Click Here for Full Article

Evidence for a modest undercount bias in early historical Atlantic tropical cyclone counts

Michael E. Mann,[1] Thomas A. Sabbatelli,[1] and Urs Neu[2]

[1] We obtain new insights into the reliability of long-term historical Atlantic tropical cyclone ('TC') counts through the use of a statistical model that relates variations in annual Atlantic TC counts to climate state variables. We find that the existence of a substantial undercount bias in late 19th through mid 20th century TC counts is inconsistent with the statistical relationship between TC counts and climate. **Citation:** Mann, M. E., 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.

1. Introduction

[2] Positive trends over recent decades have been established for various measures of the powerfulness of Atlantic TCs [*Emanuel*, 2005a; *Webster et al.*, 2005; *Sriver and Huber*, 2006; *Hoyos et al.*, 2007]. These trends appear connected with increasing tropical Atlantic sea surface temperatures (SSTs) [*Emanuel*, 2005a; *Hoyos et al.*, 2007], and the increasing SSTs are likely driven in large part by anthropogenic climate change [*Mann and Emanuel*, 2006; *Trenberth and Shea*, 2006; *Santer et al.*, 2006]. More debated, however, are long-term trends in Atlantic TC occurrence rates or 'counts' and, in particular, their reliability in earlier decades.

[3] Observations of TC behavior have improved over time, particularly since the mid 20th century with the development of more sophisticated observing systems such as aircraft reconnaissance and satellite technology. Prior to the availability of modern observing systems, detection of tropical storms was based on observations along coastlines and from ships. As the density of those observations becomes increasingly sparse back in time, it is plausible that an increasing number of TCs were missed by observers in prior decades.

[4] In some studies, a substantial undercount bias has been argued to exist, particularly prior to the availability of aircraft reconnaissance in 1944. *Nyberg et al.* [2007] used a biological proxy-based reconstruction of past TC activity to infer average counts for major (category 3 and higher) Hurricanes between 1870–1943 that exceed the high levels of the past decade. This finding would imply an average pre-1944 undercount bias of 2–3 storms a year for major hurricanes alone. Assuming that major hurricanes account for between a sixth and a third of all named storms (the range of decadal variation in this fraction from the available historical record by *Holland and Webster* [2007]), such an estimate would imply a dramatic pre-1944 average annual undercount of 6–20 named storms.

[5] Others [*Landsea*, 2005, 2007; *Landsea et al.*, 2004] have argued for a smaller, but still substantial undercount. *Landsea* [2007] estimated undercounts based on changes over time in the proportion of total reported Atlantic named storms ("PTL") that achieved landfall on islands or along the U.S. coastline. PTL was interpreted as a proxy for the underreporting of total TC counts due primarily to decreasing track density back in time, with higher PTL interpreted as reflecting increased undercount bias. The increase in average PTL over 1900–1965 (75%) relative to 1966–2006 (59%) was interpreted by *Landsea* [2007] as reflecting an undercount bias of 2.2 total named TCs for the earlier period. Combined with a speculative additional 1 storm bias that was assumed to hold prior to 2003, this led *Landsea* [2007] to estimate to an aggregate undercount bias of 3.2 named TCs prior to 1966.

[6] Yet still other studies conclude that the long-term record of Atlantic TC counts is likely reliable back through the late 19th century [*Emanuel*, 2005a; *Mann and Emanuel*, 2006; *Holland and Webster*, 2007]. Some analyses [*Neumann et al.*, 1999; *Holland and Webster*, 2007; *Chang and Guo*, 2007] estimate an average undercount of at most only about one TC through the early 20th century. Indeed, an alternative analysis of long-term PTL changes to that described above implies modest undercount. Extending PTL prior to 1900, we find that PTL actually *decreased* prior to 1900 (see auxiliary material),[1] implying, by the reasoning of *Landsea* [2007], a decrease, rather than increase, in undercount bias prior to 1900. Yet logically the undercount bias should be higher prior to 1900 owing to a further decrease in ship tracks, and coastal population density. The mean value of PTL was 67% over 1851–1899, and lower still (61%) for 1851–1885 (see *Holland* [2007] for a similar analysis). This latter value is statistically indistinguishable from the average PTL over the modern period 1966–2006 (59%), which by Landsea's reasoning, would imply a minimal undercount over 1851–1885.

