U.S. Tropical Cyclone Activity in the 2030s Based on Projected Changes in Tropical Sea Surface Temperature

TIMOTHY M. HALL,[a] JAMES P. KOSSIN,[b] TERENCE THOMPSON,[c] AND JAMES MCMAHON[c]

[a] *NASA Goddard Institute for Space Studies, New York, New York*
[b] *NOAA/National Centers for Environmental Information, Madison, Wisconsin*
[c] *The Climate Service, Durham, North Carolina*

(Manuscript received 14 May 2020, in final form 11 November 2020)

ABSTRACT: We use a statistical tropical cyclone (TC) model, the North Atlantic Stochastic Hurricane Model (NASHM), in combination with sea surface temperature (SST) projections from climate models, to estimate regional changes in U.S. TC activity into the 2030s. NASHM is trained on historical variations in TC characteristics with two SST indices: global–tropical mean SST and the difference between tropical North Atlantic Ocean (NA) SST and the rest of the global tropics, often referred to as ''relative SST.'' Testing confirms the model's ability to reproduce historical U.S. TC activity as well as to make skillful predictions. When NASHM is driven by SST projections into the 2030s, overall NA annual TC counts increase, and the fractional increase is the greatest at the highest wind intensities. However, an eastward anomaly in mean TC tracks and an eastward shift in TC formation region result in a geographically varied signal in U.S. coastal activity. Florida's Gulf Coast is projected to see significant increases in TC activity relative to the long-term historical mean, and these increases are fractionally greatest at the highest intensities. By contrast, the northwestern U.S. Gulf Coast and the U.S. East Coast will see little change.

KEYWORDS: North America; Tropical cyclones; Risk assessment; Probability forecasts/models/distribution; Stochastic models

1. Introduction

Tropical cyclones (TCs) are among the most hazardous natural catastrophes to impact the United States (Smith and Katz 2013). How this hazard evolves with climate variation, both natural and anthropogenic, is of crucial importance to coastal populations and is an extremely active area of research. Other factors being constant, coastal storm-surge hazard is certain to rise with sea level (Tebaldi et al. 2012), TC rainfall is projected to increase with high confidence (Kossin et al. 2017; Knutson et al. 2020), and the interaction of freshwater flooding and storm surge will be exacerbated (Moftakhari et al. 2017). Theory indicates that a warming climate causes TC wind intensity to increase, especially at the highest intensity levels, a result borne out by numerical model simulations (Villarini and Vecchi 2013; Sobel et al. 2016; Knutson et al. 2020) and observations (Elsner et al. 2008; Kossin et al. 2013, 2020; Knutson et al. 2019).

At least as important as intensity for coastal hazard is evolution in the geographic distribution of TC landfall via changes in formation and storm track, though the nature and degree of these changes are considerably more uncertain than sea level rise, rainfall, and intensity (Knutson et al. 2020). Most model studies find little change or decreases in global TC annual frequency (Yoshida et al. 2017; Murakami et al. 2020; Knutson et al. 2020), though the variation across studies is greater in individual ocean basins. However, several studies find an eastward shift in North Atlantic Ocean (NA) TC formation region toward the eastern tropics (Murakami and Wang 2010; Wang et al. 2011; Colbert et al. 2013; Yoshida et al. 2017). There is also evidence for eastward or northeastward TC track anomalies (Wang et al. 2011; Colbert et al. 2013; Baldini et al. 2016). In addition to changes in basinwide TC rates, both the formation region and track changes would play a crucial role in determining regional changes in U.S. landfall rates.

Regional and local spatial scales are of primary importance to disaster and adaptation planning, and it is of considerable interest to estimate the impact at these scales of global and basinwide relationships between climate and TC activity. Here, we use the North Atlantic Stochastic Hurricane Model (NASHM), a statistical–stochastic model of NA TCs and their climate variation (Hall and Jewson 2007; Hall and Yonekura 2013), along with projections of climate variation from an ensemble of climate models from phase 5 of the Coupled Model Intercomparison Project (CMIP5) (Taylor et al. 2012), to estimate U.S. regional TC activity into the 2030s. This time scale is highly relevant to a number of sectors, such as infrastructure development and the insurance and reinsurance industry. NASHM is able to replicate historical U.S. coastal TC activity on regional scales, as well as climate-induced variation of that activity (see section 2). It therefore provides a method to project large-scale climate variations onto regional TC activity.

We represent climate variation by two annually varying sea surface temperature (SST) indices: 1) global tropical SST (GSST), averaged over all longitudes and during the NA hurricane season; and 2) ''relative SST'' (RSST), the NA main development region (MDR) SST minus the tropics over all longitudes outside the Atlantic. Historical values of GSST and

Corresponding author: Timothy M. Hall, timothy.m.hall@nasa.gov

DOI: 10.1175/JCLI-D-20-0342.1

RSST are derived from the HadISST product (Rayner et al. 2003), while the future projected values are taken from CMIP5 climate model data (Taylor et al. 2012). (See methods section for further details.) GSST exhibits a secular increase since the 1970s driven largely by anthropogenic climate change (Santer et al. 2006; Gillett et al. 2008; Knutson et al. 2013). RSST can be considered as a proxy during this period for potential intensity (PI), which is a function of SST and atmospheric thermodynamic factors (Bister and Emanuel 2002). RSST has been shown to be correlated with NA TC activity (Vecchi and Soden 2007), as well as NA major hurricane activity (Murakami et al. 2018).

Large-scale SST indices have an advantage of being relatively insensitive to heterogeneities in the reanalysis data compared to historical estimation of PI (Kossin 2015). Moreover, SST data are available decades earlier than the reanalysis data required for PI construction, allowing the use of longer historical time periods in the construction of NASHM. In addition, GSST and RSST are relatively robust across climate model future projections compared to the ingredients needed to construct PI. However, these SST indices have the disadvantage that they are only indirectly related to the specific meteorological mechanisms responsible for evolving TC characteristics. Thus, we use these large-scale SST variables largely as proxies for specific meteorological mechanisms. For example, reduction in the tropical overturning circulation in a warming climate (Held and Soden 2006) and a weakening of the summer midlatitude circulation (Coumou et al. 2015) both impact TC steering winds and would be represented here only indirectly via increasing large-scale SST. As proxies, the SST variables have limitations. Emanuel and Sobel (2013) have argued that TC characteristics cannot be expressed as any unique function of SST, other than the approximate relationship between PI and RSST, and RSST is not projected to increase with climate warming (Vecchi et al. 2008; and see below). This leaves the increase in PI projected by climate models (Sobel et al. 2016) to be captured here by the increase in GSST, which is an incomplete representation (Emanuel and Sobel 2013). Therefore, our results may be conservative, as they do not capture other mechanisms that may increase PI as greenhouse gas concentrations increase.

