley's forward speed was also unusual for Gulf hurricanes. It moved at between 16 and 19 kt during its final 12 h over water, whereas speeds nearer 10 kt are typical. Charley's very small circulation and very fast forward speed combined to limit the amount of time the hurricane's eyewall spent over waters roughened (upwelled and mixed) by the leading part of the hurricane's surface circulation. We can calculate a period of roughening for the Gulf P_{rG} corresponding to the period that ≥34-kt winds are experienced by the waters over which the center passes (i.e., the 34-kt radius along the direction of motion, divided by the forward speed). For Charley, P_{rG} is 2–3 h. Because of this low value of P_{rG}, the SST cooling under the storm could have been close to zero. Assuming no SST cooling and with the low vertical shear, the modified MPI for Charley was about 145 kt, which is much larger than the sample mean for the major hurricanes shown in Fig. 5. Thus, Charley did have room for continued intensification up to landfall.

Hurricane Humberto (Brennan et al. 2009) also provided only one data point to the sample (07H in Fig. 2a). While not a major hurricane, it evolved in a manner similar to Charley's final hours over the Gulf. Humberto is important because it was a rare example of a tropical cyclone forming just offshore and then moving almost immediately inland as a hurricane overnight. Such developments pose serious public preparedness concerns. Like Charley, Humberto's surface circulation was very small. The P_{rG} for Humberto was about 6 h.

The small circulations in Charley and Humberto also meant that those cyclones had little time to ''feel'' any weakening influences of land (e.g., through increased surface roughness) before their respective central cores came ashore. This could be a secondary contributor to their intensity changes.

It is expected that larger, major hurricanes like Katrina and Ivan (2004) (Franklin et al. 2006), with a long exposure to the great expanses of upwelled and mixed waters ahead of their respective centers, could evolve differently than Charley and Humberto. The NHC 1979–2008 database contains 11 major hurricanes besides Charley 12 h before U.S. Gulf landfall. All 11 hurricanes weakened in their final hours before landfall except for Frederic (1979), which had no change in intensity. Using the approach above, the average P_{rG} was 13 h for all the major hurricanes about a half day before landfall. Excluding Charley, the periods range from 9 to 20 h. In these cases, the MPI using the adjustments for shear and cooling in Fig. 5 would apply and explain the weakening before landfall. Juan's 1985 development (Case 1986) can also be explained in part by its movement. Juan's center moved in an offshore loop extending to near the northern Gulf coast. After the loop, Juan's center moved east-northeastward across its own wake before arriving on the southeastern Louisiana coast. Juan's weakening before landfall, which stands out as data point 85J2 in Fig. 3a, is consistent with the storm spending its last 3 days passing over waters cooled by its own circulation. It is also consistent with the conclusion of M08 that a negative feedback is especially effective when storms are (nearly) stationary for a few days.

In summary, the P_{rG} would likely be most important in areas where OHC is high enough to support hurricane intensity, but where a rather large vertical gradient in water temperatures (steep thermocline) keeps it below peak values. This is the case for the Gulf common water between the U.S. coast and the area of maximum OHC over the Gulf. It would not be expected to be as important over the central Gulf where OHC is high because warm waters extend to much greater depths (Shay and Uhlhorn 2008).

b. *Hurricane Isidore*

Hurricane Isidore (2002) (Pasch et al. 2004) represents a class of tropical cyclones where the effects of a previous passage over land disrupt the storm's convective

structure and circulation to the point where recovery is delayed for hours to days, or precluded altogether. Isidore weakened from a hurricane at the high end of category 3 to a tropical storm during its center's ~36 h over the Yucatan Peninsula. It never regained hurricane strength on its 48-h overwater trek northward to Louisiana.

c. *Hurricanes before 1979*

Several category 5 hurricanes contribute to the 1979–2008 records. At landfall, however, no category 5 and only two category 4 hurricanes (Frederic and Charley) occurred. NHC's archives begin in 1851 and indicate as many as eight Gulf landfalls stronger than Charley and Frederic before 1979 (cf. Blake et al. 2007), with the exact number depending on whether the wind speed or central pressure is the standard. It is of interest to know how well the tendencies identified in this study apply to earlier or to future especially intense hurricanes. This interest, and the observation that the very large 36- and 48-h correlations of near one are based on rather small numbers of cases, prompted us to extend some parts of the analysis to before 1979.