[7] Of course, estimation of undercount based on the assumption of a fixed relationship between total TC counts and the number of landfalling storms is perilous. Such an approach assumes, in particular, that the large-scale atmospheric steering which determines the trajectories of TCs once they've formed is constant, when there is in fact strong

[1]Department of Meteorology and Earth and Environmental Systems Institute, Pennsylvania State University, University Park, Pennsylvania, USA.
[2]ProClim, Swiss Academy of Sciences, Bern, Switzerland.

[1]Auxiliary materials are available in the HTML. doi:10.1029/2007GL031781.
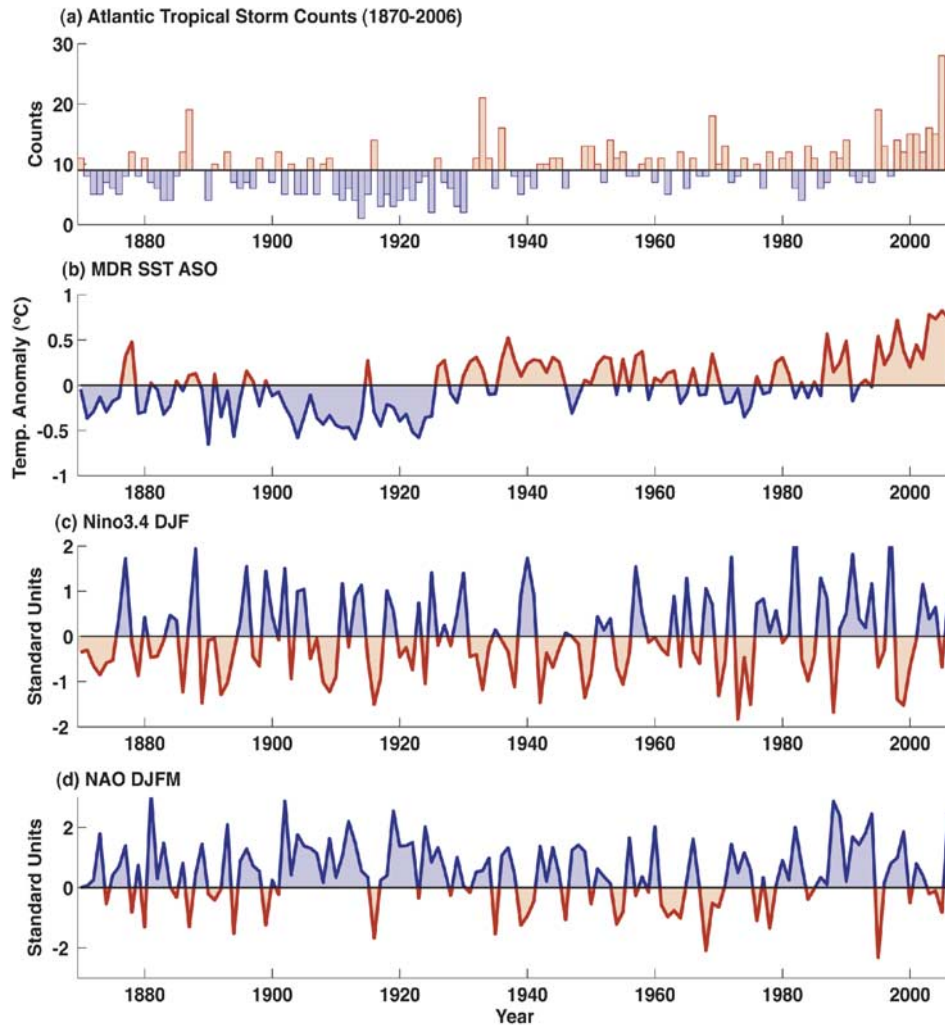
Figure 1. Time series (1870–2006) of (a) annual Atlantic TC counts, (b) MDR ASO SST time series, (c) Niño3.4 DJF SST index, and (d) NAO DJFM SLP index. Red (blue) indicates positive (negative) anomalies in TC counts and hurricane-favorable (unfavorable) conditions in the three indices (MDR SST, Niño3.4, and NAO). Note that year convention applies to the 'D' in DJF and DJFM for both Figures 1c and 1d.

