On Estimates of Historical North Atlantic Tropical Cyclone Activity*

GABRIEL A. VECCHI AND THOMAS R. KNUTSON

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

ABSTRACT

In this study, an estimate of the expected number of Atlantic tropical cyclones (TCs) that were missed by the observing system in the presatellite era (between 1878 and 1965) is developed. The significance of trends in both number and duration since 1878 is assessed and these results are related to estimated changes in sea surface temperature (SST) over the "main development region" ("MDR"). The sensitivity of the estimate of missed TCs to underlying assumptions is examined. According to the base case adjustment used in this study, the annual number of TCs has exhibited multidecadal variability that has strongly covaried with multidecadal variations in MDR SST, as has been noted previously. However, the linear trend in TC counts (1878–2006) is notably smaller than the linear trend in MDR SST, when both time series are normalized to have the same variance in their 5-yr running mean series. Using the base case adjustment for missed TCs leads to an 1878–2006 trend in the number of TCs that is weakly positive, though not statistically significant, with $p \sim 0.2$. The estimated trend for 1900–2006 is highly significant ($+\sim 4.2$ storms century^{-1}) according to the results of this study. The 1900–2006 trend is strongly influenced by a minimum in 1910–30, perhaps artificially enhancing significance, whereas the 1878–2006 trend depends critically on high values in the late 1800s, where uncertainties are larger than during the 1900s. The trend in average TC duration (1878–2006) is negative and highly significant. Thus, the evidence for a significant increase in Atlantic storm activity over the most recent 125 yr is mixed, even though MDR SST has warmed significantly. The decreasing duration result is unexpected and merits additional exploration; duration statistics are more uncertain than those of storm counts. As TC formation, development, and track depend on a number of environmental factors, of which regional SST is only one, much work remains to be done to clarify the relationship between anthropogenic climate warming, the large-scale tropical environment, and Atlantic TC activity.

1. Introduction

There is currently disagreement within the hurricane/climate community on whether anthropogenic forcing (greenhouse gases, aerosols, ozone depletion, etc.) has caused an increase in Atlantic tropical storm or hurricane frequency. Santer et al. (2006) and Knutson et al. (2006) have presented model-based evidence that the twentieth-century rise in tropical Atlantic SSTs is outside the range expected from internal climate variabil-

ity, with a likely discernible warming from anthropogenic forcing. Mann and Emanuel (2006) and Holland and Webster (2007) argue that the close association of tropical Atlantic SSTs with the observed record of basin-wide tropical cyclone (TC) counts from the late 1800s or ~1900 to the present is evidence for a strong emerging anthropogenic signal on Atlantic TC activity. On the other hand, Landsea (2007) has argued that the existing Atlantic TC count database is seriously deficient, and that when adjusted for likely missing storms, no significant trend is evident, consistent with the analysis of Solow and Moore (2002) for 1900–98 basin-wide hurricane frequency. Both of these studies utilized storm landfalling records to infer basin-wide behavior in earlier periods. Recently, Chang and Guo (2007) used historical ship-track data and satellite-era storm-track locations to estimate the number of missing TCs in the 1900–65 period and found that, while it was likely that some storms were missed by the observing network
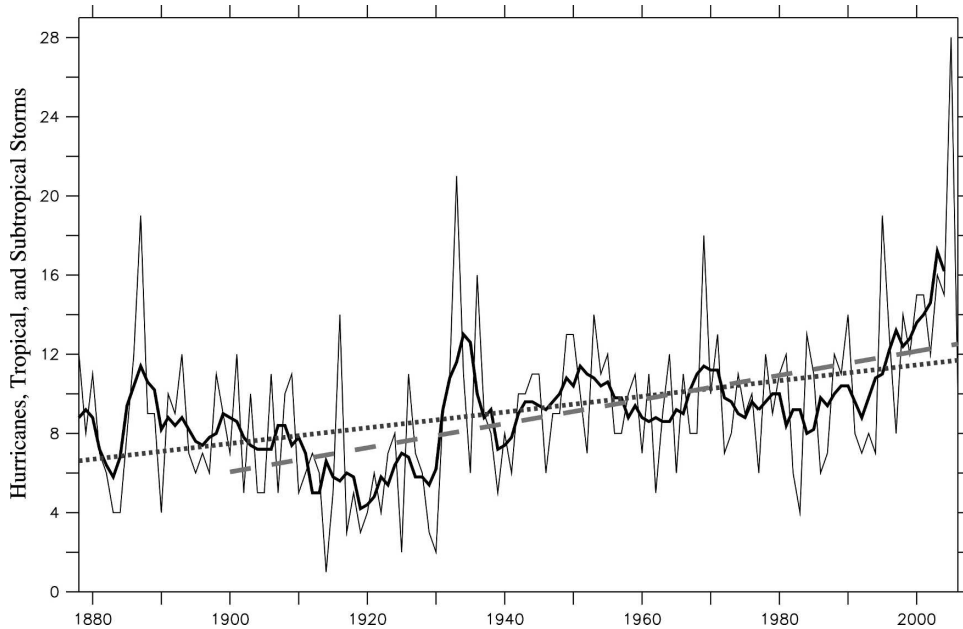
FIG. 1. Time series of unadjusted HURDAT Atlantic basin TC counts over the period 1878–2006. Black line shows the annual count of tropical and subtropical storms, and hurricanes in the HURDAT database. Dashed lines indicate the linear least squares trends computed over the periods 1878–2006 and 1900–2006.

prior to basin-wide monitoring by satellite, there was still an increase in TC counts over the period 1900–2006.

Modeling evidence indicates that anthropogenic greenhouse gas forcing may result in stronger North Atlantic TCs in a future climate (e.g., Shen et al. 2000; Knutson and Tuleya 2004; Bengtsson et al. 2007). However, analysis of climate model projections for the twenty-first century indicates that, in addition to making the tropical environment generally more favorable to TCs by warming tropical SSTs, increasing greenhouse gases may influence other factors (such as vertical wind shear, midtropospheric relative humidity, and atmospheric stability) in a way to make the environment less favorable to TCs in parts of the tropical Atlantic (e.g., Shen et al. 2000; Vecchi and Soden 2007a,c). Meanwhile, individual climate model projections of the response of Atlantic TC counts to anthropogenic forcing are mixed, with some studies indicating an increase (e.g., Oouchi et al. 2006), others a decrease (e.g., Bengtsson et al. 2007; Gualdi et al. 2008; Knutson et al. 2008), and the response in some depending on the details of the large-scale response of the climate system to increased CO_2 [e.g., Emanuel et al. (2008); the ensemble mean response of which is for an increase in storm counts]. Existing modeling work has largely focused on projections of future climate, characterized by large tropical SST changes (2°–4°C), rather than the historical period examined in this paper, which has a

more modest tropical SST increase on the order of 0.5°C (e.g., Knutson et al. 2006; Solomon et al. 2007).

Figure 1 shows a time series of the National Oceanic and Atmospheric Administration's (NOAA's) Atlantic basin hurricane database (HURDAT; see section 2a below) TC count in the Atlantic basin from 1878, when the U.S. Signal Service began tracing all West Indian hurricanes (Fernández-Partagás and Diaz 1996), to 2006. The Atlantic TC count record from HURDAT in the Atlantic exhibits variability on many time scales, as well as a clear long-term rise. There is prominent interannual variability—partly associated with El Niño–La Niña events in the tropical Pacific—as well as lower-frequency decadal to multidecadal variations: for example, the period between 1910 and 1930 was uniquely quiet, while the period starting in the mid-1990s has had activity unprecedented in this record. Using a least squares linear trend as our statistic of long-term change (a choice discussed further in section 4), the unadjusted HURDAT database exhibits a statistically significant (at $p = 0.05$) increase in TC counts, both between 1900–2006 and 1878–2006. See section 2 for a description of the statistical significance tests used. The slope of the linear trend from 1878 to 2006 represents an increase in annual storm counts of about 60% century^{-1}. Interpreted as a long-term increase in TC frequency in the Atlantic, this increase in HURDAT storm counts is quite large.
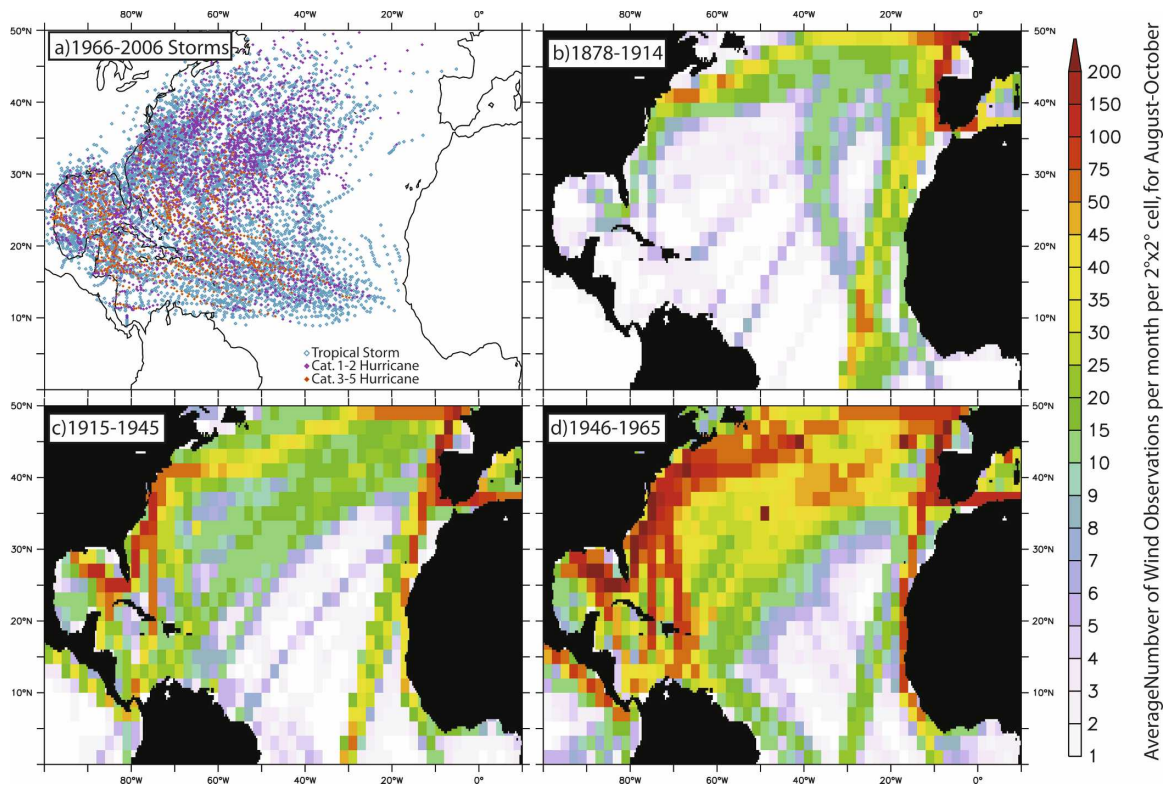
FIG. 2. (a) The TC tracks from HURDAT over the 1966–2006 satellite era. Blue symbols indicate positions of TCs with tropical storm intensity (17–33 m s^{-1}), violet symbols indicate category 1 and 2 hurricane intensities (33–55 m s^{-1}), and red symbols indicate intense hurricane positions (winds >55 m s^{-1}). Densities of August–October wind speed observations from ships in the ICOADS dataset, averaged for three different periods: (b) before the opening of the Panama Canal, (c) between the opening of the Panama Canal and the end of WWII, and (d) from the end of WWII to the satellite era. Units are observations per 2° × 2° cell month^{-1}.

However, there have been changes to the methodology used to observe TCs over the period 1878–2006. Before 1944, the main method for identifying TCs was by records of landfalling storms or by records of ships at sea. Between 1944 and 1965, there were aircraft reconnaissance flights complementing observations by ships at sea, although aircraft coverage did not extend over the entire basin. Basin-wide monitoring via satellite began in 1966 (Landsea 2007). Even during the "ship observation era" (pre-1944) there were significant modifications to the preferred tracks of ships (e.g., Fig. 2). Before the opening of the Panama Canal in 1914, most of the recorded ship traffic tended to be concentrated in the northern and eastern tropical Atlantic and near the east coast of North America (Fig. 2b), leaving a conspicuous "hole" in many regions of frequent TCs (Fig. 2a). After 1914, the ship-recorded track density in the Gulf of Mexico, Caribbean Sea, and western tropical Atlantic increased dramatically (Fig. 2c). Following World War II (WWII) the recorded ship density increased further (Fig. 2d). In addition, both disruptions to shipping and

missing records from ships during both World Wars resulted in minima of data availability between 1914–18 and 1939–45. Thus, it is plausible that some of the secular increase in TC counts recorded in HURDAT may have resulted from changes in observational practices.