We partly buffer the results here against the limitations of solely using SST-based climate variables by restricting the time horizon of the projections. NASHM is a statistical model that exploits historical relationships between aspects of TC activity and the SST indices. Once the relationships are established in the training and calibrating, we use NASHM to estimate TC activity at projected values of the SST indices. Such projections become increasingly less reliable as GSST or RSST move increasingly outside the historical range of variation, and, as mentioned above, projections in SST may not adequately capture longer-term trends in PI. Here, we make future projections only up to the year when the GSST anomaly, as determined from the ensemble mean of CMIP5 models, reaches 0.75°C, or 3.2 standard deviations above the 1900–2017 mean. For RCP4.5 this occurs in 2038; for RCP8.5 it occurs in 2030. Prior to these years NASHM's TC-activity projections via its GSST dependence can be tested. Out-of-sample tests on the highest historical GSST years verify NASHM's forecast skill compared to a naive SST-independent model (see below).

In section 2 we describe NASHM, evaluate its historical coastal TC activity and perform out-of-sample forecast tests, quantify model uncertainty, and present the historical and climate model–generated SST indices. In section 3 we analyze the TC sensitivity to the SST indices and present the forecast TC activity in the 2030s, driving NASHM with SST projections. In section 4 we discuss mechanism by isolating the SST dependence of NASHM's formation, track, and intensity components.

2. Methods

a. NASHM

The NASHM is a statistical–stochastic model that simulates the life cycle of TCs from formation through termination. NASHM was developed by Hall and Jewson (2007) and Hall and Yonekura (2013) and has been applied to study coastal TC activity and its relationship to climate (e.g., Hall and Hureid 2015; Hall and Tippett 2017). The model is trained on SST data and HURDAT2 archives (Landsea and Franklin 2013), in this application from 1900 to 2017. The SST indices are discussed below. In the construction of NASHM, attention has been paid to avoid overfitting. The spatial and temporal scales of HURDAT2 data used to estimate local rates (formation at a point, propagation from a point, intensity along a track) is determined objectively by maximizing out-of-sample likelihood. This procedure strikes the balance between resolving true structure, which pushes for smaller scales, and including sufficient data to avoid sampling error, which pushes for larger scales.

NASHM shares many characteristics with other statistical TC track models (e.g., Vickery et al. 2000; James and Mason 2005; Bloemendaal et al. 2020), many of which are used as part of commercial hurricane risk assessment and are calibrated to historical observations. Unlike many such models, NASHM is constructed to be sensitive to indices of climate variation, here GSST and RSST, via its three primary components—TC formation, track, and maximum-sustained wind speed V_{max}. Thus, it can employ information from global climate model projections to forecast future TC activity. In this regard, NASHM is similar to statistical–dynamical TC models (Emanuel et al. 2008; Lee et al. 2018), though such models require environmental fields as input that are more directly related to TC mechanisms than NASHM's SST indices.

Figure 1, top row, shows the 1900–2017 HURDAT2 data on which NASHM is trained and an example stochastic simulation of the period. The basinwide structure, westward storm movement in the subtropics, recurvature, and eastward return at midlatitudes, is well captured. This agreement between the basic basinwide structure in HURDAT2 and NASHM is further revealed in the track-density plots in the bottom row of Fig. 1. Once model training is performed on HURDAT2 and the SST indices, a calibration step is applied to long-term average landfall rates on several North American regions. The calibrated model is then used to estimate climate-driven variations on coastal TC activity. Figure 2 shows regional landfall rates as functions of V_{max}. A necessary criterion for lack of model landfall bias is that the stochastic spread in rates bound the historical rates, a criterion that NASHM meets, as illustrated in the figure.

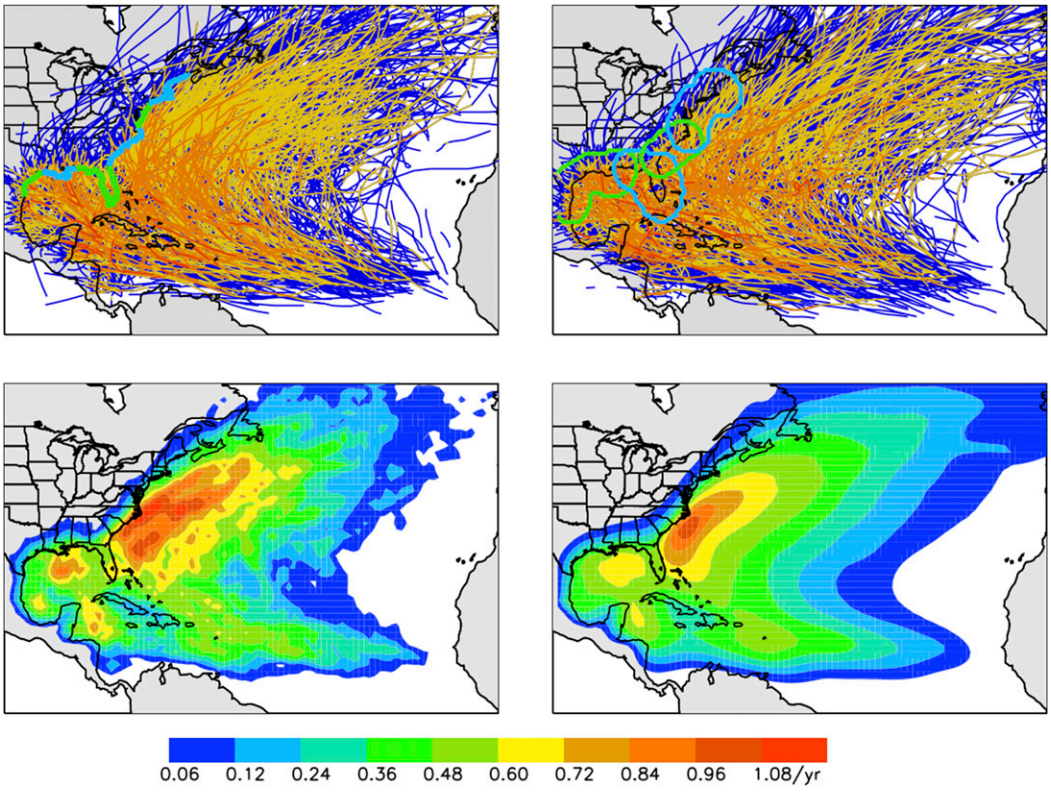

FIG. 1. Illustration of the historical TC data used to train NASHM, NASHM simulations, and evaluation of NASHM's U.S. TC landfall characteristics: (top left) HURDAT2 1900–2017 TC tracks upon which the stochastic model is trained. Tracks are color coded by maximum sustained wind speed V_{max}, showing TS and higher (blue), Saffir–Simpson Cat1 and higher (yellow), Cat3 and higher (orange), and Cat5 (red). Also shown are the landfall regions used in subsequent analysis. (top right) An example NASHM stochastic simulation of the 1900–2017 period. Also shown are the skill-assessment regions used in subsequent analysis. Also shown is annual track density (occurrence per year) in 1° boxes for TS and higher TCs: (bottom left) HURDAT2 1900–2017, corresponding to the tracks of the upper-left panel. (bottom right) The average over 1000 1900–2017 simulations, one of which is shown in the upper-left panel. White regions correspond to rates below the indicated dark-blue contour.

