



Supporting Online Material for

Modeled Impact of Anthropogenic Warming on the Frequency of Intense Atlantic Hurricanes

Morris A. Bender,* Thomas R. Knutson, Robert E. Tuleya, Joseph J. Sirutis, Gabriel A. Vecchi, Stephen T. Garner, Isaac M. Held

*To whom correspondence should be addressed. E-mail: Morris.Bender@noaa.gov

Published 22 January 2010, *Science* **327**, 454 (2010)
DOI: 10.1126/science.1180568

This PDF file includes:

Materials and Methods
SOM Text
Figs. S1 to S9
Table S1
References

SUPPORTING ONLINE MATERIAL FOR:

Modeled Impact of Anthropogenic Warming on Intense Atlantic Hurricane Frequency

MATERIALS AND METHODS

As outlined in (16), the control ZETAC regional model simulation is forced by observed SSTs and nudged on large spatial scales toward a spatially smoothed component of the NCEP/NCAR reanalysis I data (*S1*), while still enabling tropical cyclones to develop within the model domain. In the ZETAC simulations for the warmed climates, the seasonal mean SSTs and large-scale seasonal mean climate to which the model is nudged are modified by the anomalies projected by the global climate model, or model ensemble, being downscaled. For the present study, each tropical cyclone produced in these regional simulations was downscaled to the GFDL or GFDN operational hurricane model. These hurricane models are triply nested moveable mesh models with minimum grid spacing of 8 km centered near the storm. Model details have been summarized in previous studies (*S2*). The hurricane models were run for 120 hours for each storm case, beginning three days before the storm in the parent regional model integration reached maximum intensity. If a storm did not exist or was weaker than tropical storm force as of 72h before it reached maximum intensity, the downscaling simulation began at a later time when the cyclone had attained at least tropical storm intensity. We assumed that the 120 hour period is sufficient to capture the period in which maximum intensity would likely be attained.

To evaluate the soundness of this experimental design, the histogram of the maximum wind speed for each observed storm during the 2006-2009 Atlantic seasons was plotted and compared with the operational GFDL forecast run for the time period 72h before the observed maximum intensity (Fig. S1). The close agreement between the modeled and observed histograms increases our confidence in the validity of this approach.

Unlike the operational forecast system, no synthetic vortex replacement initialization method was utilized; the model used ZETAC fields directly as initial conditions. The same procedure was applied uniformly to both control and warm-climate storms. The ocean model in the GFDL hurricane prediction system for the control runs was initialized from the Generalized Digital Environmental Model (GDEM) monthly ocean temperature and salinity climatology (*S3*). For the warm climate runs, the ocean structure at each vertical level was modified by the temperature change estimated by the 18 climate model ensemble, while keeping the salinity unchanged. Each of the simulated storms was run separately using the GFDL (NWS) and GFDN (Navy) versions of the coupled hurricane prediction system. The change in occurrence of major hurricanes is shown in Figure S5 for the combined GFDL and GFDN. The overall distribution of intense hurricane activity in both the control and warmed climate runs showed relatively little sensitivity to the use of the two hurricane model versions (*Fig. S7*).

In the hurricane model experiments, the mean intensity change (using one maximum surface wind speed value per storm) was -1.2% for all tropical storms and +0.7% for storms reaching hurricane intensity. However, the change in the wind speed distribution (*Fig. 1e*) is best described as a change in its shape (with more weak and very strong systems, and with fewer storms overall) rather than a simple rightward or leftward shift of the distribution.

We have attempted to address concerns about robustness of results and the separation of radiatively forced climate change responses from internal variability through several experimental design choices: (1) We have run multi-season experiments to expand our sample size. For the CMIP3 18-model ensemble changes and control, we have run 27 separate seasons, and for the individual CMIP3 model downscaling experiments, 13 seasons (only odd years, in order to reduce computing requirements). (2) We have combined results from two separate versions of the hurricane model, which use different physics settings, to test for robustness. (3) We use a large climate change signal (80 yr forcing under IPCC Scenario A1B) to enhance our ability to distinguish the

model's climate change signal from internal variability noise. Although all of the models use the IPCC A1B scenario, there are differences between models in the radiative forcing resulting from this scenario. (4) We compare the results from the 18-model CMIP3 composite climate change signal with results using climate change perturbations from four individual CMIP3 models. (5) For the individual CMIP3 model results, we used the linear trend over 2000-2100 for each model, scaled to 80 yr magnitude, to compute the climate change perturbation. This linear trend reduces the influence of internal variability on the climate change signal derived from the individual model realizations. Such multi-decadal variability in the models could have confounded our results if we had simply used pairs (control and perturbation runs) of 20-yr segments to define the climate change perturbation. (6) For the CMIP3 18-model ensemble tests, the climate change signal was derived as the difference between the 20-yr periods 2000-2019 and 2080-2099, averaged across all of the 18 models. In this case, since the multi-model ensemble averaging itself strongly reduces the influence of internal multi-decadal variability from the individual models, the linear trend technique that was used for the individual models was not needed.

Supporting Text

In this section we discuss our method of estimating an “emergence time scale” for the anthropogenically forced increase in Category 4-5 hurricanes as simulated in our modeling study (A1B scenario). A number of complicating factors preclude us from drawing definitive conclusions, but a discussion of these calculations and associated issues provides some guidance on interpreting our model results. The complicating factors include the lack of a confident estimate of internal variability levels for Atlantic Cat 4-5 hurricane counts; uncertainties in data quality for historical counts since 1944; and uncertainties in our future projected Cat 4-5 climate change signal. These are further discussed in the main text and in the following supplemental text. The analysis below is based only on the 18-model CMIP3 ensemble results (combined GFDL and GFDN

downscalings). The model sensitivity discussed in the main text should be kept in mind when interpreting the results.