Given the central role that historical datasets of TC activity and data homogeneity questions play in our understanding of the connection between climate and hurricanes, we here estimate a correction to TC counts in the presatellite era using ship-track data from the presatellite era and TC locations from the satellite era, and explore long-term changes in TC activity measures in the tropical Atlantic. In section 2 we describe the datasets used (2a), the TC activity measures we evaluate (2b), the statistical significance tests we apply (2c), and our method to estimate missing tropical storms (2d). In section 3 we describe the principal results of this paper, focusing on long-term changes to TC activity and the impact of our storm count adjustment. Finally, in section 4 we offer some discussion of our results and discuss possible future work.

2. Data and methods

a. Datasets used

As our historical TC track data, we use the National Hurricane Center (NHC) HURDAT "best track" dataset. Data are archived 6 hourly (at 0000, 0600, 1200, and 1800 UTC) and include reports of storm position and maximum winds from 1851 to 2006 (Jarvinen et al. 1984; Landsea et al. 2004). We focus on the period 1878–2006, and only consider storms while they are in either their "tropical" or "subtropical" stages (as designated in the HURDAT dataset). To compute the distance of a particular storm to a ship observation or land point, the 6-hourly HURDAT best-track data are linearly interpolated to a 2-hourly grid, in order that storms are less likely to "hop" over an observation in the discrete analysis method used here (since storms can move many tens of kilometers in a 6-h step).

Not only have the methodology and distribution of observations changed since the late-nineteenth century, but some of the recording practices in HURDAT have also changed with time. Of relevance to the study of TC activity is the change in the number of "tropical depression days" recorded for each storm (tropical depression days are those for which a TC has maximum winds below gale force, 17 m s^{-1}). In the presatellite-era records in HURDAT, it is quite common for tropical storms to have no record of their existence as tropical depressions, while in the satellite era practically all tropical storm records include a substantial number of days as a tropical depression (Fig. 3a). After 1966, most TC records have at least 30% of their recorded lifetime as tropical depressions, with many spending most of their recorded lifetime as a tropical depression. This change in fraction of "tropical depression days" has resulted from changes in the identification and recording practices used to generate HURDAT (e.g., Landsea et al. 2004). Thus, assessment of historical changes in the duration of TCs, or of quantities that are integrated through the lifetime of a TC [such as accumulated cyclone energy (ACE) and power dissipation index (PDI)], must take into account this artificial increase in recorded storm lifetimes after the advent of satellites (i.e., by excluding tropical depression periods from the analysis). The impact of this bias is likely to be small for PDI since it is the sum of the cube of the wind speed, to which the depression stage of a storm contributes little.

We use ship observation positions from the International Comprehensive Ocean–Atmosphere Dataset (ICOADS; Worley et al. 2005) version 2.3.2a (data available online at http://icoads.noaa.gov/products.html). This dataset includes the ship position and date of observation from 1754 to 2005. For this analysis all

ships are taken to be perfect measurement platforms and unable to alter their course in response to the presence of a nearby TC. To define coastlines, we use the Smith and Sandwell 2-min topography dataset (available online at http://www.ngdc.noaa.gov/mgg/bathymetry/predicted/explore.html) and assume land points to be "perfect observers."

We use three historical SST reconstructions: the Kaplan (Kaplan et al. 1998), Hadley Centre Sea Ice and Sea Surface Temperature (HadISST; Rayner et al. 2003), and NOAA-Extended (Smith and Reynolds 2004). We do this because each of the SST reconstructions exhibits distinct long-term trends of tropical SST over the instrumental record (e.g., Vecchi and Soden 2007b; Vecchi et al. 2008). These three products have distinct techniques, involving different statistical techniques, corrections to the raw data, and slightly different data sources. Although each product shows a clear overall tendency for tropical warming since the 1880s, there are discrepancies in the spatial structure of the changes in all three tropical basins. Until the disagreement between the various SST records is resolved, we believe it is prudent to explore multiple datasets.

b. Tropical storm activity measures

We explore three different, but related, basin-wide measures of Atlantic TC activity: annual TC counts, annual tropical storm days, and average TC duration. The annual tropical storm count, or N_{TS}, is the number of systems each year than reach gale force winds or higher (17 m s^{-1}); the annual tropical storm days, or D, is the sum over all TCs present in a year of the total days each system's maximum winds exceed gale force; and the average TC duration, or d, is the average number of days each TC has maximum winds exceeding 17 m s^{-1}, or D/N_{TS}. The accuracy of each of these measures depends on the detectability of historical TCs, and the last two measures also depend on the ability of the life cycle of a storm to be accurately described by the observations.

In addition to these basin-wide measures, we explore a spatially dependent measure of TC activity: storm-track density. Storm-track density is defined, on a 2.5° × 2.5° latitude–longitude grid, as the total number of days that there is a TC record inside each grid cell, based on the HURDAT "best track" latitude and longitude data. To compute TC density, we exclude periods when storm intensities were less than gale force (17 m s^{-1}), as discussed in section 2a.

c. Statistical significance tests

We here use the terminology that a particular statistic is "significant" if it is estimated to be distinguishable
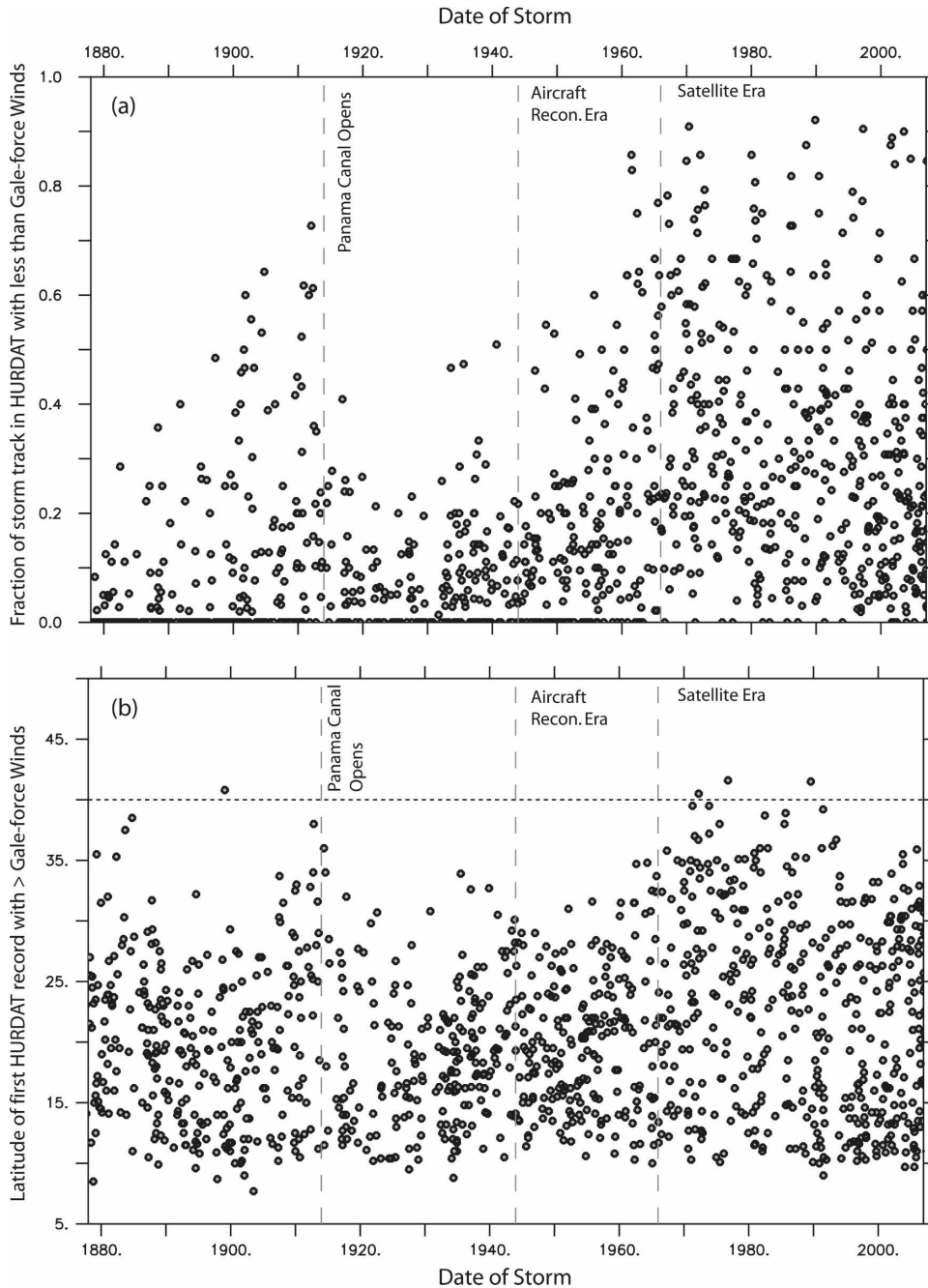
Date of Storm

FIG. 3. (top) Percentage of the recorded lifetime of a TC in HURDAT for which its maximum winds are below 17 m s^{-1}. (bottom) Latitude north at which the first TC maximum winds above 17 m s^{-1} are recorded in HURDAT. Each symbol represents a TC in HURDAT. Vertical lines indicate significant changes to the observing practices in the Atlantic basin. Notice that before the satellite era it is common for the recorded history of TCs not to include any time as tropical depressions, and most TCs have their first recorded winds greater than 17 m s^{-1} south of 30°N.

from zero at $p = 0.05$ using a two-sided test, and we will list the estimated p values for nonsignificant statistics explicitly. Three different statistical testing methods for trend have been applied, all addressing the temporal correlation in the data.

1) t TEST

The t test on the trend slope uses the linear trends computed using ordinary least squares regression. The lag-one autocorrelation coefficient, r_1, of the residual

time series (after removing the trend) is used to adjust the temporal degrees of freedom for the effects of persistence in the data, using the following formula: $DOF' = N[(1 - r_1)/(1 + r_1)]$, where N is the sample size, as discussed, for example, in Wilks (2006, p. 144). This is used in the formula in computing the t statistic and identifying critical t values. For storm count and duration series, in order to address concerns over the skewness of the distributions (e.g., storm counts have a lower bound of zero and no a priori upper bound), the standard t tests were performed on time series of the square root of the annual series.

2) t TEST ON RANKS

A second test for significance was based on the Student's t test, but applied to the ranks of the time series rather than their numerical values. For storm statistics, this test was used as an alternative to the square root transformation method.