evidence that it is highly variable over time [*Elsner*, 2003]. Yet, other methods of estimating the undercount bias in early TC count data (e.g. interpretation of ship logs, or use of proxy climate data) rest on what are arguably equally tenuous assumptions.

[8] Seeking to resolve the discrepancies between the different assessments of Atlantic TC undercount bias discussed above, we have turned to an alternative approach that does not require any of the above assumptions. Instead, our approach draws upon a recently developed statistical model that conditions expected total Atlantic TC counts on underlying climate variables.

2. Methods and Data

[9] While a significant component of the variation in TC counts from year to year represents the chance fluctuations of a random Poisson process, systematic changes over time in the mean expected rate of occurrence (i.e. annual TC counts), are believed to result from large-scale climate

forcing. Past work has shown that the El Niño/Southern Oscillation (ENSO) [*Gray*, 1984; *Bove et al.*, 1998; *Elsner et al.*, 2000, 2006; *Elsner*, 2003], the North Atlantic Oscillation (NAO) [*Elsner*, 2003; *Elsner et al.*, 2000, 2006], and tropical Atlantic Sea Surface Temperatures (SST) [*Gray*, 1968; *Shapiro*, 1982; *Shapiro and Goldenberg*, 1998; *Saunders and Harris*, 1997; *Goldenberg et al.*, 2001; *Emanuel*, 2005b; *Mann and Emanuel*, 2006; *Holland and Webster*, 2007] each lead to variations in annual Atlantic TC counts. The first two factors influence the amount of vertical wind shear in the atmosphere (less shear is more favorable for TC development), and the tendency for re-curvature of storms (which influences the likelihood that TCs encounter a favorable thermodynamic environment for development). Their impacts on TC development are limited primarily to the latter part of the storm season (the boreal fall and early winter) and can therefore be measured using the boreal winter seasonal means conventionally used to define ENSO and NAO indices. The third factor represents the local thermodynamic factors influencing development, and is
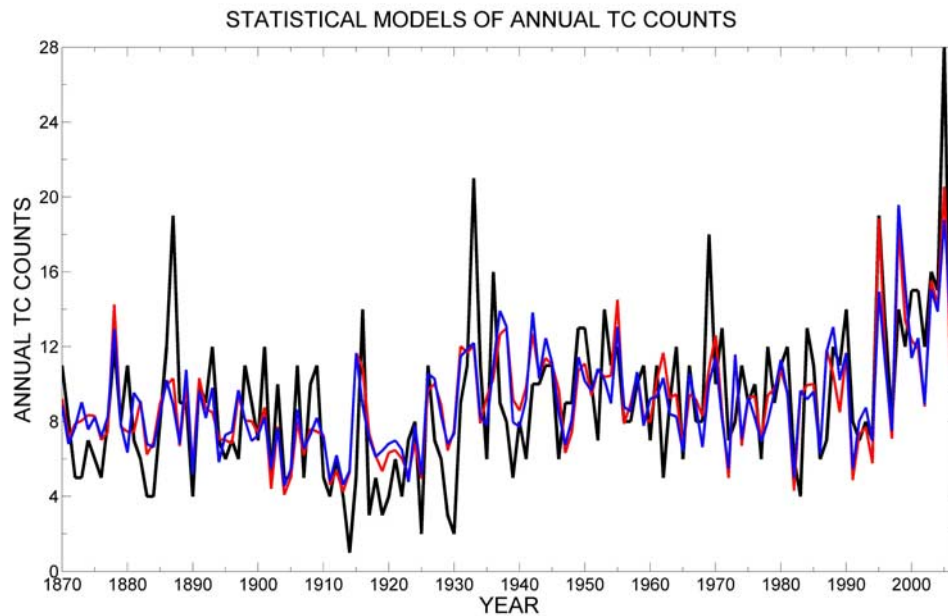
Figure 2. Regression model for annual Atlantic TC counts using the MDR ASO SST, Niño3.4, and NAO series as predictors. Shown are the statistical fits over 1870–2006 based on all three predictors (red) and only the first two predictors (blue) along with the observed TC counts for 1870–2006 (black).