According to our model, the number of Category 4-5 hurricanes is projected to increase by 81% over 80 years (Table 1), which we approximate here as a future linear trend projection of +10.1% per decade (hereafter referred to as the “signal”). Since our model under-predicts the observed number of Cat 4-5 storms by more than a factor of two (Table S1), we use the model’s percent change projection rather than absolute change projection, assuming that the absolute change would likely be biased low by the model’s low bias in absolute numbers in the control. The percent change is then applied to a baseline number of storms from observations, which we take as the long-term mean over 1944-2008 (1.34/yr) yielding a projected trend of 0.136/yr/decade. Alternatively, as a sensitivity test, we use the 2008 endpoint of a linear trend over 1944-2008 (Fig. S9) as the base line value (1.86/yr) yielding a projected trend of 0.188/yr/decade.

We define an “emergence time scale” T , as a length of time in years such that computation of trends for a large sample of plausible future Cat 4-5 series of length T or greater, based on our model projected change, yields a sample of trends with less than 5% having negative values. As such, the emergence time scale of the projected climate change signal (10.1%/yr/decade) depends on the noise in which the signal is embedded. At present, we do not have long control run estimates of category 4-5 hurricane variability from our climate models; even if available, it is unclear that these would be reliable. Therefore, in order to roughly estimate the emergence time scale, we use various filtered versions of the historical record of category 4-5 hurricane counts to derive plausible estimates of the internal variability noise.

Our base case estimate uses the observed Category 4-5 hurricane counts from 1944 to 2008 (Figure S9) directly (with no further filtering) to construct synthetic noise series. By re-sampling these 65 observed values (bootstrap with replacement) we can construct synthetic Atlantic Cat 4-5 hurricane time series of arbitrary length. To each of these synthetic noise series, we add a linear trend of 0.136 storms/yr/decade, and then create a distribution of the net trends. We created 1000 such synthetic time series for

each time horizon tested, and tested horizons of length 20 years and longer. For our base case assumptions, the emergence time scale was estimated as ~60 yr. Using the alternative baseline value (1.86/yr, the observed linear trend evaluated at 2008) shortens the emergence time to about 50 yr or about 10 yr shorter than our base case. We have also performed our trend tests for our distributions using storm counts to which a square root transformation has been applied—to reduce the impact of skewness in the time series—and find that the results are not strongly sensitive to whether this procedure is used or not

One could argue that sampling the raw observed annual values in Fig. S9 overestimates the true internal variability levels in the basin, leading to an overestimate of the emergence time. To test this possibility, we examined an alternative viewpoint (6) which holds that the multi-decadal variations in the Atlantic (such as in the observed Cat 4-5 in Fig. S9) are due mostly to changes in radiative forcing (e.g., greenhouse gases, and volcanic and anthropogenic aerosols). From this viewpoint (6), the period of low hurricane activity in the 1970s and 80s was due primarily to strong aerosol forcing, and we might therefore expect less of such multi-decadal variability in the future than during the past 60 years, apart from an accelerated quasi-linear warming due to increasing greenhouse gases. In that case, alternative synthetic noise time series can be obtained by randomly re-sampling the elements of a residual time series formed by removing a smoothed background variation, such as a 4th order polynomial, from the data series in Figure S9. To each of 1000 such random series formed from these residuals, we added a linear trend of 0.136 (storms/yr)/decade. This resulted in about a 5-yr shorter emergence time (i.e., ~55yr) than our base case, suggesting that our results are not strongly dependent on this aspect of the noise assumptions used.

Our resampling procedure creates synthetic noise time series which by design will have essentially no systematic temporal correlation, whereas a number of studies have reported that the Atlantic basin may have pronounced multi-decadal variability associated with the so-called Atlantic Multi-decadal Oscillation, or AMO (4, S4). To explore the possible impacts of autocorrelated noise in the internal variability series, we tried as a

sensitivity test a variant on the bootstrap resampling procedure (*S9*) in which we sampled “chunks” of the original observed time series rather than individual years, as done in the base case, in order to build up the synthetic noise series. We used chunk or segment lengths of 3 and 7 yr in separate tests and found that the emergence time increased by 5-10 yr beyond the base case (i.e. to 65-70 yr).

In separate tests, we have found that the presence of large multi-decadal (~60 yr) internal variability could substantially affect the detection time scale results, particularly depending upon its phase at the beginning of the series examined. For example, if a trend analyses were begun near a minimum (maximum) phase of a ~60 yr multidecadal variability “event”, the emergence time obtained use our idealized procedure might be considerably shortened (lengthened). Although Figs. 1c and *S9* and ref *S4* suggests that there may be substantial low-frequency internal variability of intense Atlantic hurricanes, unfortunately a robust separation of the radiatively forced and internal variability components of this variability is not yet available (*4*), leading us to rely on the simpler noise assumptions used here for our preliminary analysis. We suspect, however, that if the multi-decadal internal variability of hurricanes is quite pronounced in the Atlantic basin as suggested by some data analyses (see Figs. 1c, *S9* and ref *S4*), then our emergence time scale estimates are probably lower bounds for the true values.

The actual emergence time (calendar year) may be decreased by a few decades from the above estimates through the inclusion of historical data, as there has already been some warming and presumably some anthropogenic contribution toward a detectable trend in the Cat 4-5 frequency. However, this would require the historical Cat 4-5 data (1944-2008) to be carefully assessed for homogeneity problems and the contributions of various influences on past Cat 4-5 variability (e.g., greenhouse gases, aerosols, internal climate variability, etc.) to be quantified, which are tasks beyond the scope of our study.

In summary, our preliminary analysis suggests an emergence time scale for our model’s projected 21st century trend in Cat 4-5 numbers (~10%/yr/decade) of roughly 60 yr, which suggests that the signal may emerge from climate variability noise in the latter

half of the 21st century. This preliminary estimate is subject to the caveats discussed above and in the main text, including the climate model dependence of our projected climate change signal, limitations of the hurricane model, and uncertainties in both the past observations of Cat 4-5 numbers and their internal multidecadal variability levels.