HURDAT2 is known to have underreporting biases in the earlier part of the record prior to the satellite era, and even more so prior to era of routine aircraft reconnaissance (Landsea 2007; Landsea et al. 2010; Vecchi and Knutson 2011). To buffer NASHM from these biases, at least in part, in the training data we only use TCs that last at least two days, reach at least tropical-storm status (at least 63 km h^{-1} 1-min-sustained V_{max}, referred to throughout as "Cat0+"), and the origin of tracks is taken as the point at which Cat0 intensity is achieved, rather than the first reported latitude–longitude fix. To examine further the sensitivity to these potential early record HURDAT2 biases, much of the analysis reported below is subsequently repeated for a version of NASHM trained on 1970–2017 HURDAT2, and the results lead to the same conclusions about evolution regional U.S. TC activity.

b. SST indices

The climate indices used here are derived from SST: 1) GSST, averaged over all longitudes and latitudes 23°S–23°N, and each year from July through October (JASO), and 2) RSST, the NA MDR (20°–60°W and 6°–18°N) JASO SST minus the JASO global 23°S–23°N region outside the Atlantic. As noted in section 1, SST indices have the advantage of being relatively insensitive to heterogeneities in the historic record and fairly robust across climate model projections. A disadvantage of simple SST indices is that their physical link to key TC characteristics is not direct. SST itself plays a direct role in formation and intensification via the PI (Bister and Emanuel 2002), but RSST likely underrepresents long-term trends in PI and thus may provide conservative estimate of future trends in TC activity. For example, ensemble-mean RSST does not increase in CMIP5 projections, as shown below, even though PI increases (Sobel et al. 2016). SST variation are also associated with steering wind variation and, therefore, TC tracks, but the relationship is indirect and mediated by intermediate mechanisms, for example, Wang et al. (2011) and other references from the discussion in section 4. Similarly, SST indices have indirect relationships with wind shear and spatial SST patterns, important influences on TC activity, for example the Atlantic meridional mode (Vimont and Kossin 2007). There is no

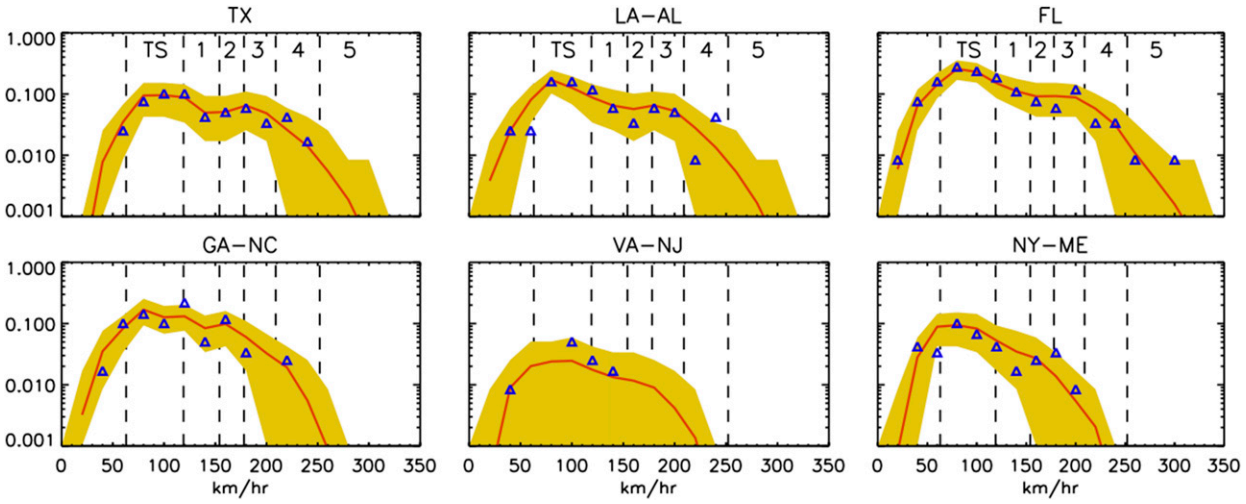

FIG. 2. Distribution of annual landfall rates as functions of V_{max} in 20 km h^{-1} bins on six U.S. coastal regions, as labeled by U.S. states and as mapped in the top-left panel of Fig. 1: Texas, Louisiana–Alabama, Florida, Georgia–North Carolina, Virginia–New Jersey, and New York–Maine. Red is the NASHM mean from 1000 simulations of 1900–2017, yellow is the inner 90% about the mean, and blue is computed directly with 1900–2017 HURDAT2. Vertical dashed lines indicate Saffir–Simpson category boundaries, as labeled in the Texas, Louisiana–Alabama, and Florida panels: TS and categories 1, 2, 3, 4, and 5. Note that for bins with zero historical events a blue symbol does not appear.

guarantee that the intermediate mechanisms providing a relationship between SST, steering winds, and wind shear in the historical period will continue to operate in a future climate. These limitations are the reason we limit our future projections to the 2030s. Beyond this time, GSST will likely be well above the highest value in the historic training period.

The projected SSTs are computed from the eight CMIP5 climate models (CCSM4, CNRM-CM5, NorESM1-M, ACCESS1.0, GISS-E2-R, HadGEM2-CC, HadGEM2-E, and MRI-GCM3S) that were ranked highest in a model–observation comparison of tropical SST anomalies (Brown et al. 2015). We use one ensemble member for each model and each emission scenario RCP4.5 and RCP8.5 and an additional RCP4.5 and RCP8.5 ensemble member for each of the models ranked in the top three (CCSM4, CNRM-CM5, NorESM1-M) by Brown et al. (2015). This results in a total of 11 ensemble members for each emission scenario. The CMIP5 SST projections start in 2006 (Taylor et al. 2012), thereby providing 12 years of overlap (2006–17) with the historical SST. We use this overlap period to adjust the projected GSST series such that each ensemble member's 2006–17-mean anomaly matches the historical 2006–17 mean anomaly, resulting in an ensemble-mean adjustment down of 0.07°C. For RSST no adjusting was necessary, as the ensemble-mean 2006–17-mean RSST already closely matched the historical mean value over this period.

Figure 3 shows the historical GSST and RSST and the CMIP5 ensemble-mean projections for the two climate change emission scenarios, RCP8.5 and RCP4.5. The ensemble-mean projected GSST continues the historical increase that began in the 1970s, with RCP8.5 rising faster than RCP4.5. For RSST, the RCP4.5 and RCP8.5 ensemble-mean projections remain essentially constant, equal to the historical 2006–17 value, which is elevated compared to the 1900–2017 RSST mean. Although individual ensemble members display considerable interannual variability in RSST, the ensemble means do not. The low-frequency variation in historical RSST, closely related to the Atlantic multidecadal oscillation (AMO), is not