3) BOOTSTRAP TEST

The third method uses a bootstrap resampling (with replacement) technique, in which synthetic time series are constructed from subsegments of the original time series. The trend analysis was performed on large sample ($n = 10^4$) of such synthetic series to determine how unusual the magnitude of the linear trend from the original series was relative to trends in the synthetic series (i.e., the percentile rank of the original series trend value within the cumulative distribution of randomly generated trends). The use of segments rather than randomly selected individual samples allows us to retain aspects of the persistence of the full time series, accounting for persistence in a complementary manner to the first two tests. Wilks (2006, p. 170) provides some guidance on the selection of the segment length for this method, and we report a typical p value based on averaging results from tests using segment lengths ranging from three to eight.

d. Estimate of historical storm count adjustment

We assess the impact of changing observational practices on measures of TC activity prior to the satellite era, using historical ship tracks from the presatellite era combined with storm-track information from the satellite era. For our analysis we must define a proximity rule defining when TCs are "detected" in the resampling experiments. We have used the statistics of the observed radius of 17 m s^{-1} winds (R17) compiled by Kimball and Mulekar (2004, hereafter KM04) to develop a statistical model for R17 to be used in our analysis. An intensity-dependent model of R17 as a log-

normal distribution adequately represents the statistics of R17 described in KM04, when applied to the 1966–2006 tropical storm record (see Fig. 4), and its functional form is

$$R17(u_{\max}, \xi)$$

$$= \begin{cases} 0 & u_{\max} < 17 \text{ m s}^{-1} \\ 90(e^{(\xi/1.3)}) + 70 & 17 \text{ m s}^{-1} \leq u_{\max} < 33 \text{ m s}^{-1} \\ 90(e^{(\xi/1.3)}) + 150 & 33 \text{ m s}^{-1} \leq u_{\max} < 50 \text{ m s}^{-1} \\ 90(e^{(\xi/1.3)}) + 170 & 50 \text{ m s}^{-1} \leq u_{\max} \end{cases}$$

where ξ is a normally distributed random number with a mean of zero and variance of 1 [i.e., $\xi = N(0, 1)$]. In this formulation the units of R17 are kilometers, and is forced to have a maximum at 700 km by making the R17 for all radii calculated larger than 700 km to be 1400-km R17. With this parameterization, R17 is larger for major hurricanes (categories 3–5) than minor hurricanes, and for minor hurricanes (categories 1 and 2) than tropical storms, in agreement with the statistics of KM04. For our analysis, since we assume that TCs are radially symmetric, we adjust the R17 values parameterized above by a factor of 0.85 to convert from maximum extents to mean extent (J. Knaff and M. DeMaria 2007, personal communication).

Using the 2-hourly storm-track data from the satellite era (1966–2006), for each 2-h segment we compute the two closest ICOADS ship-track positions on a given calendar day for each presatellite-era year—whether there is a wind observation in ICOADS or not, and making sure that the two closest ship locations are independent (i.e., we make sure that the two observations are not the same observation on the same day identified twice). We then repeat this process, but shift the storm-track calendar dates forward and backward in 5-day intervals from −30 to +30 days. This gives 13 samples for each storm observation in the satellite era, for each presatellite-era year, for each storm radius seed.

Then, randomizing the radius seed, ξ, in the R17 model above 50 times (so $13 \times 50 = 650$ iterations per storm per presatellite-era year), we compare the ship positions for each presatellite-era year with the positions of each satellite-era TC; we do this over the 41 satellite-era years, yielding $650 \times 41 = 26\,650$ sampled positions for each presatellite-era year. We then estimate the adjustment to the TC count for each presatellite-era year as the average number of tropical storms "missed" in each sampling year. This will be referred to as our additive adjustment. We also compute the probability that a particular satellite-era TC would have been missed (p_m) by each presatellite year. Finally, for each presatellite year we estimate a method uncertainty
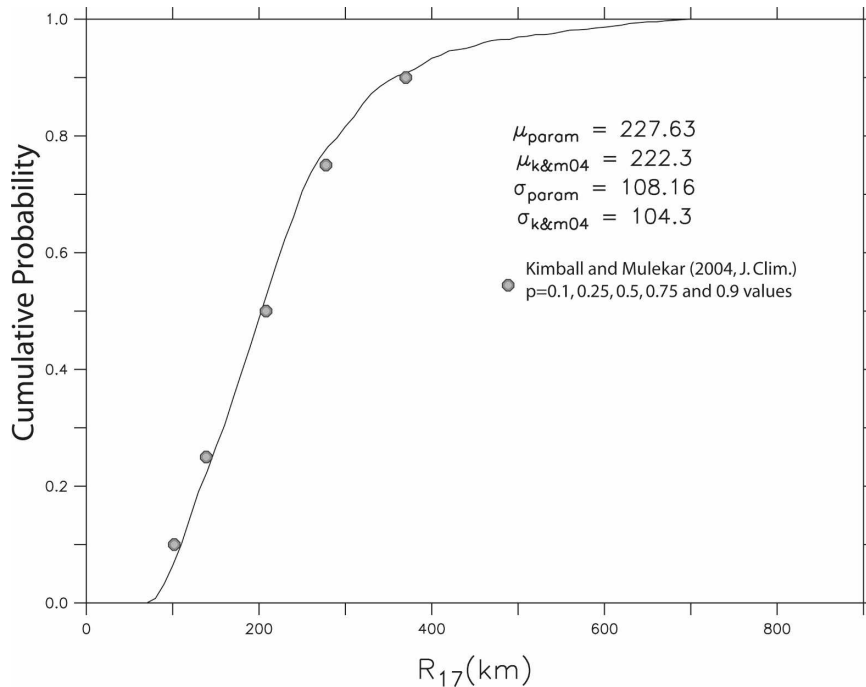
FIG. 4. Cumulative distribution function of the statistical fit to the radius of gale force winds (R17) based on the observational analysis of Kimball and Mulekar (2004). Dots show the statistics from Kimball and Mulekar (2004) for $p = 0.1, 0.25, 0.5, 0.75,$ and 0.9. The R17 model is used to assess the proximity of satellite-era TCs to historical ship positions.

for our additive adjustment using the cumulative distribution function of annual "missed" storm counts across the 26 650 samples. This adjustment assumes that the number of storms that is likely to have been missed is a function of the observing system present, and that the probability of one of the (relatively unusual) storms that is able to "slip" through the observing system is stationary and represented by the storms from the period 1966–2006.

An alternative adjustment, referred to as the multiplicative adjustment, assumes that the number of storms "missed" in each presatellite-era year is proportional to the number of storms in HURDAT for that year. The scaling is computed based on the ratio of storms missed to those "seen" across the 26 650 samples. That is, for each presatellite-era year (i),

$$A_i = N_i' R_i,$$

where A_i is the adjusted storm count number, N_i' is the HURDAT-recorded number of storms, and R_i is the scaling factor

$$R_i = 1 + \frac{\sum_{j=1}^{26\,650} M_{i,j}}{\sum_{j=1}^{26\,650} N_j - M_{i,j}},$$

where $M_{i,j}$ is the number of missed storms in each of the 26 650 samples for the particular presatellite-era year and N_j is the actual number of satellite-era storms in each of the samples. We consider this adjustment to be less plausible than the additive adjustment because a resampling of the 1966–2006 storms, using the observing systems of 1878–1965, does not indicate the positive correlation between storms missed and storms "detected" that this adjustment implies.

A storm is considered to be detected if a land point is within a radius 0.85R17 or if there are two independent occurrences of ships approaching within 0.85R17. A ship/land encounter with a storm track must occur equatorward of 40°N in order for a tropical storm "detection" to occur, since the first latitude at which each TC in HURDAT reached gale force was poleward of 40° only once before 1966 (see Fig. 3b). In deciding whether to include a new candidate storm in the official HURDAT database, the HURDAT team used as criteria two independent ship observations of gale force winds (or pressure equivalent): evidence of a closed circulation and evidence of nonfrontal character (Landsea et al. 2007). We have not attempted to incorporate the latter two criteria in our detection scheme.

Here, we reiterate some of our key assumptions, along with a rough assessment of the expected errors due to these assumptions.

1) The method assumes that all land points have been perfect storm detectors over the period under consideration. This presumes that all land is populated at sufficient density and with sufficient technological development and reporting capabilities to record and report all TCs that pass over land. If sparsely settled land allowed landfalling storms to go undetected or unreported in reality, the adjustment would be biased low. This assumption is likely to be too strong, since as recently as Landsea et al. (2007) four new landfalling storms have been discovered in the 10 yr of 1911–20. Methods of estimating the extent to which landfalling storms were likely to have been missed in certain regions in the past should be developed.

2) A second key assumption is that sufficient relevant ship tracks are contained in the ship-track database in the ICOADS. If there were in fact other ships not in ICOADS that would have reported TCs, the adjustment for missing storms would be biased high. However, the inclusion of such additional ship data might result in the discovery of new TCs for inclusion in HURDAT (which would raise the unadjusted storm count). For example, additional ship log data have recently been digitized, and will be included in forthcoming versions of ICOADS (S. Woodruff 2007, personal communication). As these data become available, they should be used both to identify historical storms and to recompute the expected storm count adjustment.

3) An assumption related to assumption 2 above is that all of the relevant storms that would be detectable from the ICOADS have been included in the HURDAT dataset. Errors in this assumption would tend to bias the adjustment low. In fact, a reanalysis of the historical ship-track data and storm-track data is currently under way (e.g., Landsea et al. 2007), and during the preparation of this manuscript 13 additional storms were identified in the period 1911–20 (five of these storms are included in our analysis). If the rate of new storm identification of Landsea et al. (2007) is representative of that for other periods in the early twentieth century, one may expect around an additional storm per year to be identified as other periods are reanalyzed. However, the correction computed here would remain applicable to a revised storm database provided that data comparable to ICOADS have been used as its basis.

4) TCs are assumed to be radially symmetric, as detailed information about storm structure is unavailable for most of our analysis period. Errors in this assumption will likely be random—rather than systematic—and presumably not result in a significant bias.

5) Ships and land are always able to perfectly measure the wind. We expect the largest errors in ship and land sampling to be random, but any systematic (under-) overestimate of wind speed would lead to an (under-) overestimate of TC activity.

6) We assume that the ships' crews did not attempt to avoid, or were unable to avoid, chance encounters with TCs (at least to gale force strength). Errors in this assumption would lead to an underestimate of the adjustment.

7) We assume that modern-day TCs are representative of the TCs in the past, in terms of their number and location. This assumption would tend to make the adjustment err against any real trend in TC counts. If the modern era is in fact more active than the early period, the storm adjustment will be biased high. Alternatively, if a negative trend in storm counts existed, the adjustment would be biased low.

8) We assume that if a storm is "detectable" through observations within the radius of gale force winds (see above), then there will be sufficient ancillary observations to identify the system as a closed circulation and nonfrontal in character, which are the other criteria necessary for a system to be included in HURDAT (Landsea et al. 2007). Errors in this assumption would lead to an underestimate of the adjustment.

9) We assume that single storms have not been counted as two separate storms in the HURDAT database. If double counting occurred, it would tend to bias our adjustment high.

Overall, errors in most of the assumptions would tend to lead to either random errors (assumptions 4 and 5) or an underestimate of the adjustment (assumptions 1, 3, 6, and 8). The sign of the error produced by assumption 2 is not clear, and that of assumption 7 would be to oppose any real trend. Assumption 9 could lead to an overestimate of the adjustment, but no evidence has been published that indicates that this error is substantial in size (and the duration statistics presented below argue against it being large). On this basis, while the relative and cumulative impacts of errors in these assumptions are difficult to quantify, we speculate that our adjustment is more likely to be an underestimate than an overestimate of the true number of missing TCs in HURDAT.

e. Adjustment to other activity measures

In addition to adjusting annual TC counts, we can use our estimate of missed TCs to adjust other annually aggregated statistics of tropical storm activity. To do this we use the probability a storm was missed in a

particular year (p_m) to weight the value of the statistic to be aggregated (e.g., total storm days, or storm density). We then add to the value computed from the unadjusted data, the probability-weighed value of the statistic, averaged over the satellite-era years. That is, for any annually aggregated statistic ($\Phi \equiv \Sigma_s \phi$) computed over the tropical storms (s) of a particular pre-satellite-era year, then $\tilde{\Phi}$ (the adjusted estimate of Φ) is

$$\tilde{\Phi} = \Phi + \frac{\sum_{i=1966}^{2006} \sum_{s_i} p_m(s_i) \cdot \phi(s_i)}{41},$$

where s_i are the storms of each satellite-era year, $p_m(s_i)$ is the probability the storm is "missed" in the given presatellite-era year, and $\phi(s_i)$ is the statistic to be aggregated from the particular storm.

To estimate the adjusted average storm duration ($\tilde{d}$), we divide the adjusted total storm days per year ($\tilde{D}$) by the adjusted tropical storm count, $\tilde{N}_{TS}$. Similarly, to compute the time-smoothed value of d (e.g., in Fig. 7, below), rather than smooth the time series of d, we compute the ratio of time-smoothed values of D and N_{TS}.

3. Results

a. Tropical storm counts

Figure 5 summarizes the estimated ship-track-based adjustment to historical storm counts, based on the assumptions and methods described in section 2c. The adjustment gradually increases going back in time, from about ¼ storm yr^{-1} in the 1950s and 1960s to about 3.4 storms yr^{-1} by around 1880. Local maxima are also apparent around the World War periods. Aside from these local maxima, the additive adjustment changes gradually, without suggesting a natural "cutoff date." The effect of the multiplicative and additive adjustments on the trends is quite similar, though the multiplicative adjustment exhibits substantially larger interannual variability than does the additive adjustment.