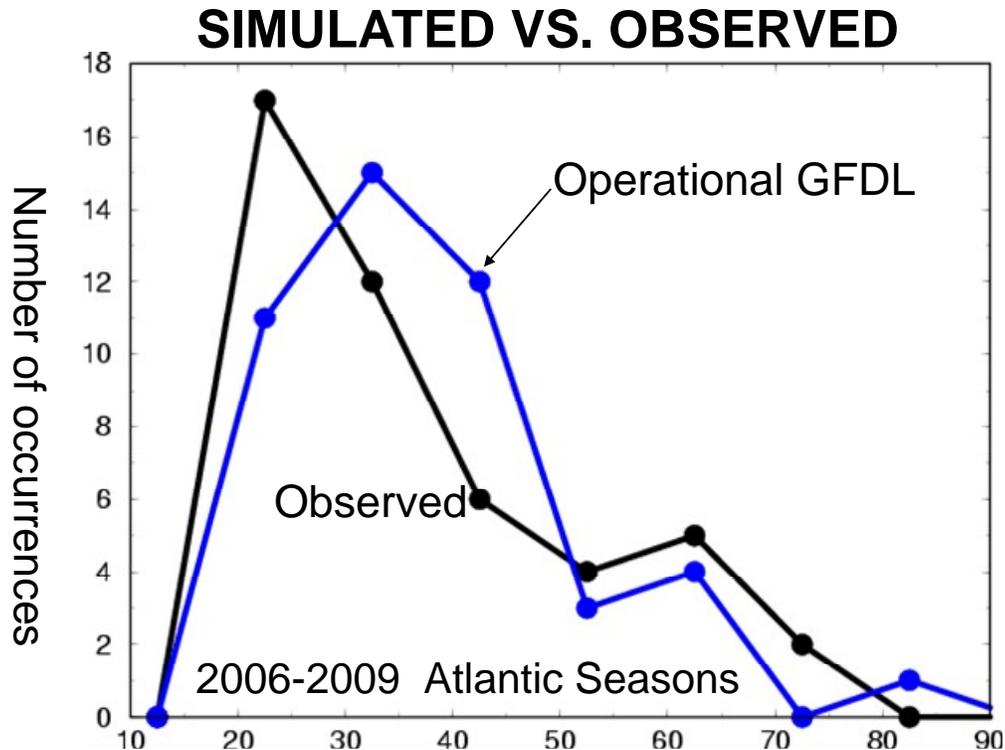


Figure S1 Histogram of maximum surface wind speed (m/s) in the Atlantic basin for each observed storm that reached at least tropical storm strength, during the 2006-2009 hurricane seasons (black line), compared to the operational GFDL hurricane model forecast made three days before the observed maximum intensity (blue line). This three-day lead time comparison is presented here, as it mimics the procedure (starting three-day prior to maximum intensity) used in our hurricane/climate downscaling studies.

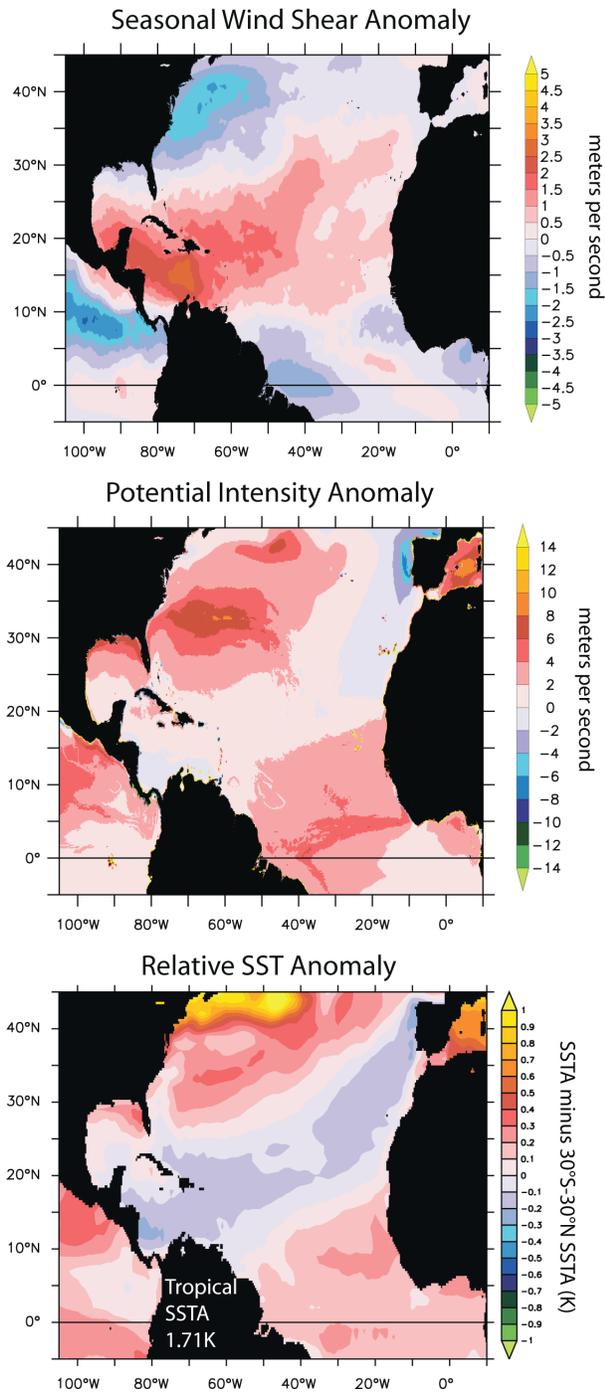


Figure S2 Seasonally averaged wind shear (top), potential intensity (middle) changes (warm climate minus control) and relative SST anomaly, computed from the regional simulations of the ZETAC model, averaged for the 13 odd years (1981 to 2005) used in this study. The anomaly fields are based on the warm climate conditions provided from the 18-model CMIP3 ensemble.