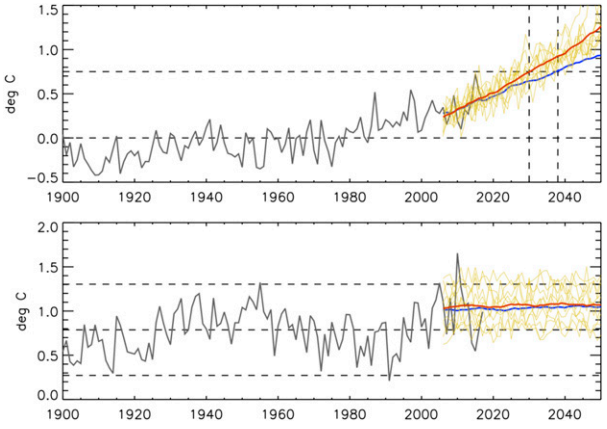

FIG. 3. The SST index covariates for (top) GSST (anomaly relative to the 1900–2017 mean) and (bottom) RSST. Black represents the historical annual time series. Blue represents the RCP4.5 CMIP5 ensemble annual mean, and red represents the RCP8.5 CMIP5 ensemble annual mean. The yellow lines are individual CMIP5 ensemble members for RCP8.5. In the top panel, the lower dashed line indicates the zero value and the upper dashed line indicates the highest value over the historical period. The vertical dashed lines indicate the years 2030 and 2038 at which the ensemble means exceed the highest historical value for RCP8.5 and RCP4.5, respectively. In the bottom panel, the center dashed line is the 1900–2017 mean and the lower and upper dashed lines indicate −2 and +2 standard deviations about the mean.

projected to continue in a coherent fashion by these models, possibly because the models replicate most of the historical AMO signal as an externally forced variation, rather than a natural mode of variability (Mann et al. 2020).

c. NASHM-SST forecast skill assessment

We use out-of-sample tests to evaluate NASHM's ability to forecast TC rates, given values of GSST and RSST. The changes projected for the 2030s depend primarily on GSST, and so we focus on this covariate. We select as a subset of historical data (the in-sample years) all of the years whose GSST is less than 1.5 standard deviations above the 1900–2017 GSST mean. There are 109 such in-sample years of the 118 total years. The remaining 9 years (the out-of-sample years: 1987, 1997, 1998, 2003, 2009, 2014, 2015, 2016, and 2017) have GSST greater than 1.5 standard deviation above the mean. We then train two versions of NASHM on the in-sample years: 1) a version that depends on GSST and RSST and 2) a naive version, whose components are trained on 1900–2017 HURDAT2 means and variances assuming an annually repeating mean state, as described in Hall and Jewson (2007), independent of interannual variation in GSST and RSST. For each model version, we perform 118 000 simulations for each of the out-of-sample years.

We evaluate the skill of these out-of-sample forecasts by comparing the rates with the historical events in the following manner: A 0.5° grid is defined covering the western North Atlantic and coastal North America, spanning 65°–100°W and 17°–46°N, 50-km-radius circles are drawn around each grid point, and TC tracks passing through the circles are counted. Because of the small area of the circles, passage rates are low and, to a good approximation, equivalent to the annual probability of TC passage through the circles. We take the passage rate averaged over the 11 800 simulations of each out-of-sample year as the mean rate of a Poisson distribution and evaluate that distribution at the historical count in the circle for the year. This results in a model-predicted probability for the historical event in that year and circle.

We then evaluate the skill in two ways:

1) The log probabilities are summed over all the out-of-sample years and over specified evaluation regions and intensity levels. We repeat the procedure using the naive model and compare the sums of the log probabilities. The climate-dependent model is deemed more skillful if its total log probability is greater than that of the naive model. This procedure is sensitive to the overall counts in a region, as well as the geographic distribution of counts within a region.
2) An F score is computed from calculations of true positive (TP), false positives (FP), and false negatives (FN). The climate-sensitive and naive model Poisson distributions are sampled at each location. A model forecast of a TC occurrence in the intensity category of interest is recorded TP if such an event occurred in the out-of-sample historical record, FP if such an event did not occur. A model forecast of no TC occurrence in the intensity category of interest is recorded FN if such an event did occur. From these, the precision $P = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP})$ and recall $R = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$, are computed, followed by the F score, $F = 2\mathrm{PR}/(P + R)$. We then document the percent difference in F between the climate-sensitive and naive models. Higher F indicates better performance. When counting TP, FP, and FN, the Poisson sampling is random. Values are accumulated over many repeated realizations of the procedure until the ratio of the climate-sensitive to naive F scores converges at the 1% level.

We choose five regions within 350 km of the coast to evaluate skill: northwest Gulf, Florida, the U.S. Southeast, the U.S. Northeast, and the full U.S. Gulf and Atlantic Coasts (Fig. 1b). We also evaluate skill on the complete 65°–100°W and 17°–46°N domain. The evaluations are performed for Saffir–Simpson category Cat0, Cat1–2, and Cat3–5 intensity levels. (The regions and intensity ranges are chosen to balance the desire to assess skill regionally but still have historical counts available for scoring.) Table 1 summarizes the results of the out-of-sample evaluation. In all but one case (15 of 16 region–intensity combinations), the climate-dependent model is more skillful than the naive model at forecasting TC occurrence in high GSST years according to both the log-likelihood comparison and the F-score comparison, though the U.S. Southeast Cat0 difference is marginal. The exception is Cat1–2 storms in the U.S. Northeast. Many such storms have transitioned from tropical to extratropical cyclones, which derive their energy from different meteorological mechanisms, and therefore may be less sensitive to the tropically based RSST and GSST indices used here. In general, the increase in skill is greatest for high-intensity TCs. Skill scores have the greatest potential for differentiating models where the differences between the out-of-sample and in-sample data are greatest, and it is at high intensity where the high-GSST years are most different fractionally from the in-sample years.

Although testing forecast skill for high GSST years is our primary focus, we have also performed similar out-of-sample tests on the five years (1908, 1909, 1910, 1911, and 1916) whose GSST anomaly is lower than −1.5 standard deviations. In this case, all of the region–intensity combinations show increased total log likelihoods using the climate-sensitive model. The F-score tests cannot be performed on all of the region–intensity combinations because there are not always historical events available to allow for TPs. In all cases where there is the possibility of a TP, the climate-sensitive model's F score is higher than the naive model. For cases where there is not the possibility of a TP, the climate-sensitive model has lower FN + FP than the naive model.

To evaluate further NASHM's skill, we compute the landfall rates on the same six regions of Fig. 2, but now only for the nine out-of-sample years of the skill testing. The results are shown in Fig. 4. The stochastic spread is here larger than in Fig. 2, because individual simulations are only nine years long, rather than 118 years. It is common for a 20 km h^{-1} V_{max} bin not to be populated over the nine years. Nonetheless, the historical values fall inside the inner 90% of the simulations, further evidence for NASHM's skill in making probabilistic forecasts.

TABLE 1. Out-of-sample predictive skill assessment on the nine years with GSST anomaly above 1.5σ by coastal region and intensity range according to two measures: accumulated log likelihood and F score based on precision and recall. In each entry, a plus indicates that the climate-dependent model has a higher total log likelihood than the naive model at predicting rates of TS occurrence within 350 km of the indicated region and intensity range, and minus indicates that the climate-dependent model is not more skillful than the naive model. Numbers for F scores are the percentage increase or decrease in the climate-sensitive F score relative to the naive F score.