conventionally [e.g., *Emanuel*, 2005a] measured by its state during the peak of the storm season.

[10] The impact of these factors on expected TC counts can be accounted for [*Elsner*, 2003; *Elsner et al.*, 2000, 2006], through a variant on linear regression known as 'Poisson Regression' which is appropriate for modeling the influence of some set of independent variables ('covariates') on the expected rate of a Poisson distributed random process. We employ a recently developed and validated Poisson regression model [*Sabbatelli and Mann*, 2007] for conditional expected Atlantic annual TC counts, updated here using a new blended SST product derived by averaging three alternative published SST datasets [*Rayner et al.*, 2003; *Smith and Reynolds*, 2003; *Kaplan et al.*, 1998] all available back to 1870. Historical Atlantic total named storm counts [*Jarvinen et al.*, 1984] were regressed against three covariates: (1) mean SSTs during the August-October ('ASO') peak of the tropical storm season over the main development region ('MDR': $6°–18°N$, $20°–60°W$) for Atlantic TCs, (2) the boreal winter Niño3.4 SST index of ENSO (normalized SST averaged over the region $5°S–5°N$, $120°–170°W$), and (3) the boreal winter NAO index [*Jones et al.*, 1997]. None of the three covariates are significantly correlated with each other. However, we do find a statistically significant lagged correlation relating the Niño3.4 index to the MDR SST series for the following year's storm season, consistent with the observation elsewhere [*Trenberth and*

Shea, 2006] that ENSO events influence tropical Atlantic SST in the following summer. This lagged influence of ENSO is implicitly accounted for in our analysis via the use of MDR SST as a statistical predictor. The MDR SST, Niño3.4, and NAO series used are shown along with the TC count series in Figure 1. Data used and other supplementary information can be found at: http://www.meteo.psu.edu/~mann/TC_GRL07.

[11] The statistical model for Atlantic TC counts resulting from training (or 'calibrating') over the full 1870–2006 interval is shown in Figure 2. The statistical model captures a substantial fraction $R^2 = 50\%$ (i.e., half) of the total annual variance in TC counts (statistically significant at the $p \ll 0.01$ level). The skill of the model was diagnosed through 'validation' experiments [see *Sabbatelli and Mann*, 2007], wherein the full 137 year interval 1870–2006 was divided approximately into a first (1870–1938) and second (1939–2006) half. The statistical model was trained alternatively using either the first or second half, with TC counts predicted for the other (validation) sub-interval using only the climate state variables and the statistical relationship that was developed over the training interval. In these tests, the statistical model was found to successfully predict only slightly less ($R^2 = 43\%$; $p \ll 0.01$) of the annual TC variance than was nominally resolved above during calibration over the full 1870–2006 period. Similar results were achieved for a bivariate Poisson regression using only the