We compare our adjustments to three recently derived adjustments in Fig. 5b. The 10-yr-averaged amplitude of our adjustments is similar to that derived independently, with a similar methodology, by Chang and Guo (2007), when compared over the various 10-yr intervals for which Chang and Guo (2007) report values. However, Chang and Guo (2007) do not provide estimates for missed storms in the nineteenth century, WWI, or WWII—the periods where we find the largest adjustment. Overall, our adjustments are more modest than that of Landsea (2007), with the Landsea adjustment outside the uncertainty estimates of our adjust-

ment for a substantial part of the 1900–2006 period. The temporal character of our adjustments is also different from the adjustment proposed by Landsea (2007) who inferred that 2.2 storms yr^{-1} were missing for each years from 1900 until 1965 (and 3.2 storms yr^{-1} relative to 2003–06). It should be noted that we attempt to estimate the effect on storm counts of a limited set of sources of observational uncertainty, while Landsea (2007) infers a storm undercount based on the characteristics of the HURDAT database without directly addressing the sources of uncertainty. Also, the character of our adjustment time series is different from the central estimate of Mann et al. (2007), though their estimate is within the method uncertainty estimate for ours. Over much of the twentieth century, there is general agreement between the amplitude of our mean estimate and that of Mann et al. (2007), yet our mean estimate is considerably larger in the earlier parts of the record (nineteenth century and first two decades of the twentieth century).

Based on the locations of historical ship tracks and our methodology, not all TCs in the satellite era are equally likely to have been missed. Figure 6 shows the probability that each satellite-era TC was missed by the historical ship-track locations for two different periods. As might be expected, in both periods, the TCs least likely to encounter a ship or land are those in the central and eastern parts of the basin, being both least likely to encounter land and in the region of least-dense ship sampling (Fig. 2).

For our base case time series, the linear trend over 1900–2006 in this time series of +4.22 storms century^{-1} (+50% century^{-1}) is statistically significant according to all three tests shown in Table 1 (estimated p value of 0.001 or less). However, the trend over the entire 1878–2006 period is +1.42 storms century^{-1} (+15% century^{-1}) and is not significant (estimated p value of ~0.17–0.2, in Table 1). The beginning year of 1900 has been used in previous studies, although ideally we would like to use as long a time series as possible to enhance the signal-to-noise characteristics. This notion supports our emphasis on the 1878–2006 period. However, we also recognize that the uncertainty in our storm adjustment grows larger as we go farther back in time (Fig. 5), which calls for increasing caution regarding trends computed beginning from the earlier parts of the record.

We perform a sensitivity test where the storm count is reduced by one storm per year from 2003 to 2006 to account for recent improvements in detection technologies as proposed by Landsea (2007). This test has little impact on the statistical significance results just discussed: the 1900–2006 trend remains significant while

(a) Atlantic HURDAT Storms (Adjusted for Estimated Missing Storms) 1878-2006

(b) "Missing storm" adjustments to HURDAT Tropical and Subtropical Storm Counts (1878-2006)
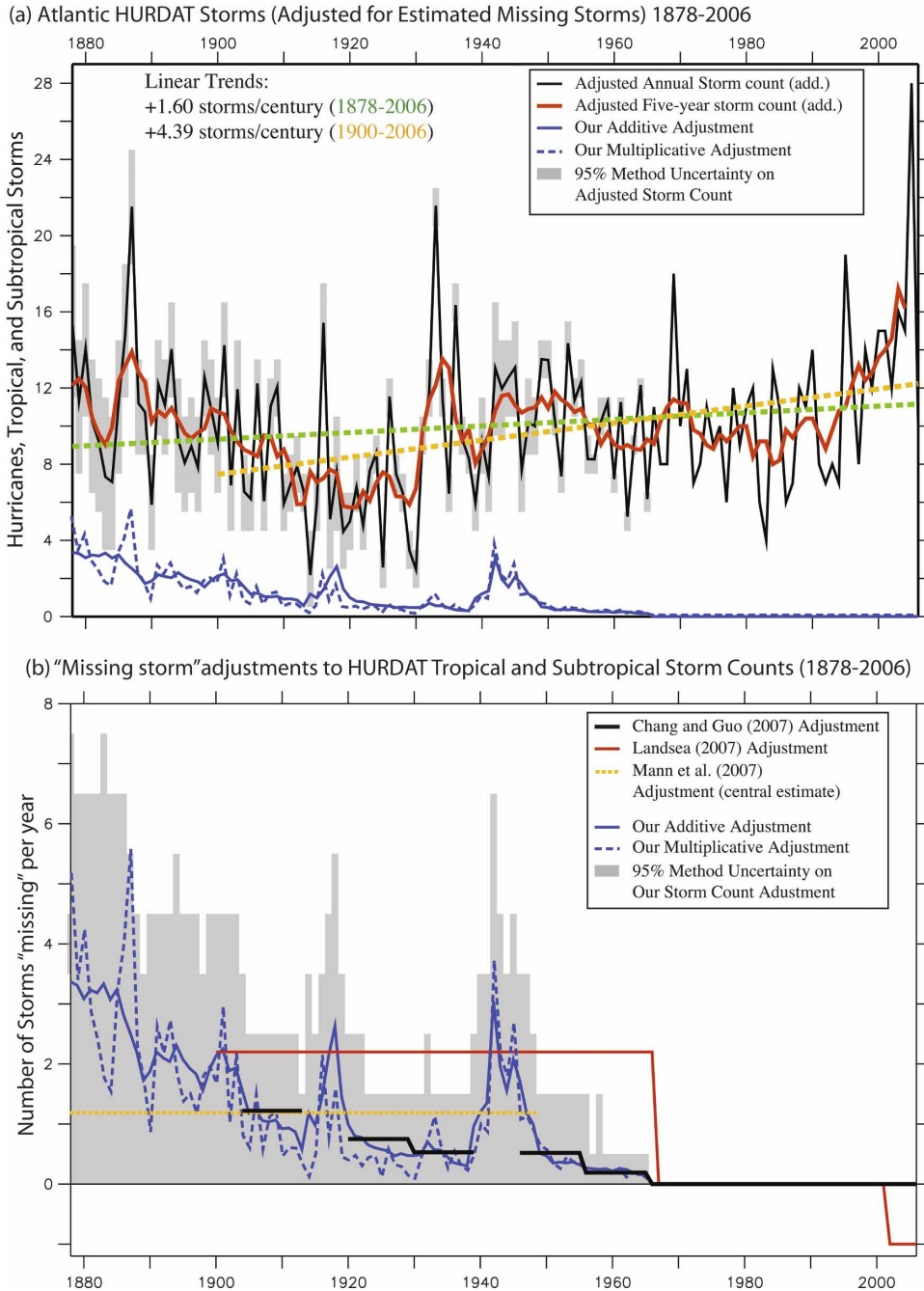

FIG. 5. (top) Time series of Atlantic TC counts, using the base-case adjustment proposed in this study (black). Red line is the 5-yr running mean adjusted storm count. The solid blue line is the adjustment applied to the data (the estimated number of missing storms each year), while the dashed blue line is a sensitivity case requiring one ship/storm encounter for detection instead of two, as in the base case. The green and orange solid lines are linear trend fits to the data for 1878–2006 and 1900–2006, respectively. Black shading indicates the two-sided, 95% method uncertainty, estimated from the probability density function of the estimated adjustments. (bottom) Time series of corrections to Atlantic TC counts, from our study and from three recent analyses (lines); black shading indicates the 95% method uncertainty in our adjustment. (See supplemental information available online at http:dx.doi.org/10.1175/2008JCLI2178.s1)
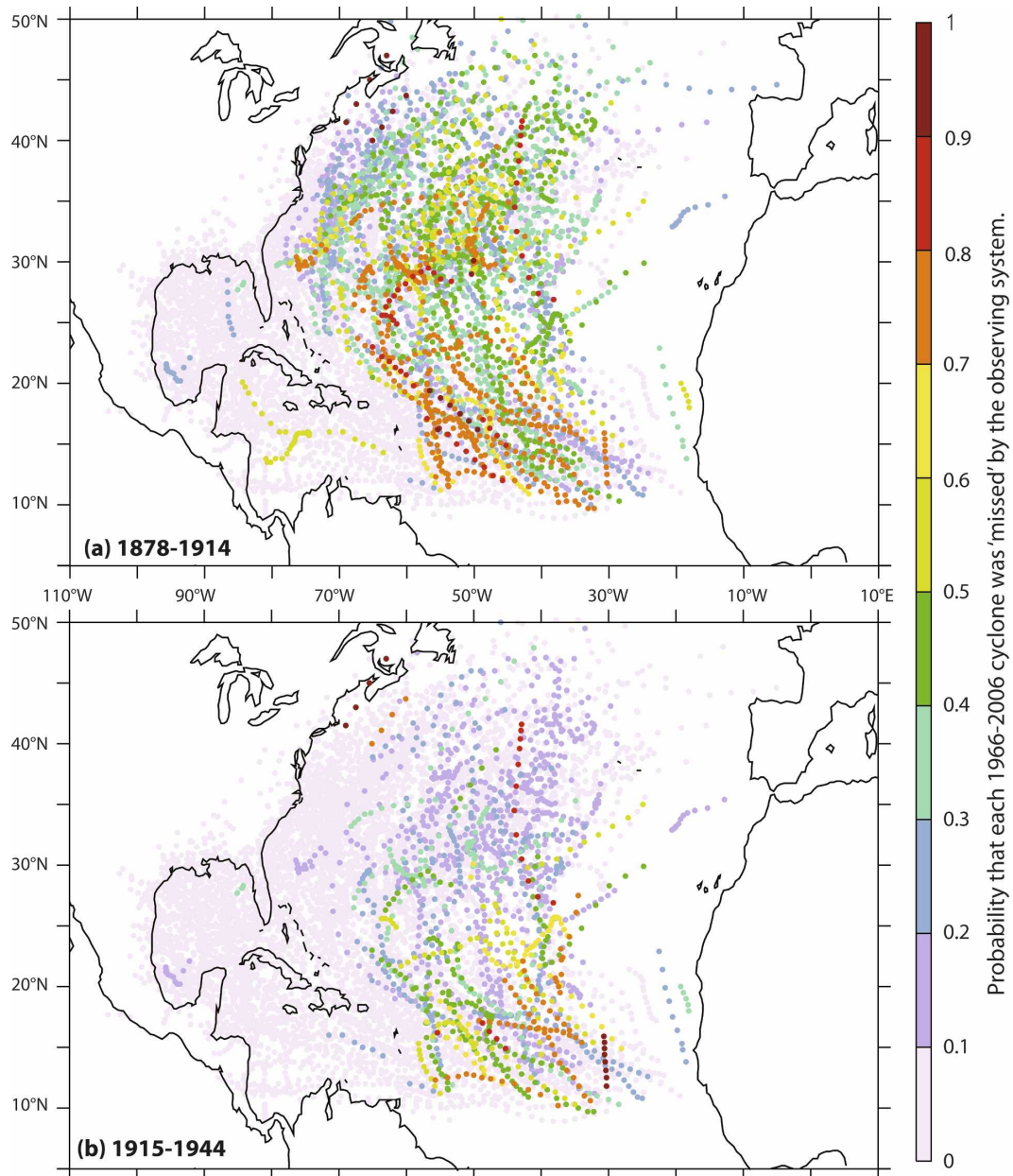
FIG. 6. Probability that a satellite-era TC was "missed" by the historical ICOADS ship observations of the (a) pre–Panama Canal era (1878–1914) and (b) preaircraft reconnaissance era (1915–44).

the 1878–2006 trend remains not significant. Another alternative adjustment is the "multiplicative adjustment" (see section 2d), for which the statistical significance (and dependence on starting dates of the significance) is similar to those for the additive adjustment. Table 1 confirms that the unadjusted tropical storm count trend is highly significant for both the 1900–2006 and 1878–2006 periods, although we believe our storm undercount estimate—as well as those of Chang and Guo (2007), Landsea (2007), and Mann et al. (2007)—

suggests this is an unlikely scenario. Landsea's (2007) proposed adjustment, which was developed only for the 1900–2006 period, leads to a trend of 2.89 storms century^{-1} and is significant according to two of our three tests, with a p value of 0.09 for the ranks test; because Chang and Guo (2007) do not develop their adjustment for the entire record, we cannot estimate its effect on the significance of the trends. Mann et al. (2007) indicate that their adjustment is consistent with a real long-term increase in tropical cyclone activity. The slight

TABLE 1. Linear trend values (century^{-1}) for various tropical cyclone measures, MDR SSTs, and global mean temperature. Results in columns 1–4 are for the period 1900–2006, while columns 5–8 are for 1878–2006. Columns 2 and 6, labeled "t test" give p values for ordinary least squares regressions of the temperature series; for tropical storm–hurricane series, these t tests are done for time series transformed by the square root function. Columns 3 and 7, labeled "Ranks t-test" give p values for ordinary least squares regressions of the ranks of the storm time series (no square root transformation). Columns 4 and 8 give the fraction of randomly generated series that have larger absolute trends than those computed for the original series and show the average percentages obtained for segment lengths of 2–8, using 10^4 random trials for each segment length. The p values represent the estimated probability of incorrectly rejecting the null hypothesis of no trend, using a two-sided test. Trend values in boldface are estimated to be significantly different from zero at the $p = 0.05$ level according to the test results presented. All t-test results use methods that account for serial correlation in the time series as described in the text.