	Northwest Gulf	Florida	Southeast	Northeast	U.S. total	North Atlantic
Cat0	+; 5%	+; 8%	+; 0%	+; 8%	+; 6%	+; 8%
Cat1–2	+; 15%	+; 15%	+; 10%	−; −6%	+; 4%	+; 12%
Cat3–5	+; 23%	+; 51%	+; 79%	+; 92%	+; 71%	+; 56%

d. *Uncertainty estimates*

In addition to testing skill, we want to estimate the uncertainty of our best-estimate occurrence rates for the historical baseline and the projections shown below. To do this, we perform a bootstrap analysis (Efron 1979). The 118 years (1900–2017) of HURDAT2 are resampled with replacement 100 times. NASHM is then trained on each of the 100 bootstrap resamples, and each NASHM-bootstrap model is used to simulate 1000 realizations of the 118-yr period (118 000 years). Mean TC occurrence rate maps (annual passage rates through the set of 50-km circles) are computed from each bootstrap simulation, resulting in bootstrap distributions of the mean occurrence rates at each location. We then compute quantiles from the bootstrap distributions for use in assessing uncertainty of the best-estimate rates. For uncertainties on the mean rates, we compute the 5%–95% quantile range of the bootstrap distribution.

To quantify the significance of differences between rates in one period and another period, a different quantile-selection scheme is used. We want to estimate how significant is the difference between two best-estimate values, given their individual bootstrap distributions. What is the probability that a value from the lower distribution is greater than a value from the higher distribution? The probability that a value in the lower distribution is in the upper 0.224 fraction (the 77.6% quantile) and the value in the higher distribution is in the lower 0.224 fraction (the 22.4% quantile) is $(0.224)^2 = 0.050$, assuming the distributions are independent. Therefore, if the 22.4%–77.6% ranges of the distributions do not overlap, then the best estimates are significantly different to at least 95% confidence. This estimate of the significance of the differences is conservative because the distributions are not, in fact, independent. In the simulations, if a 1950–80 value of a bootstrap-resample member is on the low (high) end of its bootstrap distribution, then the 2030 value also tends to be on the low (high) end of its distribution, because the long-term averaged rates from that bootstrap member are lower (higher).

3. Results

We now discuss the impact of GSST and RSST variation on TC activity and make projections into the 2030s. To analyze the impact of various levels of GSST and RSST on TC activity, both individually and in combination, long NASHM simulations are performed (118 000 years) with GSST and RSST held fixed in various combinations. We choose 30 combinations of fixed-anomaly values with respect to 1900–2017, using the six values −0.5°C, −0.25°C, 0°C, +0.25°, +0.5°C, and +0.75°C for

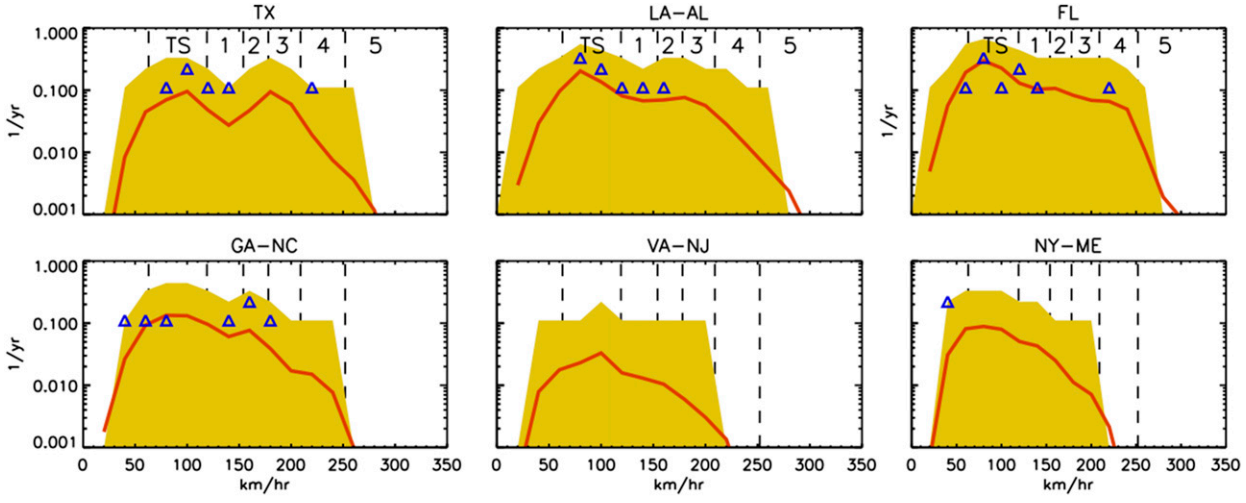

FIG. 4. As in Fig. 2, but here using the simulations of the nine out-of-sample years (1987, 1997, 1998, 2003, 2009, 2014, 2015, 2016, 2017) of the forecast skill test. Note that, for many V_{max} bins (and all of Virginia–New Jersey), there are zero events in the nine out-of-sample years; an event count falls inside the inner 90% model spread.

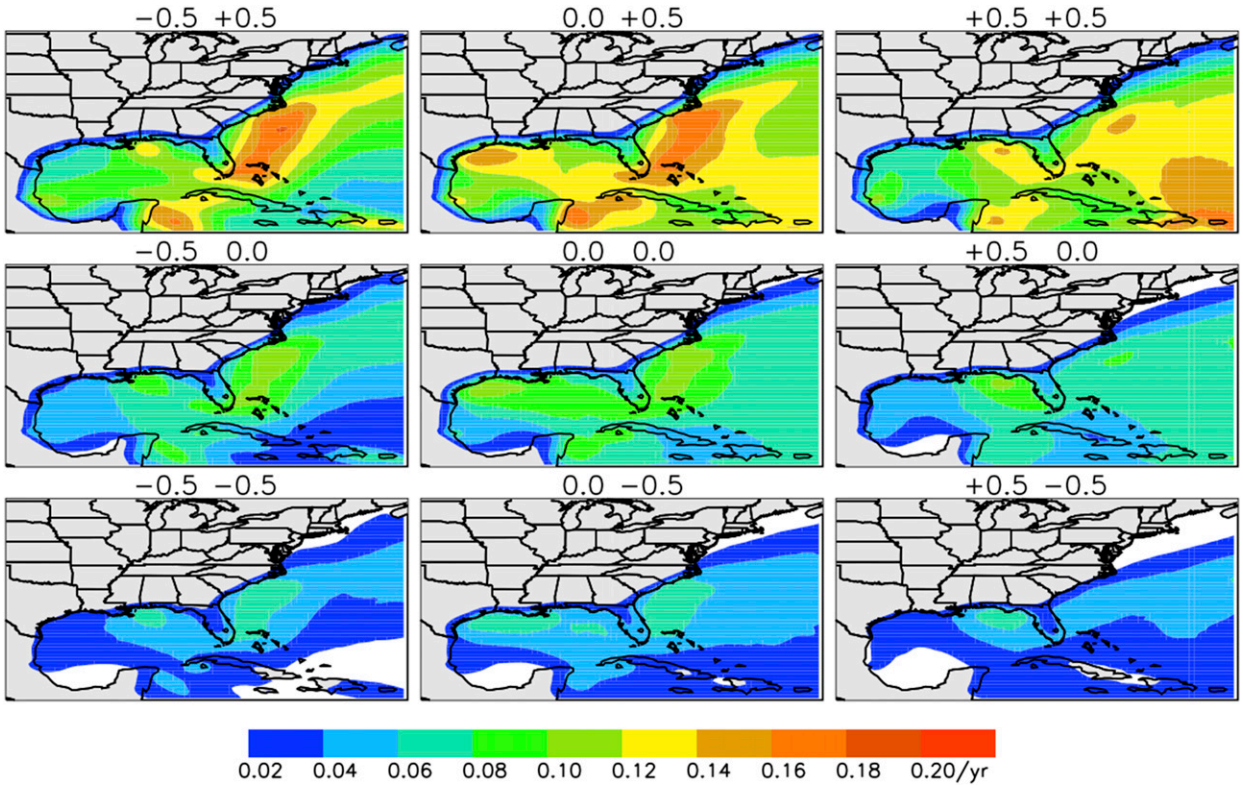

FIG. 5. Cat1+ annual occurrence rate maps as a function GSST and RSST anomaly, as labeled. Anomaly labels represent degrees Celsius difference from the 1900–2017 means. GSST is the first label, and RSST is -the second. White regions correspond to rates below the indicated dark-blue contour.