We added the 1955–78 hurricanes to our study. The initial year corresponds approximately to the time that reconnaissance aircraft began to fly routinely to the center of a tropical cyclone. The data from this period, however, were sometimes not reliable, as noted by Sheets (1990). Further, there were fewer observation systems: no Stepped Frequency Microwave Radiometers to estimate surface wind speeds, GPS dropsonde data in the eyewall, or ground-based Doppler radar data. Satellite images were less frequent and of lesser quality. The Dvorak (1984) analysis technique for estimating intensity, which came into operational use in the early 1970s, was not available in the early part of the period. Similarly, the interpretation of flight-level and dropsonde data obtained from reconnaissance aircraft has evolved (e.g., Franklin et al. 2003).

Also, as noted earlier, NHC did not then provide an estimate of the landfall intensity in its poststorm reports. For 1955–78, we used the last offshore intensity in the best-track file. Doing so excludes changes in intensity occurring up to 5 h before landfall.

Figure 3c shows in gray the eight cases at the 36-h period from 1955 to 1978. Their characteristics are representative of the data at the other forecast periods. Three of the hurricanes, Hilda (1964), Celia (1970), and Carmen (1974), evolved similarly to the hurricanes in the 1979–2008 period, falling within the envelope comprising the more recent hurricanes. Three others, Audrey (1957), Carla (1961), and Gladys (1968), lie outside that envelope. If data and analysis issues for them can be set aside, these three hurricanes give evidence that on rather rare occasions there will be a Gulf hurricane, even a critical one, in which the relationships presented and the behaviors identified to date for outlying storms will not apply. These three cases also suggest that correlation coefficients for a longer period might be lower than those found for 1979–2008.

The remaining two cases, Beulah (1967) and Camille (1969), warrant additional discussion. The report from the reconnaissance aircraft about 15 h before Beulah's landfall along the Texas–Mexico border indicated a flight-level wind of 140 kt, the highest observed for that hurricane, and a central pressure of 923 mb. Subsequent data implied significant weakening. Aircraft flight-level winds did not exceed 100 kt thereafter. The central pressure rose to 936 mb on the last center penetration about 3 h before landfall and to near 948 mb at landfall according to NHC's contemporary surface analysis. The maximum surface winds in the best-track dataset, however, do not reflect the implied weakening—remaining at 140 kt until landfall. This suggests the ''67B'' point in Fig. 3c should be discarded. In contrast, Beulah's central pressure changes fall within the envelope of the 1979–2008 pressure changes (not shown).

Camille is the only category 5 hurricane at Gulf landfall in the archive. We ask, how did Camille's intensity change prior to landfall? Was it even stronger 12–24 h before landfall, as would be implied by the 1979–2008 data shown in Fig. 3?

During Camille's final 48 h over water there were only three aircraft reconnaissance missions to the center of the hurricane, about 33, 28, and 10 h before landfall. The first of these provided no wind data near the center and the last was shortened.

There is insufficient reliable data to estimate with confidence for Camille the temporal variations in maximum wind speed this study requires. On the other hand, the reconnaissance aircraft reported central pressures of 900–910 mb on each of the three flights noted above. This suggests Camille did not weaken much as its center neared land, an outlying pattern of behavior compared to the strongest hurricanes of the 1979–2008 period.

Camille was a relatively small hurricane whose forward speed of 13 kt approaching land was greater than usual. This gives it a somewhat low P_{rG} of around 9 h. With that in mind, it is also worth noting that Camille's track followed the primary axis of maximum OHC climatology more closely than did the 1979–2008 hurricanes studied (cf. Figs. 1 and 6). The actual OHC distribution during Camille, including whether the Loop Current or a warm-core eddy was located closer to the Gulf coast than normal (as contemplated by Shay 2009), is, unfortunately, unknown due to the lack of high-resolution satellite measurements at that time.