Figure 3. Atlantic TC counts (black curves) for each of the three undercount scenarios ('unadjusted,' 'lightly adjusted,' 'heavily adjusted') considered in the text, compared against predictions using the three-variable Poisson regression model. (a, b, c) Results where the regression model was trained (green curves) on the late 1944–2006 interval and predicted (red curves) for the early 1870–1943 interval. The actual mean counts for both intervals (horizontal black lines) and model-predicted mean for the early interval (horizontal red lines) are shown. (d, e, f) The results where training and prediction intervals have been switched.
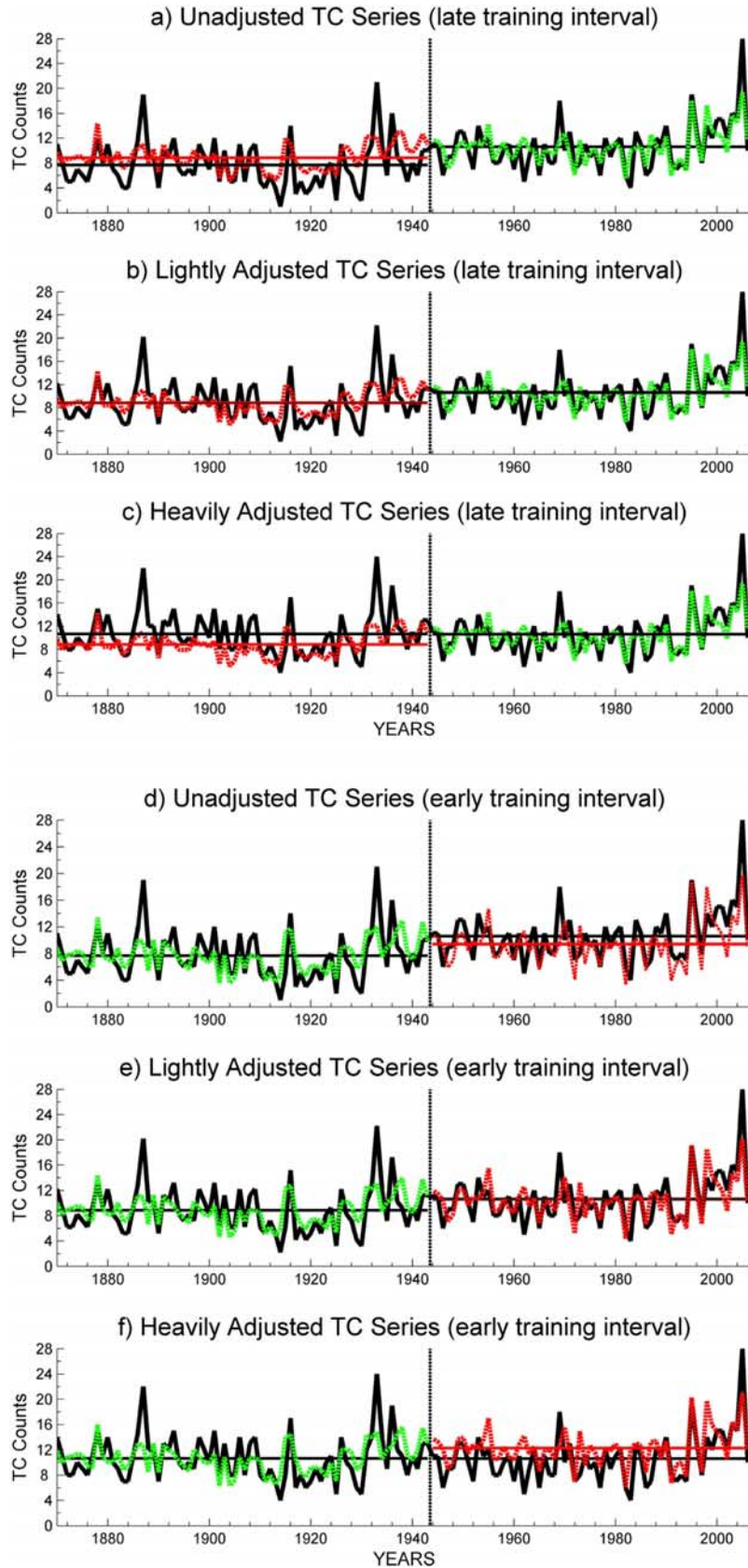
Figure 3