GSST and the five values −0.5°C, −0.25°C, 0°C, +0.25°, and +0.5°C for RSST. This represents a range from −2.1 to +3.2 standard deviations for GSST and from −1.9 to +1.9 standard deviations for RSST. We then compute from the simulated TCs for each GSST–RSST combination the average annual rates of TC passing through the set of 50-km circles on the gridded field defined previously. Figure 5 shows the annual Cat1+ (Saffir–Simpson category 1 and higher) TC occurrence rates for 9 of the 30 GSST–RSST combinations. Increasing RSST (bottom to top) is primarily associated with an overall basinwide increase in rate, with relatively little spatial variation in the increase. By contrast, increasing GSST (left to right) results in changes in geographic distribution, but comparatively little basinwide change in TC annual rates. Of particular interest is a shift eastward in Gulf and Atlantic occurrence rate as GSST is increased from a neutral to an elevated level. For example, in the Gulf the hot spot near Texas and Louisiana for GSST = 0°C anomaly (for any RSST) shifts eastward near the Florida Panhandle at GSST = +0.5°C anomaly. Similarly, the hot spot near North Carolina shifts eastward off the U.S. East Coast.

The rate maps of Fig. 5 are bilinearly interpolated in the two-dimensional GSST–RSST space to the GSST and RSST values of each year of the CMIP5 projections, 2020–38 for RCP4.5 and 2020–30 for RCP8.5. (We could have directly simulated the response to the projected CMIP5 RSST and GSST values. We chose the approach of simulating the GSST–RSST combinations then interpolating in order to document the separate dependence on each SST covariate. As the CMIP5 ensemble-mean RSST and GSST are smoothly varying, interpolating vs direct simulation yields very similar results year by year through the projected time horizon.) We then compare the rate maps at the end of the projections to NASHM's rate map from a set of 1950–80 historical baseline simulations. Note that the rates depend on time only via the SST indices, and thus the RCP4.5 rate map in 2038 is identical to the RCP8.5 rate map in 2030. The SST values at these final dates are GSST = 0.75°C and RSST = 0.20°C, expressed as anomalies from the 1900–2017 means. It may appear at first that a single projected year (2030 or 2038) is being compared to an average over a range of years (1950–80), which would be problematic due to internal variability. However, the single projected year is in fact derived from 118 000 years of stochastic simulation variability, and the driving GSST and RSST values are derived from an ensemble mean of climate model simulations, each with its individual variability.

Figure 6 shows the TC occurrence rate maps for the 1950–80 baseline, the 2038 RCP4.5 (equivalent to 2030 RCP8.5), and the difference, for intensities of Cat1+, Cat3+, and Cat4+. Basic climatological features of the maps are peaks in rates in the Gulf and U.S. mid-Atlantic, as well as a rapid decrease with intensity of rates at mid latitudes, seen by comparing Cat3+

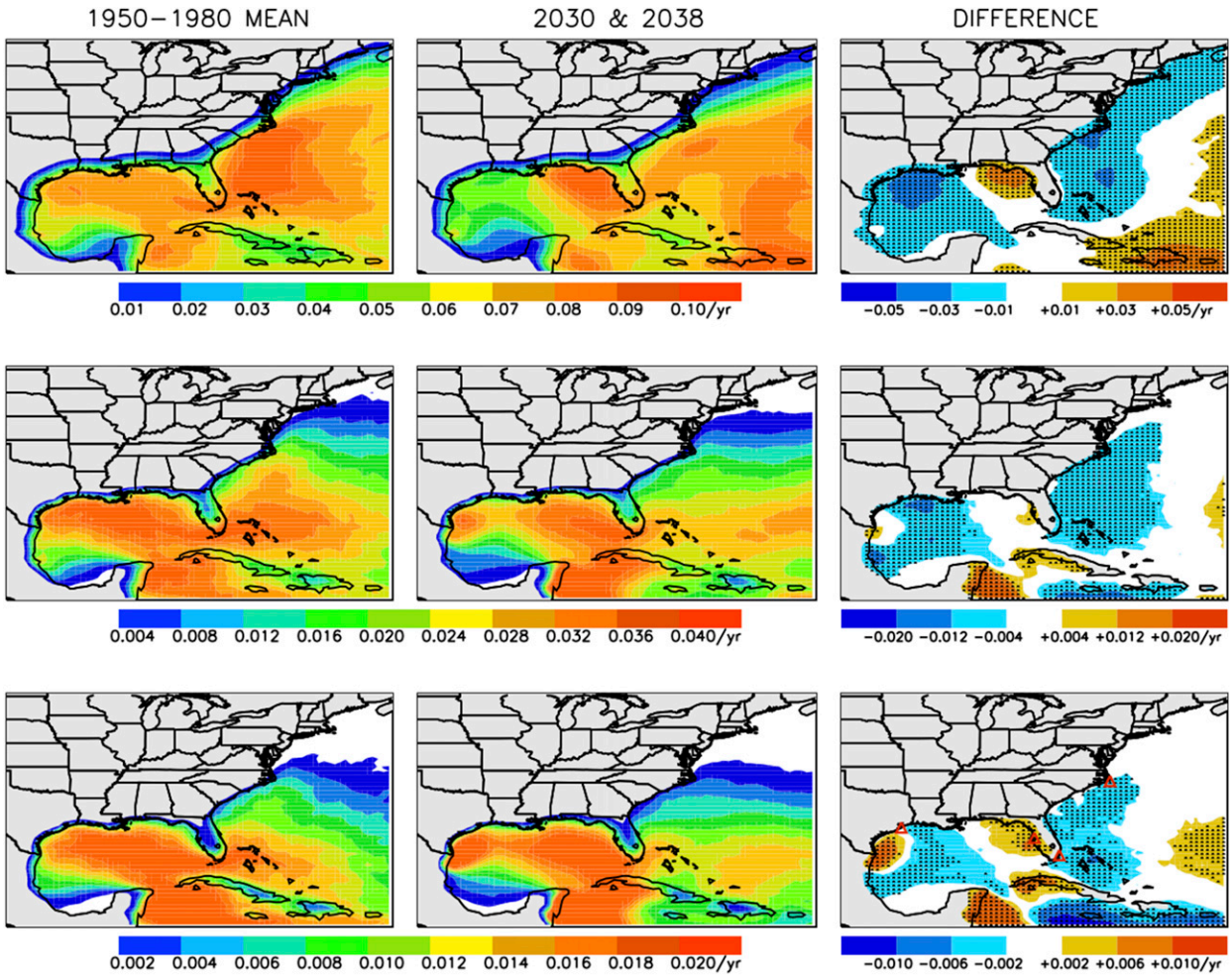

FIG. 6. Maps of annual TC annual occurrence rates (yr^{-1}) for (left) 1950–80, (center) 2030 (RCP8.5) and 2038 (RCP4.5), and (right) the difference for (top) TCs of intensity Cat1+, (middle) Cat3+, and (bottom) Cat4+. The contour scales change from row to row, as indicated by the color bars. In the left and center columns, white regions have rates below the dark-blue scale indicated on the color bars. In the right column, stippling indicates points where the differences are estimated to be significant at 95% confidence by a bootstrap analysis.