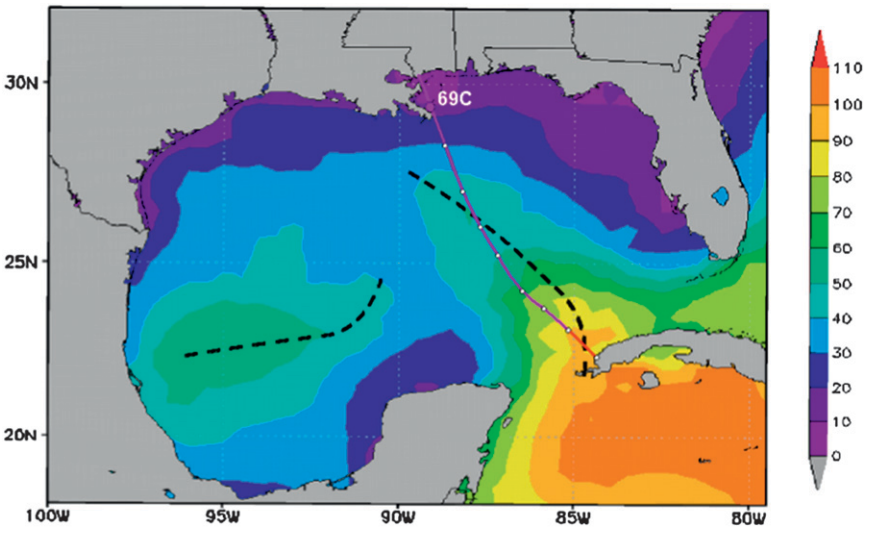

FIG. 6. Hurricane Camille (1969) track superposed on the average June–November OHC (kJ cm $^{-2}$). The purple track line is for the MH stage and the red line is for the NMH stage, with dots indicating 6-hourly positions. The overplotted and labeled purple dot indicates the location of maximum intensity. The dashed black lines show the primary axes of OHC maxima in the Gulf of Mexico.

5. Operational forecast application

Figure 3 showed that a least squares linear fit to the data explains much of the variance of the hurricane intensity change ending at landfall on the U.S. Gulf coast. In this section we explore whether the associated equations for wind speed yield predictive information of high enough accuracy to be of potential benefit to forecasters.

To measure the predictive value, we compared the forecasts made operationally by NHC to the landfall ''forecasts'' that would have come from applying the Gulf equations in Fig. 3. Because NHC does not explicitly forecast landfall information, we employed in some cases the special verification procedures described in the appendix.

The performance of the regression equations should be interpreted with knowledge of the competitive advantages and disadvantages arising from the computational process. The equations, for example, benefit from being applied to the dependent dataset. A disadvantage for the equations is that they are applied at the synoptic time. This is 3 h (was 4 h until 1992) before the time that the corresponding operational forecasts are issued. We also allowed the occasional NHC ''special'' advisory forecasts to supersede forecasts issued at the normal time in the verification process. Special advisory forecasts are anchored at the selected synoptic time, but they are issued up to 6 h after that time (as opposed to the usual 3- or 4-h offset). Further, we included in the verification the outliers discussed in section 4. This increased the errors reported below for the regression equations from what they would be in practice (i.e., in the future), because such systems can be identified a priori and be excluded from the application.

Figure 7 provides verifications for forecasts for the dependent sample from the equations in Fig. 3, and from SHIPS, LGEM, and the 5-day Statistical Hurricane Intensity Forecast model (SHIFOR5; Jarvinen and Neumann 1979; Knaff et al. 2003)—the three guidance models with a statistical basis available to NHC. For this comparison, we used today's versions of those models, which, to expand the number of cases, were rerun on cases as far back as possible (1983) in ''perfect prog'' mode. That is, we used analyses rather than forecast fields as input to SHIPS. In an analogous approach, we used the ''best track'' SHIFOR. Using the perfect prog and best-track frameworks puts the error statistics for those models in a relatively positive light.

Figure 7 indicates that mean errors for the forecasts from the regression equations for this rather small sample size ranged from about 25% to 65% less than the three operational models. These results suggest it might be possible to improve the SHIPS and LGEM levels of performance in the Gulf with additional or revised predictors.