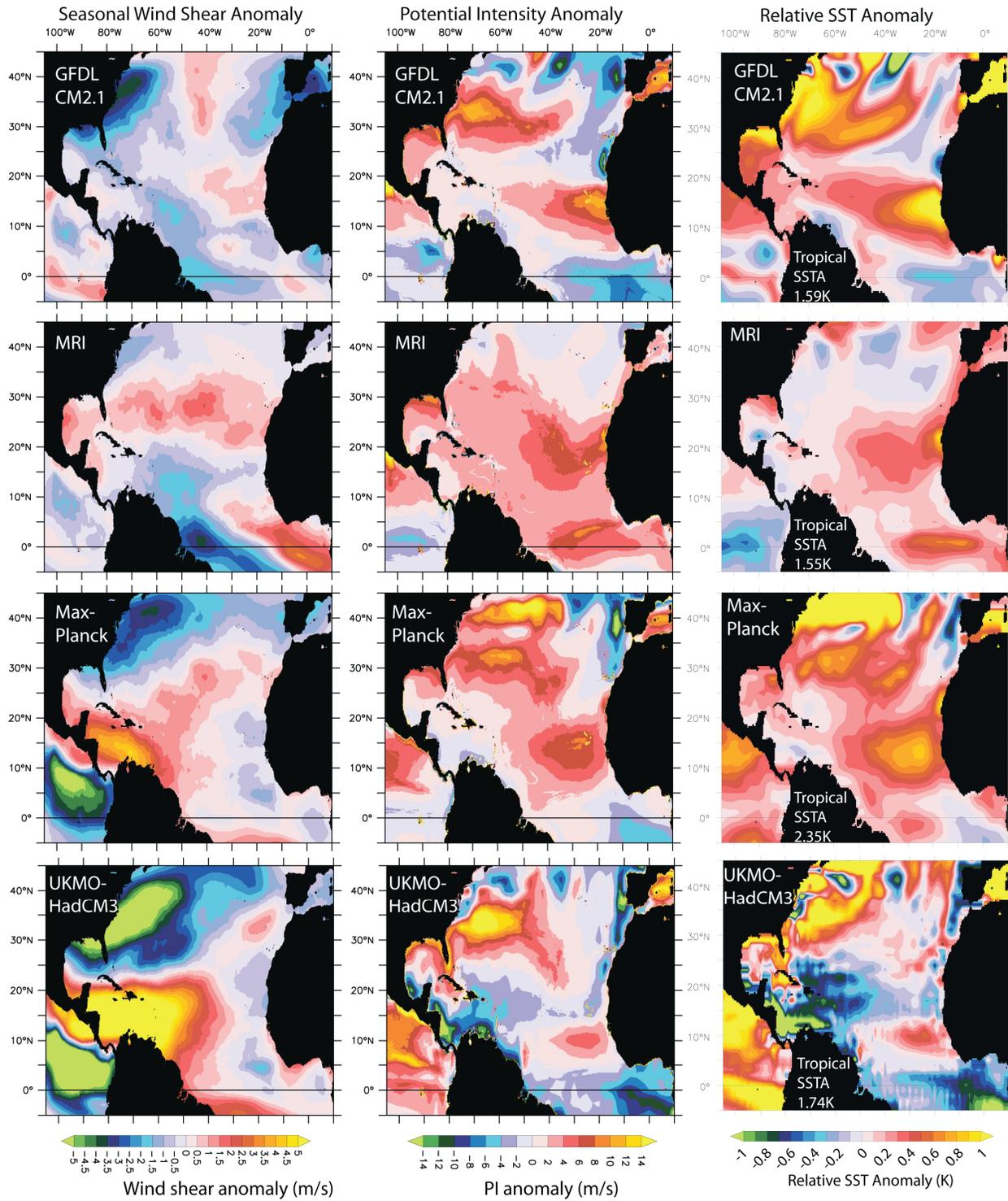


Figure S3 Same as S2 but for the four individual CMIP3 models.