Table 1. Results of t Tests for Differences Between Predicted and Observed Mean Occurrence Rates μ for Early and Late Prediction Intervals as Discussed in Text[a]

	t	Φ	p
Observed vs. Predicted Occurrence Rate (1870–1943)			
Unadjusted Series	−3.71	73	0.0004
'Lightly Adjusted' Series	0.018		0.99
'Heavily Adjusted' Series	4.76		<0.0001
Observed vs. Predicted Occurrence Rate (1944–2006)			
Unadjusted Series	3.00	62	0.004
'Lightly Adjusted' Series	0.17		0.87
'Heavily Adjusted' Series	−4.11		0.0001

[a]Early, 1870–1943; late, 1944–2006. Indicated are degrees of freedom $\Phi = n − 1$ in the t statistic and the two-tailed p value for rejection of the null hypothesis of equal means.

MDR and Niño3.4 series (Figure 2), though the skill estimates were found to be slightly lower ($R^2 = 43\%$ for calibration and $R^2 = 36\%$ for validation).

[12] We subsequently used the statistical model to investigate the issue of potential TC undercount bias. Our underlying assumption is that a properly bias-corrected record of TC counts should yield long-term relationships between TC activity and climate that are consistent over time. We therefore investigated the implications of our model given alternative possible adjustments of annual TC counts reflecting varying levels of assumed undercount prior to the aircraft reconnaissance period.

3. Results

[13] We first analyzed the scenario of no undercount bias, using the unadjusted TC count series in our analyses. We found (Figure 3a) that training over the modern 1944–2006 interval yielded predictions over the pre-reconnaissance (1870–1943) interval (mean annual count = 8.84 TCs) that slightly over-predicted the 'observed' TC counts (mean annual count = 7.65 TCs). The significance of this difference, which suggests an undercount of roughly 1.2 TCs prior to 1944, can be estimated based on the null hypothesis of a fixed mean Poisson process. Under this null hypothesis, and recognizing that the underlying samples are large (∼100), the difference between predicted and observed mean count rates can be approximated at t-distributed [see e.g., *Wilks*, 2005]. Application of Student's t-test indicates that the difference, while modest, is highly significant (Table 1). In other words, the predicted and observed means are statistically inconsistent.

[14] Motivated by this finding, we next considered a scenario which assumes a very modest early undercount of 1.2 TCs per year during the pre-reconnaissance period by constructing a 'lightly adjusted' TC series in which the annual TC count was simply increased by 1.2 over the early (1870–1943) sub-interval. In this case, observed and predicted 1870–1943 means (8.85 and 8.84 respectively) are both visually (Figure 3b) and statistically (Table 1) indistinguishable.

[15] Finally, we considered a scenario which assumes a substantial early undercount of 3 TCs per year during the pre-reconnaissance period by constructing a 'heavily adjusted'

TC series in which the annual TC count was increased by 3 over the 1870–1943 sub-interval. In this case (Figure 3c) the actual mean count (10.65) is substantially higher than the predicted annual mean count (8.84), and the difference is highly significant (Table 1).

[16] Generalizing the analysis further, we established that assumed undercounts of either less than 0.53 or more than 1.91 yielded observed means that are statistically inconsistent with the predicted means (i.e., the null hypothesis of equal means can be rejected at the $p < 0.05$ level). Our analyses consequently suggest the average pre-1944* undercount to be between 0.5 and 2, with a most likely value of 1.2.