Figure 8 presents a comparison of the equations and NHC forecasts. The statistically derived forecasts for hurricanes had smaller errors on average for each forecast period. The errors were at least 50% smaller at 36 and 48 h

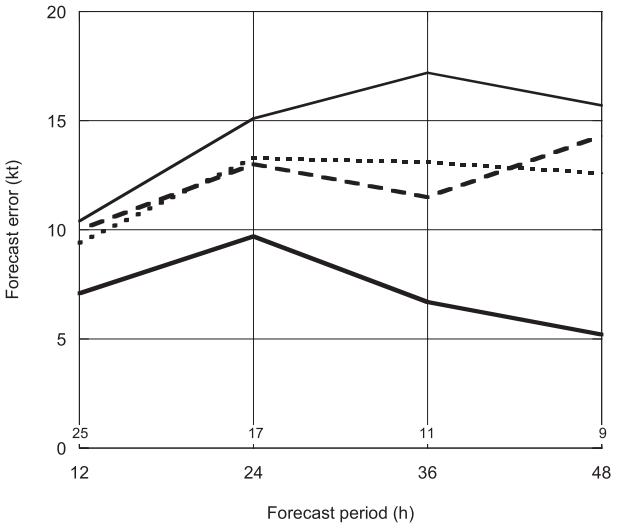

FIG. 7. Errors for SHIPS (long dashes), LGEM (short dashes), SHIFOR (light solid), and Gulf regression equations (heavy solid) from Fig. 3 for hurricanes (1983–2008). The number of verified cases for each forecast period is shown near the bottom of the panel.

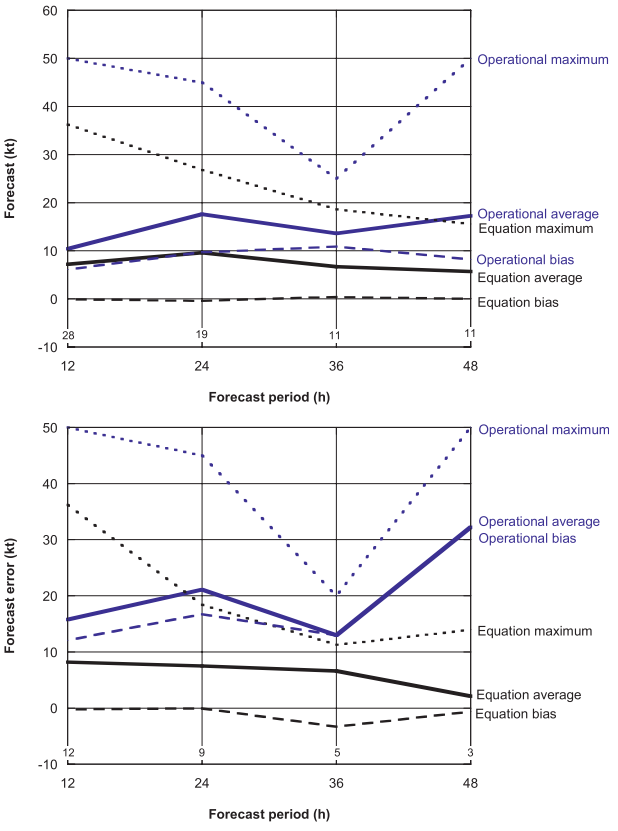

FIG. 8. Errors for operational forecasts (blue) and regression equations (black) from Fig. 3 for (top) all Gulf hurricanes and (bottom) a subset of Gulf MHs (1979–2008). The solid lines show the error average, while dotted lines indicate the maximum error, and dashed lines display the bias. The number of verified cases for each forecast period is shown near the bottom of each panel.

than the operational forecast for this still-small sample at those periods (Fig. 8, top panel). The maximum error and bias provided by the equations were also smaller than their operational counterparts for each period.

A larger disparity exists for the MH subset (Fig. 8, bottom panel). Figure 8 also shows that much of the relatively large errors in the operational forecasts for MH derive from a strong positive (''overforecast'') bias for them.

The large positive bias of NHC operational forecasts likely stems in part from forecasters' reluctance to predict hurricanes to weaken (much) upon approach to land unless the forecasters are quite certain—a level of confidence they rarely have at this juncture. Their reluctance arises out of the concern for incorrectly diminishing the perceived risk to the coastal population.