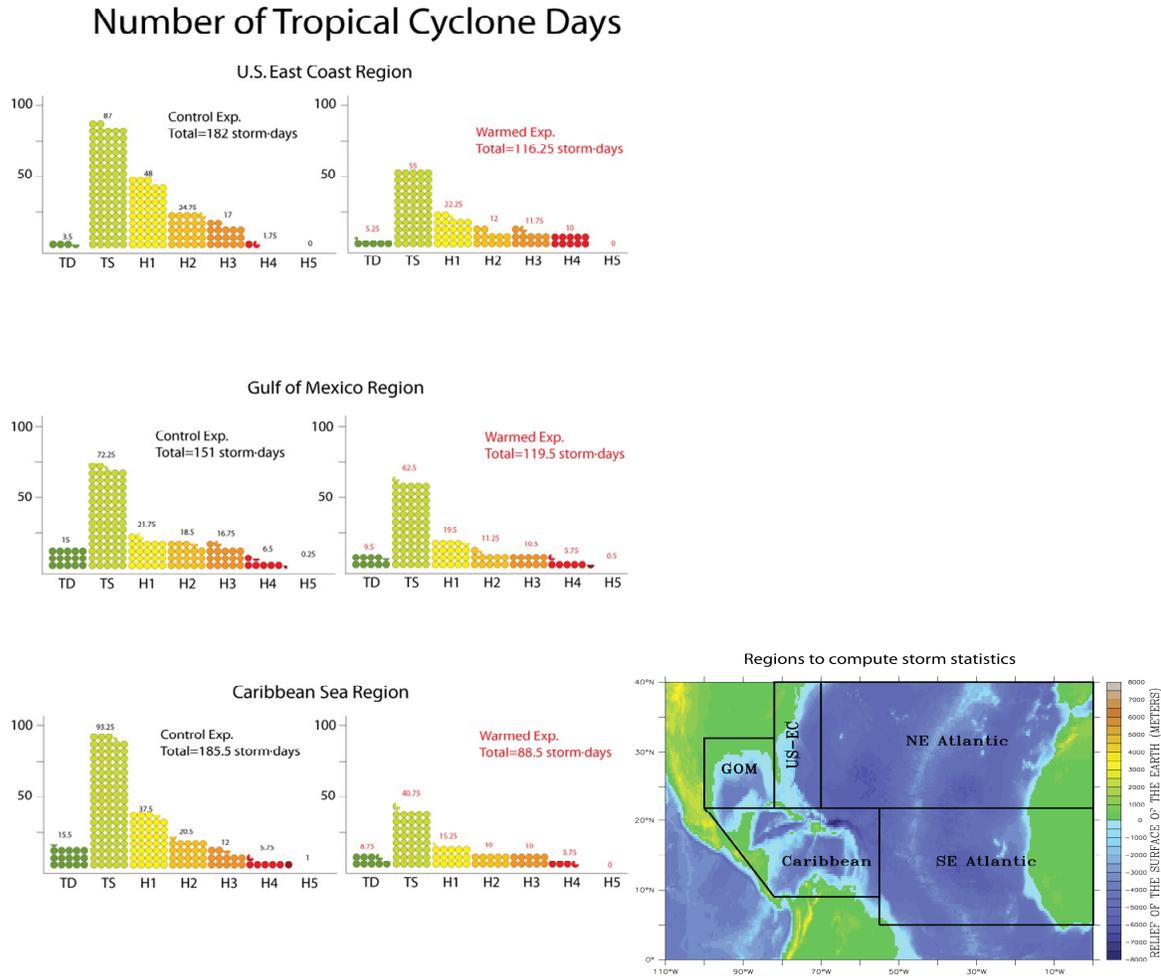


Figure S4 The total number of storm days (GFDL and GFDN models combined) plotted for the seven categories of intensity (Depression, Tropical Storm, category 1, 2, 3, 4 and 5 hurricanes), computed for the control and the 18-model ensemble warm climate scenario. The statistics were computed for three regions of the western Atlantic (Gulf of Mexico, Caribbean, and the U.S. East Coast region (see inset map)).

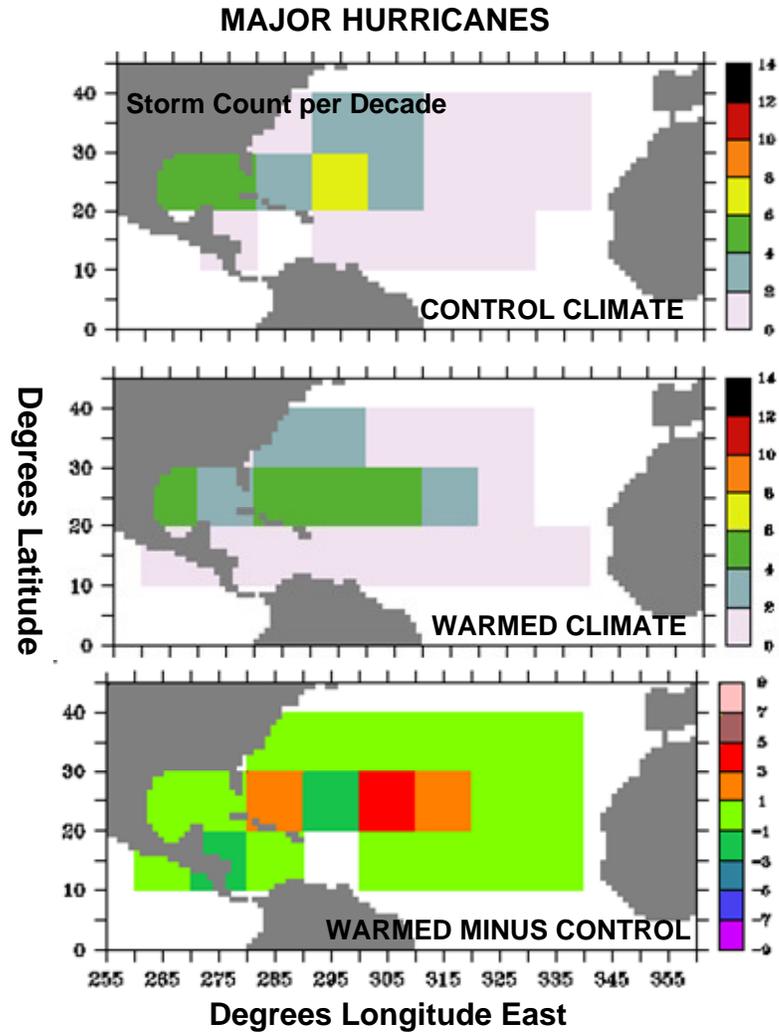


Figure S5. The spatial distribution of major hurricane (Cat 3-5) occurrence (scaled by storm counts per decade) for the combined (GFDL/NWS and GFDN/Navy) control runs (top), the combined GFDL and GFDN 18-model CMIP3 ensemble warmed climate results (middle), and the difference between the warmed climate and control Cat 3-5 hurricane occurrence (bottom).