	1900–2006				1878–2006			
	Trend (100 yr^{-1})	t-test p value	Ranks t-test p value	Bootstrap p value	Trend (100 yr^{-1})	t test p value	Ranks t-test p value	Bootstrap p value
Global mean temperature (°C)	**0.72**	$<10^{-6}$	—	$<10^{-4}$	**0.55**	3×10^{-6}	—	$<10^{-4}$
MDR SST (°C) (ERSST2)	**0.71**	0.001	—	$<10^{-4}$	**0.49**	0.004	—	$<10^{-4}$
MDR SST (°C) (HadISST)	**0.51**	9×10^{-4}	—	$<10^{-4}$	**0.41**	1×10^{-4}	—	$<10^{-4}$
MDR SST (°C) (Kaplan)	**0.43**	0.002	—	$<10^{-4}$	**0.26**	0.009	—	9×10^{-4}
TC unadjusted	**6.01**	4×10^{-6}	2×10^{-6}	1×10^{-5}	**3.84**	5×10^{-4}	3×10^{-4}	9×10^{-4}
TC adjusted base case	**4.39**	0.001	0.001	3×10^{-4}	1.60	0.178	0.197	0.172
TC adjusted 2003–06 (−1)	**4.19**	0.001	0.001	3×10^{-4}	1.46	0.201	0.222	0.207
TC adjusted (multiplicative)	**4.76**	5×10^{-4}	7×10^{-4}	2×10^{-4}	1.69	0.147	0.141	0.198
TC adjusted (multiplicative) 2003–06 (−1)	**4.56**	5×10^{-4}	7×10^{-4}	2×10^{-4}	1.55	0.166	0.152	0.234
Tropical storm adjusted (Landsea)	**2.89**	0.030	0.092	0.007	—	—	—	—
U.S. landfall tropical storm	0.369	0.744	0.921	0.502	−0.318	0.431	0.283	0.456
U.S. landfall hurricanes	−0.226	0.499	0.269	0.539	−0.556	0.140	0.058	0.098
TC duration (unadjusted; days)	**−1.47**	2×10^{-4}	2×10^{-4}	0.003	**−1.78**	$<10^{-6}$	$<10^{-6}$	1×10^{-5}
TC duration (adjusted; days)	**−0.988**	0.005	0.010	0.020	**−1.22**	9×10^{-6}	1×10^{-5}	2×10^{-4}

trend in U.S. landfalling tropical storms is not significant, and the more pronounced downward trend in U.S. landfalling hurricanes is also not significant, although for the 1878–2006 period, the p value for the ranks test alone approaches significance ($p = 0.058$).

The adjusted storm count trends can be contrasted with those of SST. The results in Table 1 confirm that the positive trend in global mean temperature is highly significant, in agreement with numerous previous studies and other methods (e.g., Solomon et al. 2007). The MDR SST trends for the three reconstructions [Extended Reconstructed SST, version 2 (ERSST2); HadISST; and Kaplan) are significant, consistent with previous studies (Knutson et al. 2006; Santer et al. 2006).

b. Other activity indices

In this section we explore the century-scale changes in other North Atlantic TC activity indices: tropical storm days per year, average TC duration, and TC density. Because these indices are more complex than TC count, and depend on more than just identifying the existence of a TC, we view the errors inherent in these indices as being larger than those in the TC counts. Nonetheless, because the character of the changes in these indices is quite interesting (and perhaps unexpected), we believe the decadal–centennial-scale variations in these indices are worth exploring. Further, since these indices are intermediate between TC counts and other frequently discussed and physically based indices, such as ACE and PDI, they provide a context for understanding changes in the various indices of TC activity.

1) TROPICAL STORM DAYS PER YEAR

Over the period 1878–2006, the time series of tropical storm days per year (D) from the raw HURDAT dataset does not exhibit a noticeable (or significant) long-term change (Fig. 7a), though the trend is nominally positive (Fig. 7b, black line). Based on D, the most active year in Atlantic TC activity was 1933 (Fig. 7a), which was about 25% more active than 2005. Also, the impact of the artificial increase in sub–gale force wind records in HURDAT (Fig. 3) can be seen in Fig. 7a. Had the contribution to D from records with wind less than 17 m s^{-1} (dotted line) been included, a spurious increase in D would have resulted.

When the time series of D is adjusted (as described in section 2e) for the ship-track-based estimate of missed TCs (Fig. 7b, red line), the nominal trend becomes negative, though, again, not significant. This lack of
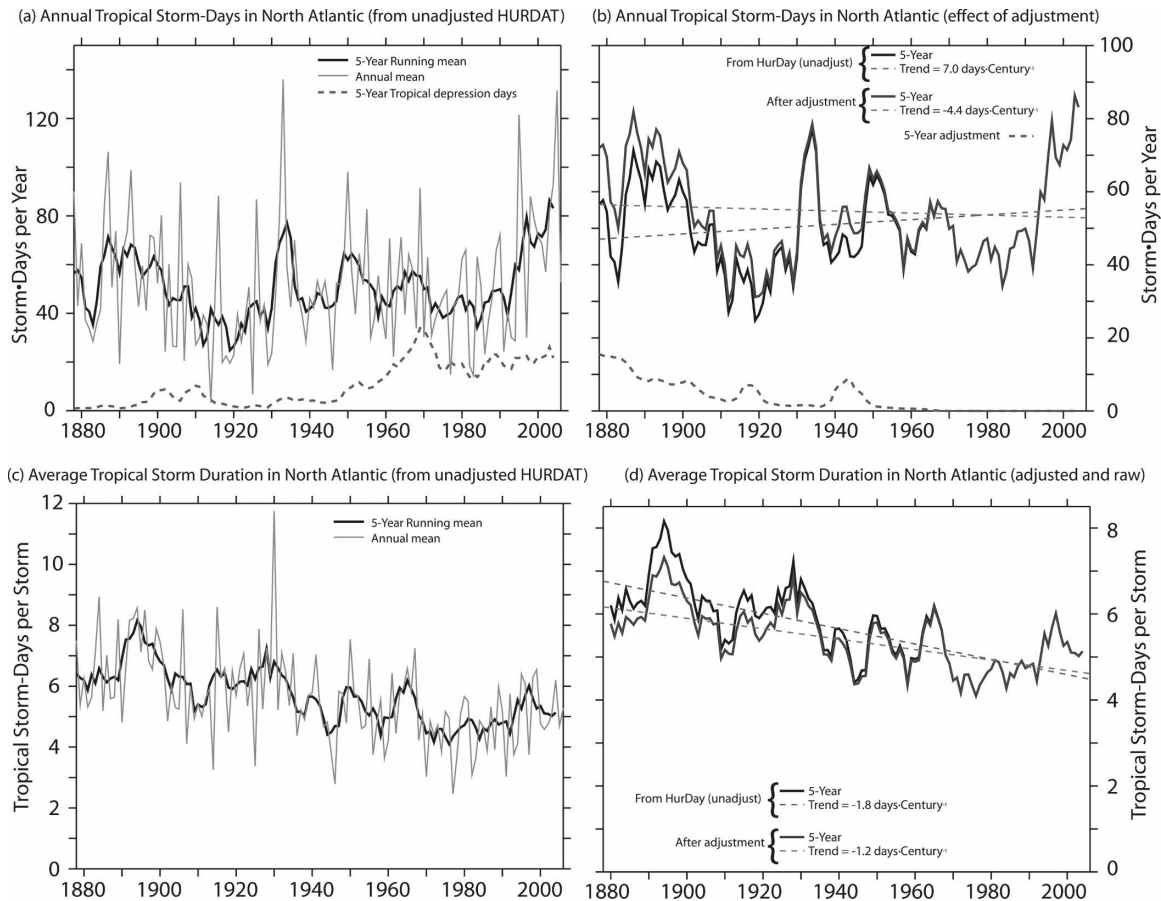
FIG. 7. Historical records of TC activity and average TC duration in the North Atlantic. (a) Time series of tropical storm days per year, 5-yr smoothed (black line), annual mean (gray line), and 5-yr smoothed annual count of HURDAT records with less than 17 m s^{-1} winds. (b) The 5-yr smoothed tropical storm days per year and linear trends (dashed lines); black lines show the values from the unadjusted HURDAT dataset, and red lines show the values when the estimated number of missed storms is included; blue dashed line shows the impact of the adjustment. (c) Average TC duration (days with wind >17 m s^{-1}) from the unadjusted HURDAT database. (d) The 5-yr smoothed (solid lines) and linear trends (dashed lines) for average TC duration from the unadjusted HURDAT database (black line) and ship-based adjusted dataset (red line).

trend in D contrasts with the time series of TC counts shown in Figs. 1 and 3, which show a long-term increase (even if not statistically significant for the adjusted 1878–2006 series). Assuming that our reconstructed time series in Fig. 7b is reliable, the lack of long-term trend in D means that any long-term trends in either ACE or PDI (if they exist) would have been due to long-term changes in intensity, because ACE and PDI are the integrated square and cube, respectively, of the maximum storm wind speed over the lifetime of a storm.

2) AVERAGE TC DURATION

The absence of a significant long-term change in D in the unadjusted HURDAT data, along with the substantial increase in storm counts, indicates that the average tropical storm duration (d) must have exhibited a sub-

stantial negative trend in the long term. Also, since in the ship-track-based adjusted dataset, the nominal trend of D is negative and the nominal trend in the storm counts is positive, d in the adjusted dataset must have exhibited a negative trend. This is confirmed by the time series of d shown in Figs. 7c and 7d, which reveal a substantial and statistically significant decrease in the average storm duration in HURDAT since 1878 (both raw and adjusted as described in section 2e).