Because of its forecast limitations, NHC trains users of its forecasts to prepare for a hurricane one category higher than NHC predicts for landfall about a day in advance. Figure 8 shows, however, that NHC's errors—at least for Gulf hurricanes threatening U.S. landfall—are dominated by a positive bias that itself approaches one category for periods of 24–48 h, and reaches a category at all forecast periods for MHs.

The authors, two of whom have experience as NHC hurricane forecasters, do not suggest that the statistically based predictions provided by the new regression equations be used verbatim. It does appear, however, that these relationships could provide useful guidance for forecasters. Forecasters can have this guidance available at synoptic time, 3 h before they issue their forecasts.

6. Summary and future work

Tropical cyclones undergo systematic patterns of behavior when approaching the U.S. Gulf coast. Intensity change of tropical storms and tropical depressions is a strong function of time to landfall, but not of initial intensity. On average, they strengthen by about 7 kt per 12 h, except for slight change during their final 12 h over water. Hurricane intensity change prior to landfall correlates strongly with initial intensity at all lead times (forecast periods) through at least 48 h. The variance explained by a linear fit to the initial intensity increases from 50% at 12 h before landfall to 92% at 48 h for hurricanes.

On average, category 1–2 hurricanes strengthen and category 3–5 hurricanes weaken by landfall, such that they approach the threshold intensity for major hurricanes. This pattern of behavior can be partially explained by consideration of the maximum potential intensity modified by the environmental vertical wind shear and hurricane-induced sea surface temperature

reduction near the storm center as storms approach the northern Gulf coast. The high OHC regions in the central Gulf are very deep mixed layers, which prevent the SST reduction. The OHC in the northern Gulf is much lower, however; so, the SST reduction lowers the average shear-modified MPI to the category 2–3 threshold.

Simple predictive linear regression equations applied to the dependent database for hurricanes yield relatively small forecast (''hindcast'') errors compared to the model guidance and official forecasts. They could provide useful guidance during future hurricanes, and even more so when systems can be identified and excluded at forecast time as belonging to one of the following outlier groups:

- hurricanes with a Gulf roughening period P_{rG} (extent of 34-kt winds in the direction of motion divided by the forward speed) $\leq \sim 6$ h;
- hurricanes with a P_{rG} of more than about 20 h, for example, by stalling, moving slowly ($<\sim 5$ kt), or looping; and
- hurricanes with an inner-core convective structure significantly disrupted by a previous passage over land.

In the first exception, a hurricane stronger than predicted by the equations, with possible RI, should be anticipated. For the latter two cases, the system can be expected to be weaker than projected by the equations. For real-time and training considerations, applying operationally the concepts and equations presented here could require some forecaster and end-user recalibration, especially for MHs.

The findings in this study suggest a follow-on line of work. The importance of Gulf tropical cyclones and their intriguing behaviors imply that a regional version of the SHIPS program (perhaps limited to hurricanes) could be useful for forecasters. For example, the impacts of OHC in this study appear to be more important than in the basin-wide version of SHIPS, which could be due to the larger gradients of OHC in the Gulf compared to the Atlantic basin as a whole. A Gulf-region version of SHIPS, potentially using a higher-resolution OHC climatology (e.g., Shay and Brewster 2010), could provide a step toward meeting the HFIP goal of a 50% improvement in intensity forecast guidance, at least for this important subset of tropical cyclones that accounts for most U.S. hurricane landfalls.

Acknowledgments. LKS acknowledges support through grants from the NSF (ATM-04-44525), the NOAA Joint Hurricane Testbed Program (NA17RJ1226), and NASA's Hurricane Science Program (NNX09AC47G).

APPENDIX

Forecast Verification Procedures

a. ''Traditional'' verification procedure

When NHC's poststorm analysis indicates that landfall did not occur at a synoptic time, the NHC best-track intensity for the last synoptic time prior to landfall was chosen as the verifying intensity for the operational forecasts. For example, Hurricane Rita (2005) made landfall at 0740 UTC 24 September. For purposes of the comparison, the best-track information at 0600 UTC 24 September provided the verifying data to evaluate the operational forecasts. [In this case, Rita had 100 kt at both times (0600 and 0740 UTC) in NHC's best track, but in some cases the intensities at the two times differed.]