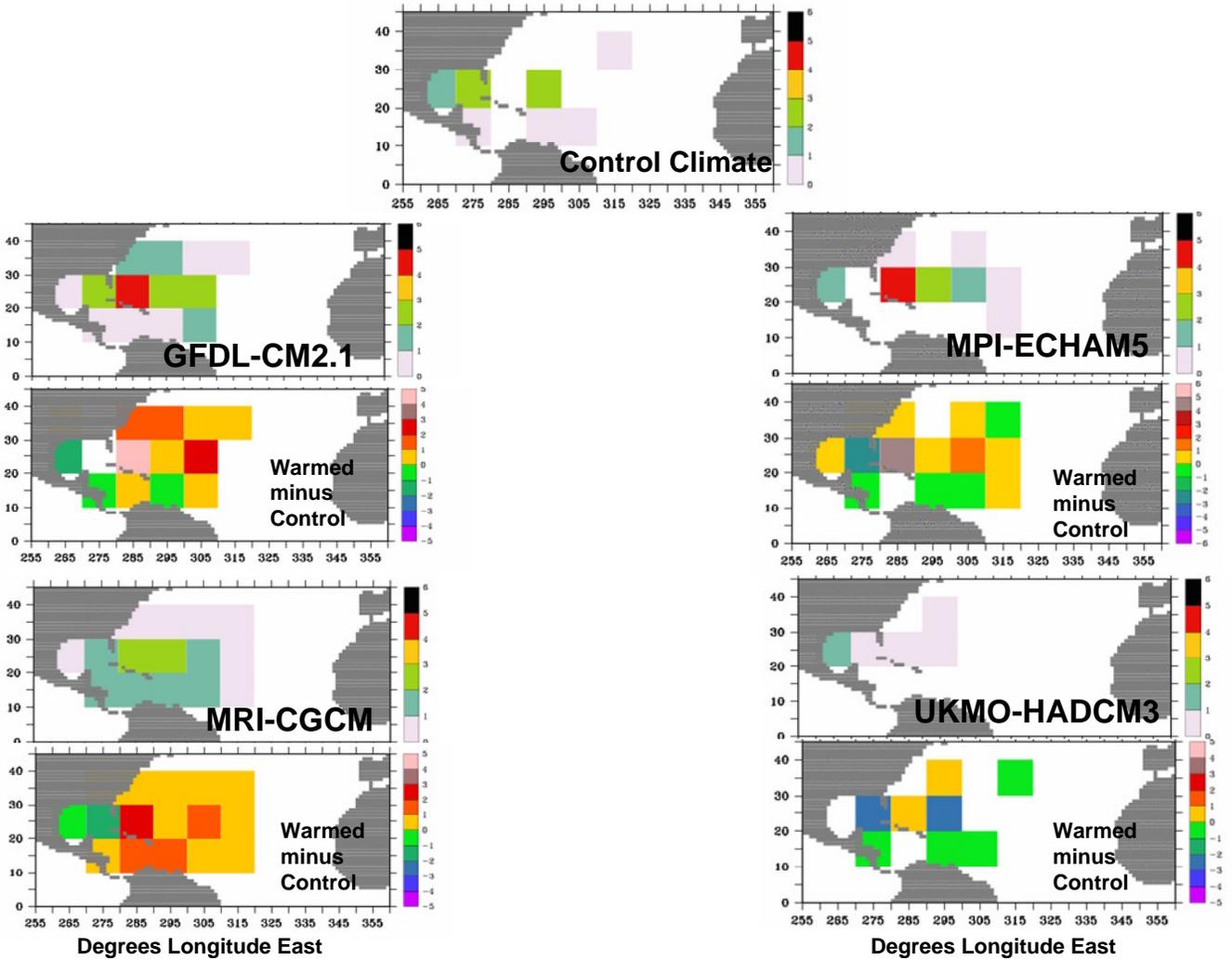


Figure S6 The spatial distribution of Cat 4-5 hurricane occurrence (scaled by storm counts per decade) for the combined control (GFDL/NWS and GFDN/Navy versions), and the occurrence maps from the integrations using the climate conditions provided by the four individual CMIP3 models used in this study. The differences in the Cat 4-5 occurrence between the control and the four warmed scenarios are also presented.