This long-term decrease in TC duration in the North Atlantic runs counter to most of the nominal trends in basin-wide storm activity discussed in the literature, which tend to describe a system that is becoming—at least nominally—more active over the twentieth century. As such, this result should be viewed with caution until it can be further verified. However, we are unable to identify any obvious reasons that this long-term de-
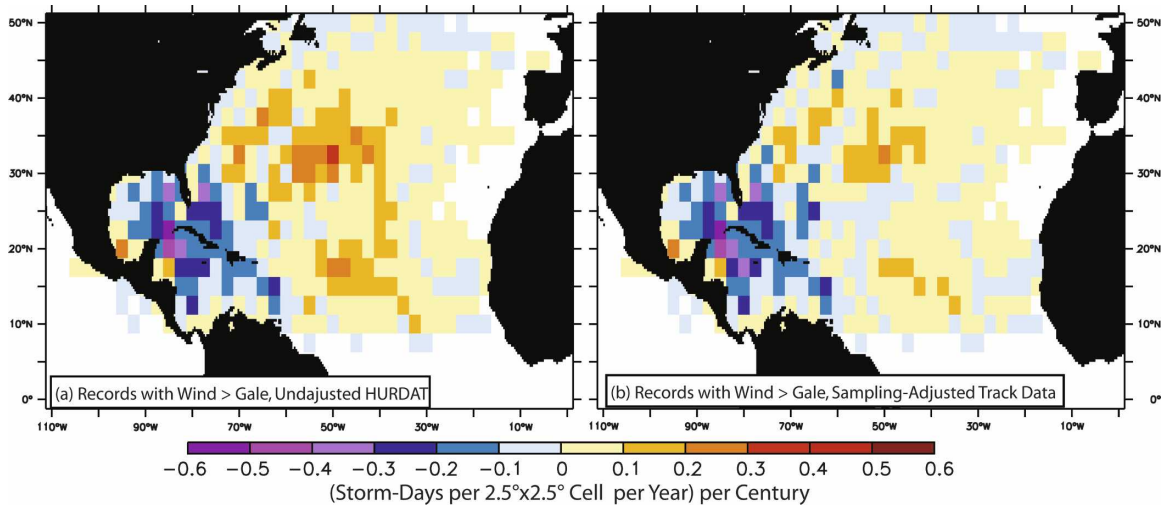
FIG. 8. Maps of linear least squares trend in TC density (density of positions where HURDAT winds exceed 17 m s^{-1}) over 1878–2006 as (a) computed from the unadjusted HURDAT database and (b) computed after adjusting the presatellite records based on the estimated number of missed storms (see section 2e).

crease in average TC duration should have arisen as an artifact of changing observing practices. In fact, because a partial sampling of the basin would be less likely to observe the entire lifetime of a TC, one may expect a spurious trend in the opposite sense. Another possibility is that there has been a systematic change in the estimate of mean wind speeds in TCs, with past observations overestimating the mean speed of storms, which would have led to an artificial reduction in d. However, we are unaware of any reports of such a change, and the reports of systematic biases in the overall wind speeds (e.g., Cardone et al. 1990) are toward a bias toward artificially small amplitudes in the early twentieth century relative to the present.

A third possibility is that short-duration storms are more likely to have been missed altogether than longer-duration storms. This is confirmed by our analysis (Fig. 7d), with the adjusted data showing a smaller change in d than the unadjusted data. However, our adjustment does not eliminate the significance of the decrease in d (see Table 1). This suggests that either (i) our adjustment adds too few storms to the record, (ii) the decrease in d is real, or (iii) an additional factor is leading to a spurious decrease in d. Until a set of sources to explain the full decrease in d is developed, we cannot be confident that it represents a "real" long-term decrease in Atlantic TC activity.

3) TC DENSITY MAPS

Maps of changes in TC density from both the unadjusted HURDAT and ship-track-based adjusted datasets allow us to explore the spatial structure of the long-term changes in tropical storm activity in the North Atlantic (Fig. 8). The changes in TC activity in the Atlantic appear to have occurred in a spatially heterogeneous manner: since the nineteenth century, the western part of the basin (including the Caribbean Sea and Gulf of Mexico) has exhibited a decrease in TC activity, and there has been an increase in activity in the eastern part of the basin.

In the unadjusted HURDAT dataset, the integrated increase in eastern Atlantic activity is nominally larger than the decrease in the western basin (Fig. 8a), resulting in the nominal increase in D (see Fig. 7b). As was shown in Fig. 6, the estimate of missed TCs from ship tracks is not spatially uniform, with storms in the eastern part of the basin being more likely to have been missed by the historical ship tracks. When an estimate of missed TC tracks is included (as described in section 2e), the character of the century-scale trend in TC density is different (Fig. 8b), with the increase in the eastern part of the basin becoming more muted and leading to the nominal decrease in D (Fig. 7b). Overall, both the adjusted and unadjusted datasets indicate that on century scales the activity in the western part of the basin has been decreasing relative to that in the eastern part of the tropical Atlantic.

4. Discussion and conclusions

We have assessed measures of TC activity prior to the satellite era and the likely impact of "missed" TCs on these measures. The long-term changes in Atlantic TC activity are mixed, with different metrics showing either increases, decreases, and no change.
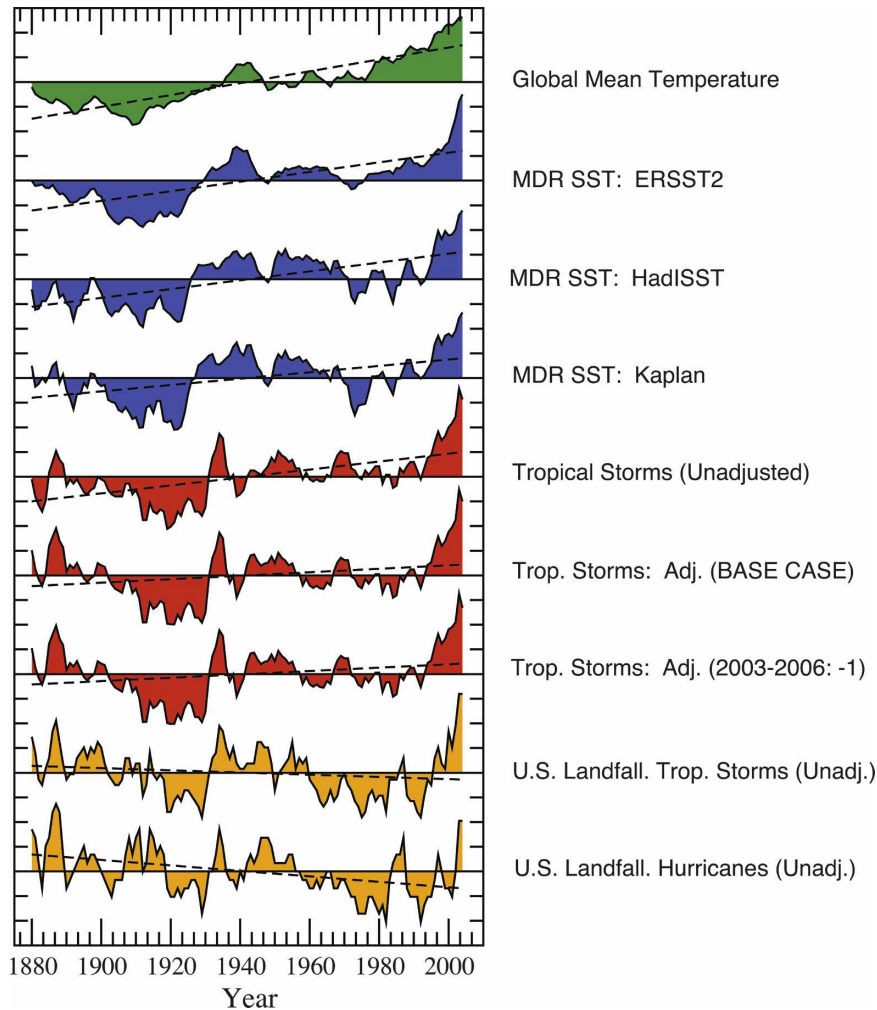
FIG. 9. Normalized tropical Atlantic indices (5-yr running means of time series from 1878 to 2006, solid black) and their linear trends (dashed black). Top (green) is global mean temperature based on HadCRUT3 (Hadley Centre–Climate Research Unit Land Surface Temperature version 3) data obtained online (http://www.cru.uea.ac.uk/cru/data/temperature/). Blue-shaded curves are for the MDR (10°–20°N, 80°–20°W) SST anomalies for August–October from ERSST2, HadISST, and Kaplan datasets. Red-shaded curves are three tropical storm series discussed in the report. The curve labeled "Adj. (2003–2006: −1)" is derived from the same series as our "base case" except that the storm count is decreased by 1 for each year from 2003 through 2006 following Landsea (2007). The orange curves are U.S. landfalling tropical storm and hurricane count series from HURDAT. The marks along the vertical axis represent 1 standard deviation (with each time series having been normalized to have a standard deviation of 1).

Our adjusted tropical TC count time series can be viewed in the context of the broader debate on possible trends in Atlantic TC counts in Fig. 9. This figure shows a progression of relevant time series ranging from global mean temperature (green curve, top), which has the most pronounced rising trend, to U.S. landfalling tropical storms and hurricane counts (orange curves, bottom), which have small (not significant) negative trends over the period 1878–2006. Each time series in Fig. 9

has been normalized by the standard deviation of its 5-yr running mean series.

The Intergovernmental Panel on Climate Change's (IPCC's) Fourth Assessment Report (AR4; Solomon et al. 2007) recently concluded that most of the observed global mean temperature rise since the mid–twentieth century is very likely due to anthropogenic increases in greenhouse gas concentrations. The MDR SST time series, based on three widely used SST reconstructions

(ERSST2, HadISST, and Kaplan), are shown by the blue curves. The MDR series all have a significant rising trend (Table 1), although these time series are not as smooth as the global mean series (i.e., they exhibit greater multidecadal departures from trend than does global mean temperature). Among the MDR time series, the trend is slightly smaller in the Kaplan SST data. Santer et al. (2006) have presented model-based evidence that the rising trend in the MDR in the ERSST2 and HadISST datasets is too large to be explained by internal climate variability alone (see also Knutson et al. 2006).

While the unadjusted storm counts have a similar trend to the MDR series, the application of our adjustments lowers the trend in storm counts to be less than that in the SST series (in terms of normalized data). This trend in our base-case adjusted series is still positive and thus larger than the linear trend in the (unadjusted) U.S. landfalling tropical storm series. The lack of a trend in U.S. landfalling hurricane activity has been noted previously. For example, Landsea (2005) presented a time series of U.S. landfalling power dissipation showing no evidence for a long-term trend. It is possible that the preferentially reduced activity in the western part of the basin (Fig. 8) results in the differing behavior in the landfalling and basin-wide storm counts, as was also suggested by Holland (2007).

The relationship of the low-frequency variability among the different curves is examined in Fig. 10b, in which all series have been detrended. The covariation of the various detrended series on long time scales supports the notion that MDR SST variations on long time scales may modulate Atlantic TC counts, either directly or perhaps indirectly through circulation changes (e.g., Goldenberg et al. 2001; Emanuel 2007; Swanson 2008). Although some discrepancies in this relationship are apparent, including the time lag between the rise in MDR SST in the late 1920s and the rise in TC counts beginning several years later, the agreement between these detrended, normalized series is quite remarkable.

A crucial question is whether this multidecadal relationship between MDR SST and TC counts (e.g., Emanuel 2007; Mann et al. 2007) also holds for the greenhouse-gas-induced warming. It is not necessary that the relationship between local SST change and storm activity be the same for both greenhouse-gas-induced climate change and multidecadal climate variations. The changes in circulation (e.g., vorticity and vertical wind shear), local atmospheric stability, and relative humidity associated with an SST increase due to internal climate variability may well differ from those due to greenhouse-gas-induced warming. For example, Vecchi and Soden (2007a) show that the relationship

between tropical cyclone maximum potential intensity and local SST changes associated with globally homogeneous warming is different from the response to a localized temperature change; Swanson (2008) indicates that Atlantic PDI shows a stronger connection to the departure of MDR SST changes from tropical-mean warming than to MDR SST changes alone. Additionally, Vecchi and Soden (2007c) show that the tropical circulation response to projected future warming includes increased vertical shear over much of the Caribbean and Gulf of Mexico in most of the IPCC AR4 and phase 3 of the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project (CMIP3) climate models. Recent studies with dynamical models showing skill at describing historical Atlantic activity (e.g., Emanuel et al. 2008; Knutson et al. 2007) indicate that a CO_2-induced warming of the tropical Atlantic need not lead to an increase in tropical cyclone counts (e.g., Emanuel et al. 2008; Knutson et al. 2008), although the overall tendency of the Emanuel et al. (2008) experiments is for an increase in counts. Thus, we must understand both the dynamical connections between the environmental conditions and tropical cyclone activity, as well as how these environmental conditions are likely to change from greenhouse gas forcing.