We then selected the NHC forecast to verify. Using the above example of Rita for the 24-h forecast, we first checked the NHC advisory originating 24 h earlier, from 0600 UTC 23 September (issued at 0900 UTC 23 September.) If the 24-h forecast position was offshore or at landfall, we could verify the forecast from that advisory. This occurred in 63 of the 69 cases.

b. Verification procedure when operational forecast position over land

In the remaining six cases (i.e., about 10% of the time), the forecast position was over land for the NHC advisory identified using the procedure described in the previous section of this appendix. This occurred when the forecast forward speed was too fast and/or the forecast track angled toward a closer land area, bringing the forecast hurricane position onshore prematurely. As an example for a 12-h forecast, Hurricane Dolly (2008) made landfall at 1820 UTC 23 July. The last synoptic time when the best-track position was offshore was 1800 UTC 23 July. The 12-h forecast position issued by the NHC in its advisory from 0600 UTC on that date (for 1800 UTC 23 July) was inland already. Verifying the corresponding 12-h 1800 UTC forecast intensity as a landfall or almost-landfall intensity would not be appropriate because NHC would have attempted in its forecast intensity to account for weakening it would expect to occur after landfall.

In this case, we stepped back to the previous advisory, the one made from synoptic time 0000 UTC 23 July. We selected that advisory to verify because the 12-h forecast position—for 1200 UTC 23 July—was still over water. This presented us with multiple verification options. Normally, the 12-h intensity forecast valid at 1200 UTC 23 July would be compared to the best-track intensity at that time. In our case, however, we were more interested in what the forecaster thought the intensity would be at or just before landfall than what the intensity would be at a particular

time. For this reason we employed a nontraditional verification procedure, in which the forecast intensity is taken from the last NHC 12-h forecast position offshore/landfall (1200 UTC 23 July), and the verifying intensity is taken from the best-track intensity at the last synoptic time offshore (1800 UTC 23 July), even though the times were 6 h apart. This procedure was required 3 times for the set of 28 12-h forecasts. In the remaining three cases, twice for the 19 24-h forecasts and once for the 11 48-h forecasts, the verifying and forecast times were 12 h apart.

c. *Verification procedure when subsequent advisory forecast position still offshore*

Of the 63 advisories identified using the procedure in section a of the appendix, 25 were not the last advisory to have an offshore or landfall position at the forecast horizon of interest. This occurred typically when the forecast forward speed in one or more of the advisories was too small. In these cases, we stepped forward in time until finding the last advisory for which the desired forecast duration had a forecast position still offshore or at landfall. Most of the adjustments were 6 h, with the few larger adjustments occurring at the largest forecast lead times.

Wilma (2005) provides an example for a 24-h forecast. Its landfall occurred at 1030 UTC 24 October. The last best-track synoptic time offshore was 0600 UTC 24 October. While NHC's 24-h forecast position issued in its advisory from 0600 UTC 23 October was still offshore, so was the 24-h forecast position in NHC's next advisory, issued from 1200 UTC 23 October. We used the 24-h forecast in the latter advisory and applied the nontraditional verification procedure described in the previous section of this appendix. That is, we verified the forecast intensity corresponding to the last 24-h offshore position (1200 UTC 24 October) using the intensity at the last offshore best-track position (0600 UTC 24 October), even though they were 6 h apart.

REFERENCES

Aberson, S. D., and M. DeMaria, 1994: Verification of a nested barotropic hurricane track forecast model (VICBAR). *Mon. Wea. Rev.,* **122,** 2804–2815.

Beven, J. L., II, and Coauthors, 2008: Atlantic hurricane season of 2005. *Mon. Wea. Rev.,* **136,** 1109–1173.

Blake, E. B., E. N. Rappaport, and C. W. Landsea, 2007: The deadliest, costliest, and most intense United States tropical cyclones from 1851 to 2006 (and other frequently requested hurricane facts). NWS Tech Memo. TPC-5, 43 pp.