(middle row) and Cat4+ (bottom row) with Cat1+ (top row). The difference maps reveal a complex geographic pattern of change that is approximately coherent across the TC intensities. In the eastern Gulf, extending from the eastern Yucatan, through western Cuba, to the Florida Panhandle, change in TC occurrence rates is either positive or neutral. Over that region, the fractional rate increase is greatest for intensity Cat4+ (bottom row). By contrast, in the western Gulf and the Atlantic off the U.S. Eastern Seaboard, change in TC occurrence rates is mostly negative or neutral, except for a positive region at higher intensities near the Texas–Mexico border. In both of these regions, the fractional decrease is smallest for intensity Cat4+ (bottom row). Farther east, into the central Atlantic, the change is again negative or neutral. This change pattern can in part be summarized as a shift eastward in rates for both the Gulf and the Atlantic, resulting in a west-to-east negative-to-positive pattern in both the Gulf and the western Atlantic. Superposed on this spatial pattern is a broad increase at the highest intensity levels.

For U.S. Gulf Coast states, the eastward shift results in a projected decrease in TC occurrence rates to the west, particularly on eastern coastal Texas and western Louisiana, and a projected increase to the east, particularly Florida's Panhandle and Gulf Coasts. In between, centered roughly on Mississippi and Alabama, is a node, with little or no change.

Figure 7 shows the TC occurrence rates averaged over four coastal regions: Texas through Alabama, Florida's Gulf Coast, Florida's Atlantic coast, and Georgia through Cape Hatteras, NC. Rates are shown for three time periods: 1950–80, 2010–19, and 2030 for RCP8.5 (identical to 2038 for RCP4.5). Consistent with the change maps of Fig. 6, rates increase on Florida's Gulf Coast and decrease Georgia to Cape Hatteras for all intensity thresholds, and the changes from 1950–80 to 2030 are significant according to the bootstrap analysis described in section 2d,

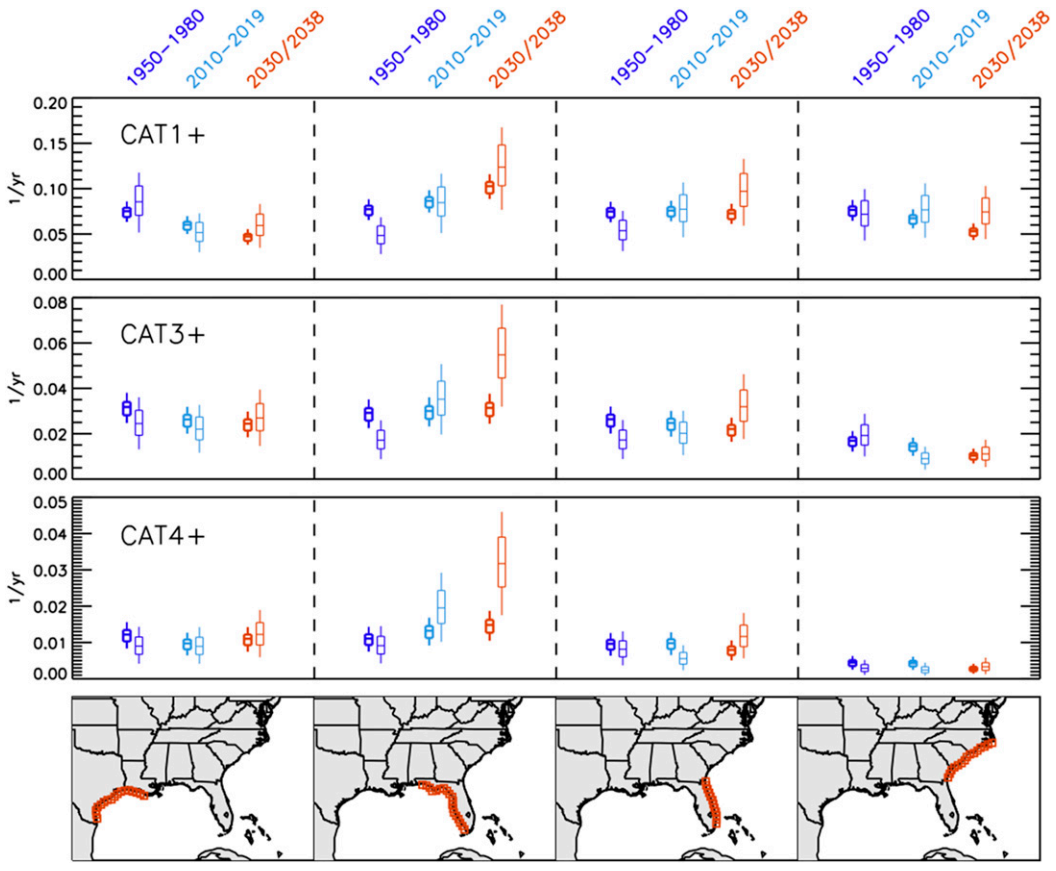

FIG. 7. Annual occurrence rates for (top) Cat1+, (top middle) Cat3+, and (bottom middle) Cat4+, averaged over 50-km-radius circles spanning (bottom) four coastal locations. Dark blue represents the period 1950–80, light blue is the period 2010–19, and red is the RCP8.5 projected rates for 2030 (identical to the RCP4.5 projected rates for 2038). The boxes and whiskers indicate the best-estimate rates (midlines) and the 22%–78% (top/bottom of the boxes) and 5%–95% confidence ranges (whiskers) about the best estimate determined from a bootstrap resampling analysis. Heavy lines indicate results that are based on 1900–2017 HURDAT2, and light lines indicate results that are based on 1970–2017 HURDAT2.

except for Cat3+ on the Florida Gulf. Rates decrease across Texas–Alabama for Cat1+ and Cat3+, but the change for Cat4+ is not significant.

We note that RSST displays considerable year-to-year and decadal variability over the historic period (Fig. 3). This in turn drives year-to-year and decadal variability in TC rates. The RSST projections based on ensemble-mean CMIP5 results, however, show no variability. The resulting projected TC rates therefore reflect only the secular increase of the ensemble-mean GSST. Natural climate variability will certainly continue to drive variability in TC rates. Using ensemble-mean projected SST, however, we have no basis to project such variability, and we focus instead on the change driven by the secular variation in GSST.