Tracks of Storms that Reached Category 4 or 5 Intensity

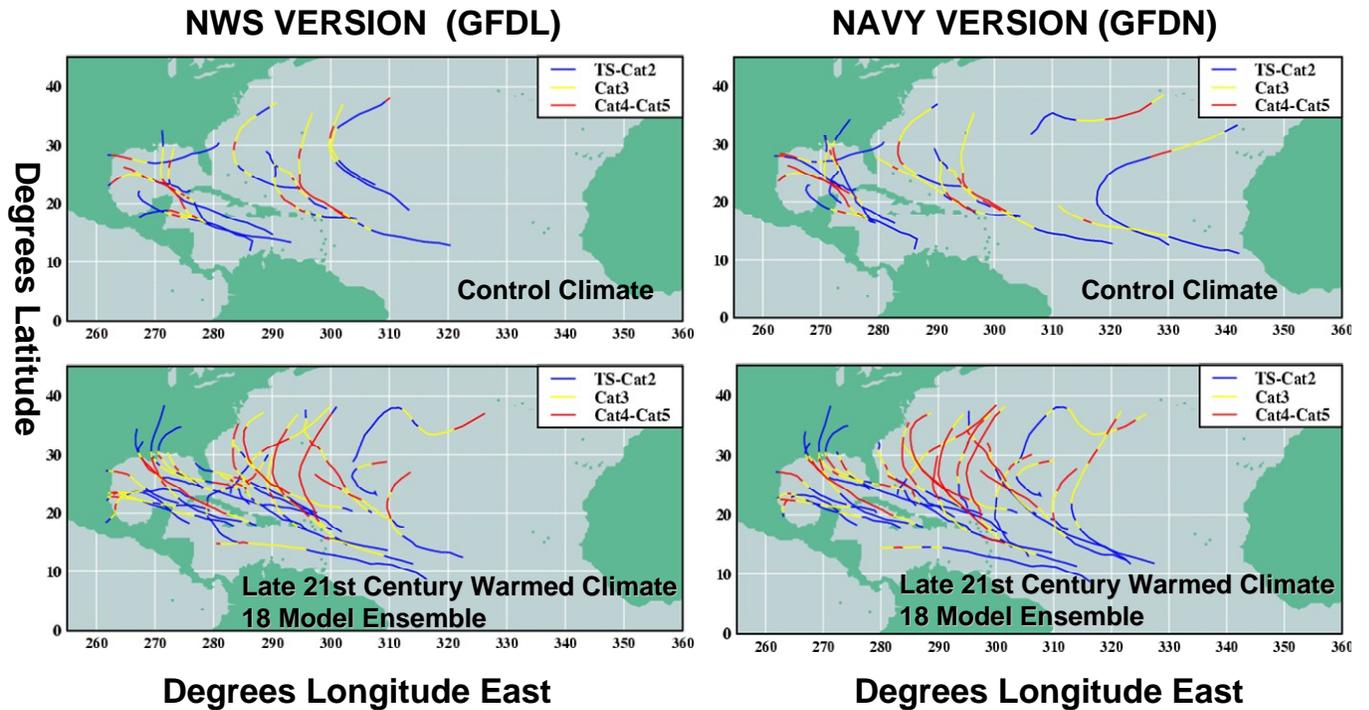


Figure S7 Comparison of the tracks for all storms reaching category 4 or 5 intensity, for the operational version of the GFDL model currently used by NOAA's NWS (left) compared with the GFDN version run by the U.S. Navy (right).

Tracks of Storms that Reached Category 4 or 5 Intensity

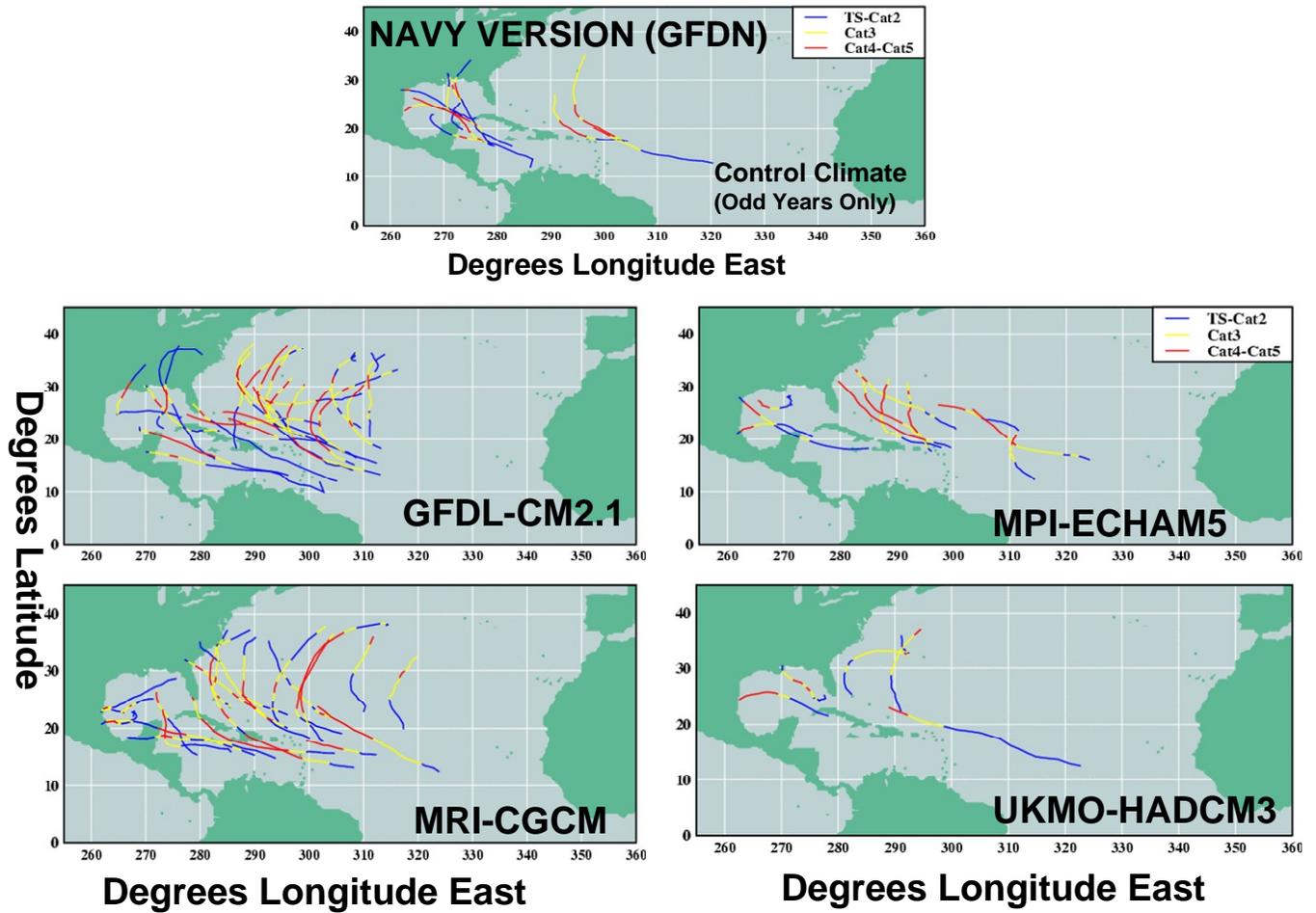


Figure S8 Comparison of the tracks for all storms reaching category 4 or 5 intensity, for the control (top) and the warm climate conditions provided by the 4 CMIP3 climate models, using the GFDN model, which is the operational version of the GFDL model currently used by the U.S. Navy.