Brennan, M. J., R. D. Knabb, M. Mainelli, and T. B. Kimberlain, 2009: Atlantic hurricane season of 2007. *Mon. Wea. Rev.,* **137,** 4061–4088.

Case, R. A., 1986: Atlantic hurricane season of 1985. *Mon. Wea. Rev.,* **114,** 1390–1405.

Chow, G. C., 1960: Tests of equality between sets of coefficients in two linear regressions. *Econometrica,* **28,** 591–605.

Cione, J. J., and E. W. Uhlhorn, 2003: Sea surface temperature variability in hurricanes: Implications with respect to intensity change. *Mon. Wea. Rev.,* **131,** 1783–1796.

DeMaria, M., 1996: The effects of vertical shear on tropical cyclone intensity. *J. Atmos. Sci.,* **53,** 2076–2087.

——, 2009: A simplified dynamical system for tropical cyclone intensity prediction. *Mon. Wea. Rev.,* **137,** 68–82.

——, and J. Kaplan, 1994: Sea surface temperature and the maximum intensity of Atlantic tropical cyclones. *J. Climate,* **7,** 1324–1334.

——, M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting,* **20,** 531–543.

Dvorak, V. F., 1984: Tropical cyclone intensity analysis using satellite data. NOAA Tech. Rep. NESDIS 11, 47 pp.

Emanuel, K. A., 1988: The maximum intensity of hurricanes. *J. Atmos. Sci.,* **45,** 1143–1153.

Franklin, J. L., cited 2009: 2008 National Hurricane Center forecast verification report. [Available online at http://www.nhc.noaa.gov/verification/pdfs/Verification_2008.pdf.]

——, M. L. Black, and K. Valed, 2003: GPS dropwindsonde wind profiles in hurricanes and their operational implications. *Wea. Forecasting,* **18,** 32–44.

——, R. J. Pasch, L. A. Avila, J. L. Beven II, M. B. Lawrence, S. R. Stewart, and E. S. Blake, 2006: Atlantic hurricane season of 2004. *Mon. Wea. Rev.,* **134,** 981–1025.

Hebert, P. J., 1978: Intensification criteria for tropical depressions of the western North Atlantic. *Mon. Wea. Rev.,* **106,** 831–840.

——, 1980: Atlantic hurricane season of 1979. *Mon. Wea. Rev.,* **108,** 973–990.

HFIP, cited 2009: Proposed framework for addressing the national hurricane research and forecast improvement initiatives. NOAA's Hurricane Forecast Improvement Project. [Available online at ftp://ftp.aoml.noaa.gov/hrd/pub/marks/HFIP/HFIP_Plan_071808.pdf.]

Holland, G. J., 1997: The maximum potential intensity of tropical cyclones. *J. Atmos. Sci.,* **54,** 2519–2541.

Jaimes, B., and L. K. Shay, 2009: Mixed layer cooling in Gulf of Mexico mesoscale eddies during Hurricanes Katrina and Rita. *Mon. Wea. Rev.,* **137,** 4188–4207.

Jarvinen, B. R., and C. J. Neumann, 1979: Statistical forecasts of tropical cyclone intensity for the North Atlantic basin. NOAA Tech. Memo. NWS NHC-10, 22 pp.

JHT, cited 2009: Joint Hurricane Testbed. Announcement of federal funding opportunity. [Available online at http://www.nhc.noaa.gov/jht/JHTFY09_Full_text_AFFO.pdf.]

Kaplan, J., and M. DeMaria, 2003: Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Wea. Forecasting,* **18,** 1093–1108.

——, ——, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and east Pacific basins. *Wea. Forecasting,* **25,** 220–241.

Knaff, J. A., M. DeMaria, B. Sampson, and J. M. Gross, 2003: Statistical, 5-day tropical cyclone intensity forecasts derived from climatology and persistence. *Wea. Forecasting,* **18,** 80–92.

——, C. R. Sampson, M. DeMaria, T. P. Marchok, J. M. Gross, and C. J. McAdie, 2007: Statistical tropical cyclone wind radii prediction using climatology and persistence. *Wea. Forecasting,* **22,** 781–791.