We here use, as a rough approximation, the linear trend in the SSTs (and in TC counts) to estimate the response of both to anthropogenic forcing, including increasing greenhouse gases. According to climate model simulations (e.g., Knutson et al. 2006), the response of tropical Atlantic SSTs to historical estimates of anthropogenic forcing (including greenhouse gases and the direct effects of aerosols only) is fairly linear from the late 1800s to the present. For example, Fig. 11 shows the ensemble response of MDR SSTs from the Geophysical Fluid Dynamics Laboratory (GFDL) coupled climate models (CM2.0 and CM2.1; Delworth et al. 2006; Gnanadesikan et al. 2006; Stouffer et al. 2006; Wittenberg et al. 2006) to estimated anthropogenic radiative forcing (well-mixed greenhouse gases, aerosols, and ozone). These model results suggest that a linear trend is a useful first-order approximation for the response of tropical Atlantic SSTs to anthropogenic forcing. The extent to which the response is linear will need to be reexamined as indirect aerosol forcing (potentially large but not included in these runs) becomes more confidently constrained. From another view, it has recently been argued that SST low-frequency variability and trend in the tropical Atlantic during late summer are primarily radiatively forced (Mann and Emanuel 2006) and, further, that the response of TC counts to monotonic greenhouse warming need not be

(a) Tropical Atlantic SSTs, Tropical Storms, and U.S. Landfall Series

5-yr running mean; Normalized
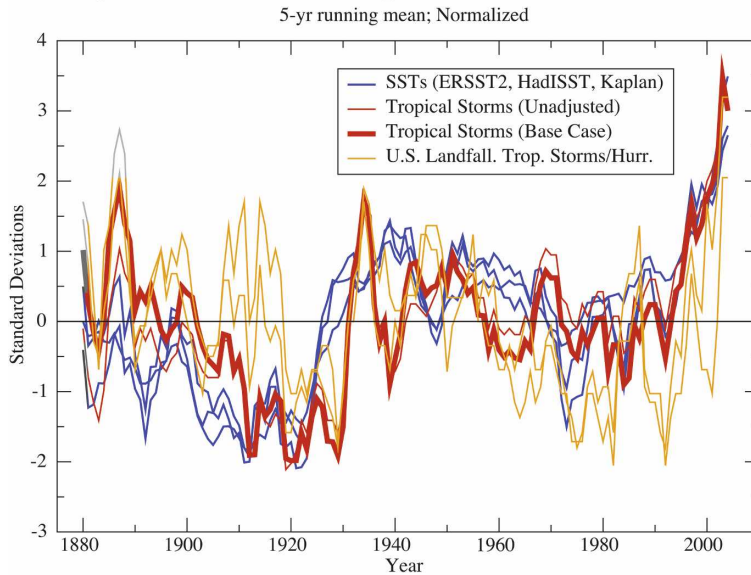
(b) Tropical Atlantic SSTs, Trop. Storms and Landfall Series: Detrended
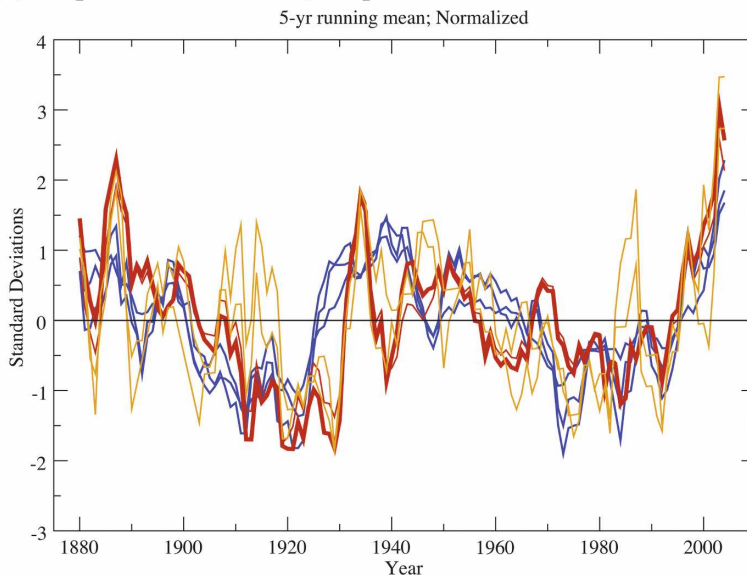
5-yr running mean; Normalized

FIG. 10. (a) The 5-yr running mean normalized Atlantic MDR SST indices from three different reconstructions of SST (see text) overlaid on Atlantic TC counts. Blue curves are the three MDR SST reconstructions. The heavy red curve is the base-case TC count. The light red curve is the unadjusted TC count. The orange curves are U.S. landfalling tropical storm and hurricane count series from HURDAT. Curves in (b) have been detrended using ordinary least squares best fits.

steady (Holland and Webster 2007). Further analyses using alternative statistical measures are on going (Smith et al., personal communication).

Figures 9 and 10 succinctly capture one of the reasons why the crucial issue of the response of Atlantic TC behavior to increasing greenhouse gases remains unsettled: multidecadal variations in Atlantic TC counts appear to be strongly correlated with SSTs, but confidence in the same quantitative sensitivity for the linear trend components of these series is quite limited. Our estimate of this sensitivity appears to depend crucially on the adjustment to the TC count series. Also, the U.S. landfalling TC record supports the notion of no detectable positive impact (and perhaps even a weak negative impact) of anthropogenic forcing on U.S. landfalling activity.
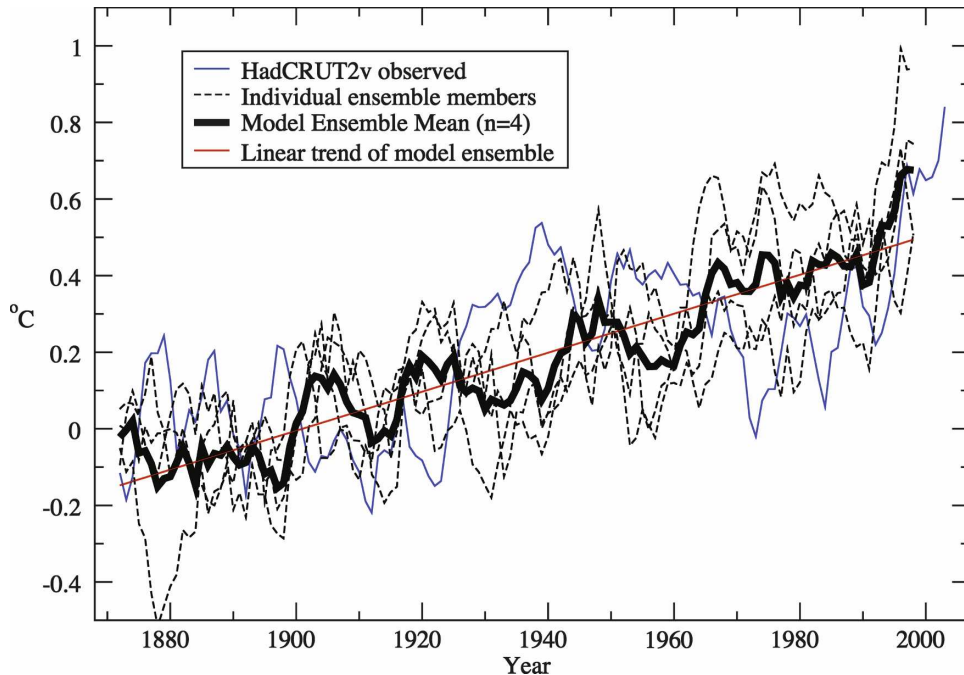
FIG. 11. The 5-yr running average surface temperature changes in the Atlantic MDR from the HadCRU air temperature dataset (blue line), and an ensemble of GFDL CM2.0 and CM2.1 global climate model integrations forced with anthropogenic radiative forcing changes (dashed lines show individual ensemble members; dark solid line shows four-member ensemble mean).

We have also explored other TC activity indices: annual tropical storm days (D), average TC duration (d), and maps of TC density (section 3b). We noted artificial increases in TC records in HURDAT with wind speeds of less than 17 m s^{-1} (Fig. 3), which must be accounted for in analyses of TC duration (e.g., Fig. 7a). After removing all records with wind speeds of less than 17 m s^{-1}, we find no significant century-scale trend in D.

Our analysis also suggests that the average TC duration (d) in the Atlantic has decreased significantly since the late 1800s (Fig. 7). It is possible that this decreasing trend in d is an artifact of changing observing practices—and not a real climate signal—though we are unable to identify a spurious source for this trend. A possible explanation is that the storms most likely to be missed in the early part of the record were remote ones that also had relatively short lifetimes, yet this would imply an adjustment to TC counts that is larger than ours. Providing some support for this conjecture, the ship-track-adjusted time series of d has a smaller negative trend than that from HURDAT—though both trends are significant. Interestingly, the time series of d does not exhibit any clear relationship to MDR SST even on multidecadal time scales, suggesting that it may be controlled by factors other than SST, or could be associated with data problems. Assuming that this decrease in average TC duration is "real," the extent to

which it represents the forced response of the climate system or internal climate variability is unclear (as is also the case for the other storm measures), since the factors driving the TC duration changes have not been clearly ascertained.

It appears that, while the total number of TCs in the North Atlantic has exhibited at least a nominal increase since the late nineteenth century, the average TC duration may have had a long-term decrease. If this represents the effect of changing climate conditions, it suggests that aspects of climate have changed (through some combination of radiative forcing and internal climate variations) in order to make the North Atlantic more favorable to cyclogenesis, while at the same time making the overall environment less favorable to TC maintenance. The model experiments of Knutson et al. (2008) indicate a modest reduction of d in response to increased CO_2, although the model sensitivity ($\sim$0.7-day reduction for 1.7-K MDR warming) is too small to explain the observed reduction ($\sim$1.8-day reduction for a 0.5-K warming). Further investigation is required to ascertain the reasons for the decreasing duration we find in both the existing and adjusted data.

Maps of TC density change (Fig. 8) indicate a reduction of TC activity in the western part of the basin and an increase in the eastern part. This reduction of TC activity closer to the common landfalling locations in

the western part of the basin may help explain the differing evolution patterns of the time series of basin-wide TC counts and that of landfalling storms. The extent to which the spatial pattern of the observed changes in TC activity may be an artifact of changing observing practices, due to internal climate variability or a result of forced changes to the global climate system, bears examination.

The apparent eastward displacement of storm activity may have resulted from an eastward shift of tropical cyclogenesis over the twentieth century, which may be related to an eastward displacement of the extent of the warmest tropical Atlantic waters (P. Webster 2007, personal communication) or to changes in the SST gradient across the equator (e.g., Vimont and Kossin 2007). There is also some correspondence between the region that shows a long-term decrease in TC density and the region in the model projections of global warming (e.g., Vecchi and Soden 2007c) that exhibits an increase in vertical wind shear and a decrease in midtropospheric relative humidity. Both of the latter would make the environment less conducive for tropical cyclone genesis, maintenance, and intensification. Indeed, the model results of Knutson et al. (2008) indicate a modest eastward shift of Atlantic cyclone activity in response to increased CO_2. It is noteworthy that this twentieth-century decrease in storm activity occurs in one of the—relatively—best observed parts of the basin. If this reduction of activity in the western part of the basin is not spurious, we speculate that it could represent the signature of century-scale changes in environmental conditions like those obtained from model projections of a warming climate.

Though the century-scale changes in the activity measures discussed here are mixed (see above), the past 30 yr have shown an increase in all of the Atlantic TC activity measures discussed here (e.g., Figs. 5, 7, 9, and 10) and others discussed elsewhere (e.g., Landsea et al. 1999; Goldenberg et al. 2001; Webster et al. 2005; Emanuel 2005). Over the satellite era, when data quality is highest and most homogeneous, the character of the changes is unambiguous, defining a clear and real recent increase in Atlantic TC activity. However, the relatively short (30 yr) record, and limitations of both models and observations, make it difficult to determine the contributions of internal climate variability (e.g., Zhang and Delworth 2006; Vimont and Kossin 2007), localized radiative forcing from aerosols (e.g., Mann and Emanuel 2006), or increasing greenhouse gases (e.g., Santer et al. 2006) to the recent increase in TC activity. While the 1878–2006 record is perhaps sufficiently long to address these issues, the decreasing data quality, changing observing practices, and mixed char-

acter of the activity changes do not allow for unqualified conclusions to be drawn at this time.