[17] While a potential criticism of our analysis is that our pre-reconnaissance statistical model estimates might be compromised by expanded uncertainties in the SST data used prior to World War II, we find this unlikely to be the case. Similar results were obtained using any one of the three individual SST products in place of the blended SST product used above (see auxiliary material). As the different SST products employ different mixes of in situ and remotely observed SST measurements, and make different assumptions about corrections for e.g. the early 20th century switch from bucket to ship intake measurements of seawater properties, our results appear to be robust with respect to uncertainties in SST data. Moreover, we obtained consistent results to those described above when the role of training and prediction period were switched in our analysis (Figures 3d, 3e, and 3f; Table 1). When the statistical model was trained on the early (1870–1943) interval for which the data quality is ostensibly poorer, the predicted TC counts for the late (1944–2006) interval slightly under-predicted observed counts with the unadjusted TC count series, over-predicted observed counts for the 'heavily adjusted' scenario, and yielded predicted and observed means that were statistically indistinguishable for the 'lightly adjusted' scenario. Similar results were also obtained (see auxiliary material) using (1) only the two predictors MDR SST and Niño3.4, (2) using Niño3 rather than Niño3.4 as a measure of ENSO, and (3) using an alternative choice of division (1939 rather than 1944) between the 'early' and 'late' sub-intervals of the TC record. We conclude that our results are robust with respect to uncertainties in climate data, and the other methodological details of our analysis.

4. Conclusions

[18] Our analyses indicate that an undercount in early TC counts approaching three storms per year is inconsistent with the observed statistical relationships between annual TC counts and the underlying climate factors that condition them. We conclude that the long-term record of historical Atlantic tropical cyclone counts is likely largely reliable, with an average undercount bias at most of approximately one tropical storm per year back to 1870. This conclusion supports other work [e.g., *Webster et al.*, 2005; *Emanuel*, 2005a; *Mann and Emanuel*, 2006] suggesting that increases in frequency, as well as powerfulness, of Atlantic TCs are potentially related to long-term trends in tropical Atlantic SST, trends that have in turn been connected to anthropogenic influences on climate [*Mann and Emanuel*, 2006; *Trenberth and Shea*, 2006; *Santer et al.*, 2006].

*The year is correct here. The article as originally published is online.

References

Bove, M. C., J. B. Elsner, C. W. Landsea, X. Niu, and J. J. O'Brien (1998), Effect of El Niño on U. S. landfalling hurricanes, revisited, *Bull. Am. Meteorol. Soc.*, *79*, 2477–2482.

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.

Elsner, J. B. (2003), Tracking hurricanes, Bull. Am, *Meteorol. Soc.*, *84*, 353–356.

Elsner, J. B., T. Jagger, and X.-F. Niu (2000), Changes in the rates of North Atlantic major hurricane activity during the 20th century, *Geophys. Res. Lett.*, *27*, 1743–1746.

Elsner, J. B., R. J. Murnane, and T. H. Jagger (2006), Forecasting U. S. hurricanes 6 months in advance, *Geophys. Res. Lett.*, *33*, L10704, doi:10.1029/2006GL025693.

Emanuel, K. (2005a), Increasing destructiveness of tropical cyclones over the past 30 years, *Nature*, *436*, 686–688.

Emanuel, K. (2005b), Emanuel replies, *Nature*, *438*, E13.

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

Gray, W. M. (1968), Global view of the origins of tropical disturbances and storms, *Mon. Weather Rev.*, *96*, 669–700.

Gray, W. M. (1984), Atlantic seasonal hurricane frequency, Part I: El Niño and 30 mb quasi-biennial oscillation influences, *Mon. Weather Rev.*, *115*, 1649–1668.

Holland, G. J. (2007), Misuse of landfall as a proxy for Atlantic tropical cyclone activity, *Eos Trans. AGU*, *88*, 349.

Holland, G. J., and P. J. Webster (2007), Heightened tropical cyclone activity in the North Atlantic: Natural variability or climate trend?, *Philos. Trans. R. Soc. Ser. A*, *365*, 2695–2716, doi:10.1098/rsta.2007.2083.

Hoyos, C. D., P. A. Agudelo, P. J. Webster, and J. A. Curry (2007), Deconvolution of the factors contributing to the increase in global hurricane intensity, *Science*, *312*, 94–97.

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, report, NOAA Natl. Hurricane Cent., Miami, Fla.