Laurmann, J. A., and W. L. Gates, 1977: Statistical considerations in the evaluation of climatic experiments with atmospheric general circulation models. *J. Atmos. Sci.,* **34,** 1187–1199.

Mainelli, M., M. DeMaria, L. K. Shay, and G. Goni, 2008: Application of oceanic heat content estimation to operational

forecasting of recent Atlantic category 5 hurricanes. *Wea. Forecasting,* **23,** 3–16.

Mainelli-Huber, M., 2000: On the role of the upper ocean in tropical cyclone intensity change. M.S. thesis, Meteorology and Physical Oceanography Department, University of Miami, Miami, FL, 73 pp.

Marks, F. M., and Coauthors, 1998: Landfalling tropical cyclones: Forecast problems and associated research opportunities: Report of the Fifth Prospectus Development Team to the U.S. Weather Research Program. *Bull. Amer. Meteor. Soc.,* **79,** 305–323.

McAdie, C. J., C. W. Landsea, C. J. Neumann, J. E. David, E. S. Blake, and G. R. Hammer, 2009: Tropical cyclones of the North Atlantic Ocean, 1851-2006 (with 2007 and 2008 track maps included). Historical Climatology Series, No. 6-2, NOAA/NWS/NESDIS, 238 pp.

Nolan, D. S., Y. Moon, and D. P. Stern, 2007: Tropical cyclone intensification from asymmetric convection: Energetics and efficiency. *J. Atmos. Sci.,* **64,** 3377–3405.

NWS, cited 2009: NWS directives system. [Available online at http://www.nws.noaa.gov/directives/sym/pd01006001curr.pdf.]

Pasch, R. J., M. B. Lawrence, L. A. Avila, J. L. Beven, J. L. Franklin, and S. R. Stewart, 2004: Atlantic hurricane season of 2002. *Mon. Wea. Rev.,* **132,** 1829–1859.

Pielke, R. A., J. Gratz, C. W. Landsea, D. Collins, M. A. Saunders, and R. Musulin, 2008: Normalized hurricane damage in the United States: 1900-2005. *Nat. Hazards Rev.,* **9** (1), 29–42.

Rappaport, E. N., and J. Fernandez-Partagas, 1995: The deadliest Atlantic tropical cyclones, 1492-1994. NOAA Tech. Memo. NWS NHC-47, Coral Gables, FL, 41 pp.

——, and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting,* **24,** 395–419.

Schott, T. J., and Coauthors, cited 2010: The Saffir–Simpson hurricane wind scale. NOAA/National Weather Service. [Available online at http://www.nhc.noaa.gov/sshws.shtml.]

Shay, L. K., 2009: Upper ocean structure: Response to strong forcing events. *Encyclopedia of Ocean Sciences.* 2nd ed. J. Steele et al., Eds., Elsevier, 4619–4637, doi:10.1016/B978-012374473-9.00628-7.

——, and E. W. Uhlhorn, 2008: Loop Current response to Hurricanes Isidore and Lili. *Mon. Wea. Rev.,* **136,** 3248–3274.

——, and J. Brewster, 2010: Oceanic heat content variability in the eastern Pacific Ocean for hurricane intensity forecasting. *Mon. Wea. Rev.,* **138,** 2110–2131.

——, G. J. Goni, and P. G. Black, 2000: Effects of a warm oceanic feature on Hurricane Opal. *Mon. Wea. Rev.,* **128,** 1366–1383.

Sheets, R. C., 1990: The National Hurricane Center—Past, present and future. *Wea. Forecasting,* **5,** 185–232.

Siegel, S. L., 1956: *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, 312 pp.

Simpson, R. H., 1974: The hurricane disaster potential scale. *Weatherwise,* **27,** 169–186.

Vickery, P. J., and D. Wadhera, 2008: Statistical models of Holland pressure profile parameter and radius of maximum winds of hurricanes from flight-level pressure and H*Wind data. *J. Appl. Meteor. Climatol.,* **47,** 2497–2517.

Willoughby, H. E., J. A. Clos, and M. G. Shoreibah, 1982: Concentric eye walls, secondary wind maxima, and the evolution of the hurricane vortex. *J. Atmos. Sci.,* **39,** 395–411.