The magnitude and statistical significance of the linear trend computed from the time series of TC counts is highly dependent on the endpoints chosen. Adjusted storm counts exhibit a strong and statistically significant positive trend over the period 1900–2006, while the trend over the longer period 1878–2006 is not significant ($p \sim 0.2$). Critical to the strong 1900–2006 trend is not just the increase in counts since the mid-1990s, but also the minimum in 1910–30; the 1878–2006 trend is damped because the adjusted data indicate that the period 1878–1900 was quite active. So a key question becomes: Which starting date for trend computation is most justified? Often a year around 1900 has been used as the beginning year for such analyses (e.g., Landsea 2007; Chang and Guo 2007; Holland and Webster 2007). Landsea (2007) argues that land-based observations may be substantially more reliable before 1900 than after. Based on relationships between weak, moderate, and strong TCs in HURDAT, Holland and Webster (2007) argue that 1905 should be used as the beginning of the reliable record for comparative TC intensity analyses. However, we are unaware of any fundamental observing technique changes that would provide an a priori distinction for 1905 (or any year between 1878 and 1914) as the beginning of the reliable record; the following years, on the other hand, correspond to substantial changes to the way TCs were observed and recorded: 1878 (U.S. Signal Corps begins monitoring and recording hurricanes), 1914 (opening of Panama Canal; shipping in Caribbean/Gulf of Mexico and across tropical Atlantic increases), 1918 (WWI ends), 1944 (aircraft reconnaissance begins in the western tropical Atlantic), 1945 (WWII ends), and 1966 (basin-wide satellite monitoring begins).

Changes to observing practices are not the only factor to consider when deciding on the period over which to compute trends. If, as has been suggested by some (e.g., Goldenberg et al. 2001; Zhang and Delworth 2006; Landsea 2007), natural multidecadal variations in the North Atlantic drive changes in tropical cyclone activity, then long records are crucial to help filter out "noise," and using start–end dates that correspond to opposite extremes of internal variations may lead to spurious significance when estimating the greenhouse-forced signal from long-term trends. This would argue against using the early 1900s as a starting date. On the other hand, if the variations in the North Atlantic are primarily radiatively forced, as has also been suggested (e.g., Mann and Emanuel 2006; Holland and Webster 2007), the problem becomes one of separating the greenhouse gas or net anthropogenic influences from

natural radiative forcings such as volcanic or solar variations. In this case, simple linear trend analysis may not be appropriate, and more detailed modeling is required. For a confident assessment, the physical character of the multidecadal variations in long-term climate conditions in the tropical Atlantic must be better understood, and used along with our knowledge of the changes in TC observing methodology, to better understand the causes of long-term changes in TC activity.

We reemphasize that while in this paper we estimate certain key sources of uncertainty in the historical Atlantic TC database, other possible sources of uncertainty remain. For example, observational errors may have led to the erroneous inclusion or exclusion of records in HURDAT, historical TCs may have been "double counted" (where one TC was misidentified as two), or perhaps storms made landfall and were unrecorded. In addition, a reanalysis of HURDAT is currently under way (e.g., Landsea et al. 2007). Thus, our current estimates of long-term changes in TC activity should be regarded as tentative, particularly when analyses span periods in which substantial changes in observing practices have occurred, and efforts should continue to update and enhance our historical records of TCs and their uncertainties.

Overall, our findings suggest that it is possible that Atlantic TC counts may have significantly increased since the late nineteenth century, although the evidence is decidedly mixed, with some other activity measures showing either no change or a decrease with time. Total storms per year and U.S. landfalling activity show no increasing trend, and average TC duration shows a significant decrease over time. Further, attribution of an increase in tropical storm counts to any particular mechanism (including increasing greenhouse gasses or natural decadal variations) would require further dynamical analysis to complement any observational results. It is noteworthy that in our adjusted record of TCs the sensitivity of basin-wide storm counts to local SST is smaller for the longest time scales (e.g., trend since 1878) than for the pronounced multidecadal variability, although the current observational "best estimate" would be that this sensitivity is positive. Additional study is needed to reconcile these findings with climate model simulations of past and future Atlantic storm activity. Future work should also focus on including more ship-track information where possible and examining assumptions about landfall detection in earlier years, and historical tropical cyclone database reconstructions should be extended to include other basins.

REFERENCES

Bengtsson, L., K. I. Hodges, M. Esch, N. Keenlyside, L. Kornblueh, J.-J. Luo, and T. Yamagata, 2007: How may tropical cyclones change in a warmer climate. *Tellus,* **59A,** 539–561.

Cardone, V. J., J. Greenwood, and M. A. Cane, 1990: On trends in historical marine wind data. *J. Climate,* **3,** 1051–1054.

Chang, E. K. M., and Y. Guo, 2007: Is the number of North Atlantic tropical cyclones significantly underestimated prior to the availability of satellite observations? *Geophys. Res. Lett.,* **34,** L14801, doi:10.1029/2007GL030169.

Delworth, T. L., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate,* **19,** 643–674.

Emanuel, K. A., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature,* **436,** 686–688, doi:10.1038/nature03906.

——, 2007: Environmental factors affecting tropical cyclone power dissipation. *J. Climate,* **20,** 5497–5509.

——, R. Sundarajan, and J. Williams, 2008: Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bull. Amer. Meteor. Soc.,* **89,** 347–367.

Fernández-Partagás, J., and H. F. Diaz, 1996: Atlantic hurricanes in the second half of the nineteenth century. *Bull. Amer. Meteor. Soc.,* **77,** 2899–2906.

Gnanadesikan, A., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part II: The baseline ocean simulation. *J. Climate,* **19,** 675–697.

Goldenberg, S. B., C. Landsea, A. M. Mestas-Nunez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity. *Science,* **293,** 474–479.

Gualdi, S., E. Scoccimarro, and A. Navarra, 2008: Changes in tropical cyclone activity due to global warming: Results from a high-resolution coupled general circulation model. *J. Climate,* in press.

Holland, G. J., 2007: Misuse of landfall as a proxy for Atlantic tropical cyclone activity. *Eos, Trans. Amer. Geophys. Union,* **88** (36), 349–350.

——, and P. J. Webster, 2007: Heightened tropical cyclone activity in the North Atlantic: Natural variability or climate trend? *Philos. Trans. Roy. Soc. Math. Phys. Eng. Sci.,* **365A,** 2695–2716.

Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic Basin, 1886–1983: Contents, limitations, and uses. NOAA Tech. Memo. NWS NHC 22, Coral Gables, FL, 21 pp. [Available online at http://www.nhc.noaa.gov/pdf/NWS-NHC-1988-22.pdf.]

Kaplan, A., and Coauthors, 1998: Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res.,* **103,** 18 567–18 589.

Kimball, S. K., and M. S. Mulekar, 2004: A 15-year climatology of North Atlantic tropical cyclones. Part I: Size parameters. *J. Climate,* **17,** 3555–3575.

Knutson, T. R., and R. E. Tuleya, 2004: Impact of CO_2-induced warming on simulated hurricane intensity and precipitation: Sensitivity to the choice of climate model and convective parameterization. *J. Climate,* **17,** 3477–3495.

——, and Coauthors, 2006: Assessment of twentieth-century regional surface temperature trends using the GFDL CM2 coupled models. *J. Climate,* **19,** 1624–1651.

——, J. J. Sirutis, S. T. Garner, I. M. Held, and R. E. Tuleya, 2007: Simulation of the recent multidecadal increase of Atlantic hurricane activity using an 18-km-grid regional model. *Bull. Amer. Meteor. Soc.,* **88,** 1549–1565.

——, ——, ——, G. A. Vecchi, and I. M. Held, 2008: Simulated reduction in Atlantic hurricane frequency under twenty-first-century warming conditions. *Nature Geosci.,* **1,** 359–364.

Landsea, C. W., R. A. Pielke Jr., A. M. Mestas-Nuñez, and J. A. Knaff, 1999: Atlantic basin hurricanes: Indices of climatic changes. *Climatic Change,* **42,** 89–129.

——, and Coauthors, 2004: The Atlantic hurricane database reanalysis project: Documentation for the 1851–1910 alterations and additions to the HURDAT database. *Hurricanes and Typhoons: Past, Present and Future,* R. J. Murname and K.-B. Liu, Eds., Columbia University Press, 177–221.

——, 2005: Meteorology: Hurricanes and global warming. *Nature,* **438,** E11–E12, doi:10.1038/nature04477.

——, 2007: Counting Atlantic tropical cyclones back in time. *Eos, Trans. Amer. Geophys. Union,* **88** (18), 197–203.

——, and Coauthors, 2008: A reanalysis of the 1911–20 Atlantic hurricane database. *J. Climate,* **21,** 2138–2168.

Mann, M. E., and K. Emanuel, 2006: Atlantic hurricane trends linked to climate change. *Eos, Trans. Amer. Geophys. Union,* **87,** 233–241.

——, T. A. Sabbatelli, and U. Neu, 2007: Evidence for a modest undercount bias in early historical Atlantic tropical cyclone counts. *Geophys. Res. Lett.,* **34,** L22707, doi:10.1029/2007GL031781.

Oouchi, K., J. Yoshimura, H. Yoshimura, R. Mizuta, S. Kusunoki, and A. Noda, 2006: Tropical cyclone climatology in a global-warming climate as simulated in a 20 km-mesh global atmospheric model: Frequency and wind intensity analyses. *J. Meteor. Soc. Japan,* **84,** 259–276.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.,* **108,** 4407, doi:10.1029/2002JD002670.

Santer, B. D., and Coauthors, 2006: Forced and unforced ocean temperature changes in Atlantic and Pacific tropical cyclogenesis region. *Proc. Natl. Acad. Sci. USA,* **103,** 13 905–13 910, doi:10.1073/pnas.0602861103.

Shen, W., R. E. Tuleya, and I. Ginis, 2000: A sensitivity study of the thermodynamic environment on GFDL model hurricane intensity: Implications for global warming. *J. Climate,* **13,** 109–121.

Smith, T. M., and R. W. Reynolds, 2004: Improved extended reconstruction of SST (1854–1997). *J. Climate,* **17,** 2466–2477.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Eds., 2007: *Climate Change 2007: The Physical Science Basis.* Cambridge University Press, 996 pp.

Solow, A. R., and L. J. Moore, 2002: Testing for trend in North Atlantic hurricane activity, 1900–98. *J. Climate,* **15,** 3111–3114.

Stouffer, R., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part IV: Idealized climate response. *J. Climate,* **19,** 723–740.

Swanson, K. L., 2008: Nonlocality of Atlantic tropical cyclone intensities. *Geochem. Geophys. Geosyst.,* **9,** doi:10.1029/2007GC001844.

Vecchi, G. A., and B. J. Soden, 2007a: Effect of remote sea surface temperature change on tropical cyclone potential intensity. *Nature,* **450,** 1066–1070, doi:10.1038/nature06423.

——, and ——, 2007b: Global warming and the weakening of the tropical circulation. *J. Climate,* **20,** 4316–4340.

——, and ——, 2007c: Increased tropical Atlantic wind shear in model projections of global warming. *Geophys. Res. Lett.,* **34,** L08702, doi:10.1029/2006GL028905.

——, A. Clement, and B. J. Soden, 2008: Examining the tropical Pacific's response to global warming. *Eos, Trans. Amer. Geophys. Union,* **89** (9), 81–83.

Vimont, D. J., and J. P. Kossin, 2007: The Atlantic meridional mode and hurricane activity. *Geophys. Res. Lett.,* **34,** L07709, doi:10.1029/2007GL029683.

Webster, P. J., G. J. Holland, J. A. Curry, and H.-R. Chang, 2005: Changes in tropical cyclone number, duration and intensity in a warming environment. *Science,* **309,** 1844–1846.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Elsevier Academic Press, 627 pp.

Wittenberg, A. T., A. Rosati, N.-C. Lau, and J. J. Ploshay, 2006: GFDL's CM2 global coupled climate models. Part III: Tropical Pacific climate and ENSO. *J. Climate,* **19,** 698–722.

Worley, S. J., S. D. Woodruff, R. W. Reynolds, S. J. Lubker, and N. Lot, 2005: ICOADS release 2.1 data and products. *Int. J. Climatol.,* **25,** 823–842.

Zhang, R., and T. L. Delworth, 2006: Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys. Res. Lett.,* **33,** L17712, doi:10.1029/2006GL026267.