Number of Cat 4+5 Atlantic Hurricanes With Emanuel Adjustment for Early Storm Intensities

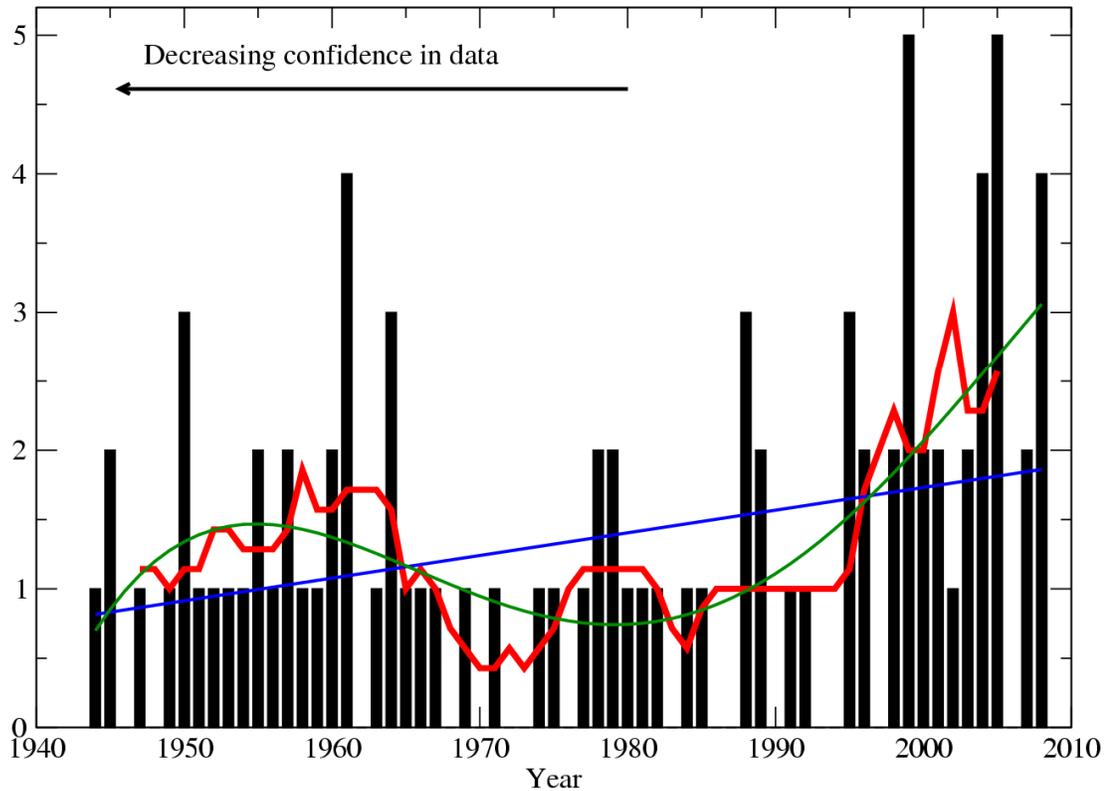


Figure S9 Observed time series of Cat 4-5 hurricane counts from the 1944 through the 2008 hurricane seasons as contained in the Atlantic HURDAT database (S6). The blue line indicates the linear trend fit. The red line is the 7 year running mean. The green line is a fourth order polynomial fit to the data (see text). Some intensities for 1944-1969 were adjusted downward according to ref 22 prior to determining the category of each storm.

	Number per Year:			Percent Changes (warm vs. control)						
Storm Category	Observed (1980-2006)	Control (1980-2006)	Control (odd years)	CMIP3 Ensemble (1980-2006)	CMIP3 Ensemble (odd years)	GFDL CM2.1 (odd years)	MRI (odd years)	MPI (odd years)	HadCM3 (odd years)	Percent of Damage by Category
Tropical Storm	3.70	2.83	2.65	-13	-19	+4	-16	-14	-14	2
Cat 1	1.89	3.41	3.31	-52	-62	-40	-45	-48	-66	5.1
Cat 2	1.04	1.91	1.81	-17	+4	-15	-28	-36	-53	7.4
Cat 3	1.00	2.13	1.81	-45	-55	+9	-34	-51	-64	37.2
Cat 4	1.00	0.56	0.69	+83	+72	+100	+72	+17	-56	41.1
Cat 5	0.37	0.02	0.04	+200	+100	+400	+800	+100	0	7.3
All Storms	9.00	10.85	10.31	-27	-28	-4	-22	-33	-49	
Percent Change in Damage Potential				+28	+13	+71	+71	-10	-54	

Table S1. Storm statistics from the hurricane model downscaling experiments (average of GFDL/NWS and GFDN/Navy versions). Numbers of storms by category and the percentage change (warm climate runs vs. control runs) by category are shown in columns 2-4 and 5-10, respectively. The far right column shows percent of total historical U.S. landfalling hurricane damage by storm category at landfall from ref 19. The bottom row shows the percent change in damage potential, estimated by combining the percent of historical damage by storm category with the percentage change (warm minus control) in the occurrence rate of each storm class (by category). This type of projection neglects important influences, such as future changes in population, wealth, damage mitigation practices, etc. onto which any climate-related changes would be superimposed.

Supplemental References

- S1. E. Kalnay *et al.*, *Bull Am. Meteorol. Soc.*, **77**, 437 (1996).
- S2. Y. Kurihara, R. E. Tuleya, M. A. Bender, *Mon. Wea. Rev.*, 126, 1306 (1998).
- S3. W.J. Teague, M. J. Carron, P.J. Hogan, *Geophys. Res.*, **95**, 7167 (1990).
- S4. S.B. Goldenberg, C. W. Landsea, A. M. Mestas-Nunez, W.M. Gray, *Science*, **20**, 474 (2001).
- S5. M. Bister, K.A. Emanuel, *Meteor. Atm. Phys.*, **52**, 233-240 (1998).
- S6. C. W. Landsea, C. Anderson, N. Charles, G. Clark, J. Dunion, J. Fernandez-Partagas, P. Hungerford, C. Neumann, M. Zimmer, In “Hurricanes and Typhoons: Past, Present and Future”, R. J. Murname and K.-B. Liu, Eds., Columbia University Press, 177-221, (2004).
- S7. D. S. Wilks, *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Elsevier Academic Press, 627 pp. (2006).