Jones, P. D., T. Jónsson, and D. Wheeler (1997), Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland, *Int. J. Climatol.*, *17*, 1433–1450.

Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan (1998), Analyses of global sea surface temperatures: 1856–1991, *J. Geophys. Res.*, *103*(C9), 18,567–18,589.

Landsea, C. W. (2005), Hurricanes and global warming, *Nature*, *438*, E11–E13.

Landsea, C. W. (2007), Counting Atlantic tropical cyclones back to 1900, *Eos Trans. AGU*, *88*(18), 197.

Landsea, C. W., C. Anderson, N. Charles, G. Clark, J. Dunion, J. Fernandez-Partagas, P. Hungerford, C. Neumann, and M. Zimmer (2004), The Atlantic hurricane database re-analysis project: Documentation for the 1851–1910 alterations and additions to the HURDAT database, in *Hurricanes and Typhoons: Past, Present and Future*, edited by R. J. Murname and K.-B. Liu, pp. 177–221, Columbia Univ. Press, New York.

Mann, M. E., and K. A. Emanuel (2006), Atlantic Hurricane trends linked to climate change, *Eos Trans. AGU*, *87*(24), 233.

Neumann, C. J., B. R. Jarvinen, C. J. McAdie, and J. D. Elms (1999), Tropical cyclones of the North Atlantic Ocean, 1871–1998, *Hist. Climatol. Ser. 6-2*, Natl. Clim. Data Cent., Asheville, N. C.

Nyberg, J., B. A. Malmgren, A. Winter, M. R. Jury, K. H. Kilbourne, and T. M. Quinn (2007), Low Atlantic hurricane activity in the 1970s and 1980s compared to the past 270 years, *Nature*, *447*, 698–702.

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*(D14), 4407, doi:10.1029/2002JD002670.

Sabbatelli, T. A., and M. E. Mann (2007), The influence of climate state variables on Atlantic Tropical Cyclone occurrence rates, *J. Geophys. Res.*, *112*, D17114, doi:10.1029/2007JD008385.

Santer, B. D., et al. (2006), Forced and unforced ocean temperature changes in Atlantic and Pacific tropical cyclogenesis regions, *Proc. Natl. Acad. Sci. U. S. A.*, *203*, 13,905–13,910.

Saunders, M. A., and A. R. Harris (1997), Statistical evidence links exceptional 1995 Atlantic hurricane season to record sea warming, *Geophys. Res. Lett.*, *24*, 1255–1258.

Shapiro, L. J. (1982), Hurricane climatic fluctuations. Part II: Relation to large-scale circulation, *Mon. Weather Rev.*, *110*, 1014–1023.

Shapiro, L. J., and S. B. Goldenberg (1998), Atlantic sea surface temperatures and hurricane formation, *J. Clim.*, *11*, 578–590.

Smith, T. M., and R. W. Reynolds (2003), Extended reconstruction of global sea surface temperatures based on COADS data (1854–1997), *J. Clim.*, *16*, 1495–1510.

Sriver, R., and M. Huber (2006), Low frequency variability in globally integrated tropical cyclone power dissipation, *Geophys. Res. Lett.*, *33*, L11705, doi:10.1029/2006GL026167.

Trenberth, K. E., and D. J. Shea (2006), Atlantic hurricanes and natural variability in 2005, *Geophys. Res. Lett.*, *33*, L12704, doi:10.1029/2006GL026894.

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

Wilks, D. S. (2005), *Statistical Methods in the Atmospheric Sciences: An Introduction*, 2nd ed., 520 pp., Academic Press, New York.

————————————
M. E. Mann and T. A. Sabbatelli, Department of Meteorology and Earth and Environmental Systems Institute, Pennsylvania State University, 523 Walker Building, University Park, PA 16802, USA. (mann@psu.edu)

U. Neu, ProClim, Swiss Academy of Sciences, Schwarztorstrasse 9, CH-3007 Bern, Switzerland.