4. Discussion

To explore the causes for the pattern of increased and decreased TC rates, we look at the GSST and RSST dependence of the three major components of NASHM: formation, track, and intensity. In NASHM, TC formation occurs by sampling Poisson densities whose mean rates are dependent on GSST and RSST (Hall and Yonekura 2013). Figure 8 (bottom row) shows spatial maps of change in mean rates as a function of GTSST and RSST compared with the neutral state of zero GSST and RSST anomalies. GSST and RSST induce different formation responses. Elevating GSST by 0.75°C (the 2030 RCP8.5 and 2038 RCP4.5 value, 3.2 standard deviations above the 1900–2017 mean) with RSST held neutral, results in a modest increase in overall annual formation count (from 10.4 to 12.6). However, there is a distinct eastward shift in formation, illustrated by the large increase in the eastern NA MDR. By contrast, overall storm count is more sensitive to RSST. Elevating RSST by the much smaller 0.20°C (the 2030 RCP8.5 and 2038 RCP4.5 value, roughly 1 standard deviation above the 1900–2017 mean), with GSST held neutral, results in about the same increased storm count (from 10.4 to 12.5). However, in contrast to GSST, the formation increase induced by RSST is spatially more uniform throughout the tropical North Atlantic. The combined GSST and RSST increases, reflecting the 2030

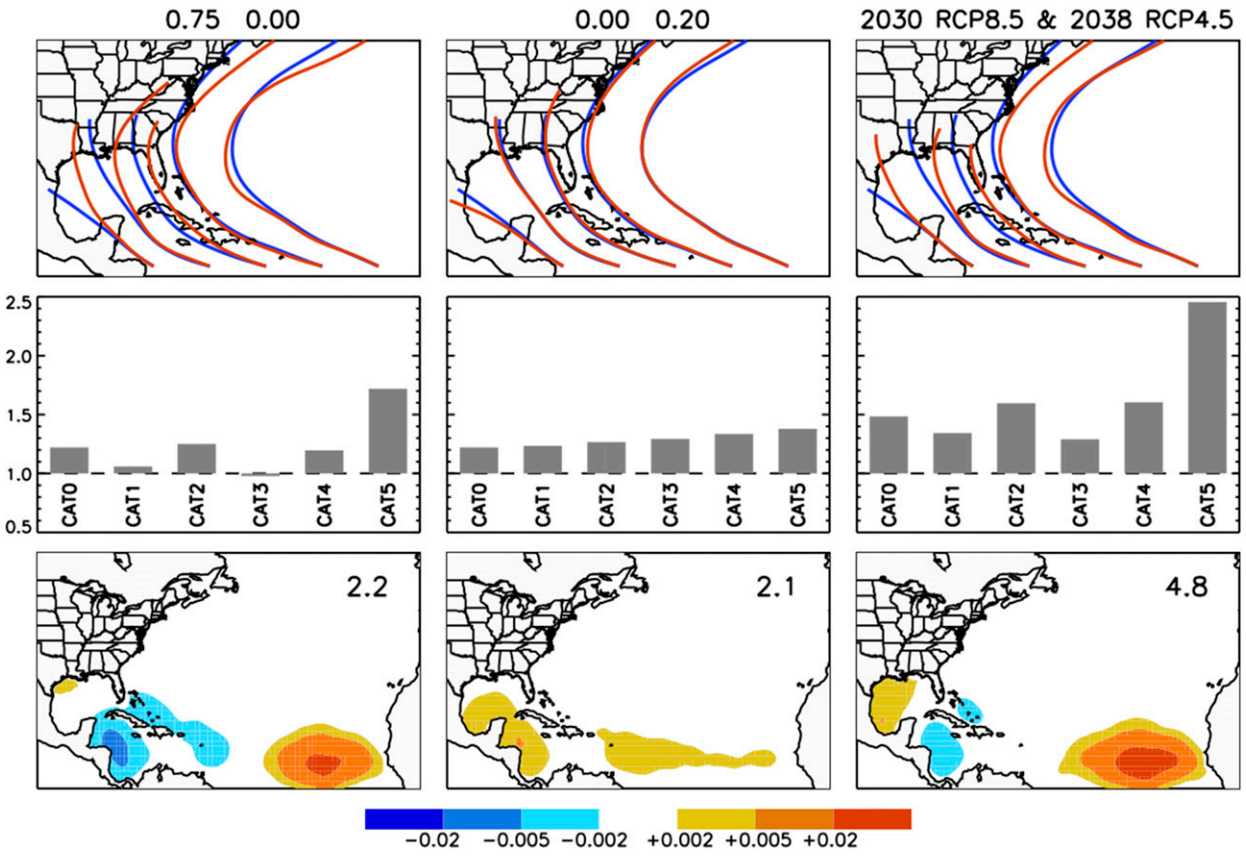

FIG. 8. Illustration of influence of covariates GSST and RSST on (top) TC tracks, (middle) LMI, and (bottom) formation rate. GSST–RSST anomaly values with respect to 1900–2017 are (left) 0.75° and 0.00°C, (center) 0.00° and 0.20°C, and (right) 0.75° and 0.20°C. The 0.75° and 0.20°C combination corresponds to the 2030 RCP8.5 and 2038 RCP4.5 values, as labeled at top. For tracks, the stochastic component of NASHM is turned off for clarity, leaving only mean tracks, which are launched from five specified locations. In each panel, blue is the neutral state and red is the GSST–RSST combination. For LMI, intensity is binned in Saffir–Simpson categories, and the ratios with respect to the neutral state are plotted. For TC formation, the difference between the GSST–RSST combination and the neutral state is mapped, with red and blue being positive and negative, respectively. Contour units are per year per 1° latitude–longitude grid box. Differences in basinwide annual TC formation are inset the top right of each bottom panel.

RCP8.5 and 2038 RCP4.5 values, result in a basinwide count increase from 10.4 to 15.2, combined with the eastward shift induced by RSST.

TC mean track in NASHM also depend on the SST covariates (Hall and Yonekura 2013). To better visualize the impact of the SST covariates on the TC tracks, we turn off the stochastic component of NASHM and release mean tracks from a fixed set of locations. Figure 8 (top row) shows the results for the same set of GSST and RSST as the formation results. There is an eastward anomaly in the mean track for elevated GSST, but little change for RSST. For example, the mean track that landfalls in Louisiana in the neutral case instead landfalls in the Florida Panhandle in the 0.75° and 0.0°C and 0.75° and 0.20°C cases. In addition, in midlatitudes, mean tracks near the U.S. East Coast recurve away from the U.S. Northeast more in the elevated GSST cases.

TC intensity in NASHM is modeled by a weighted resampling scheme with perturbation of V_{max} time series (Hall and Yonekura 2013). Weights are applied such that it is more likely to resample a V_{max} time series from a historical year whose SST covariate values are similar to the simulation year. The weights are selected objectively to minimize out-of-sample V_{max} forecast error. Figure 8 (middle row) shows histograms of simulation lifetime maximum intensity (LMI) in the elevated GSST and RSST states compared to the neutral state. Elevated RSST induces an increase at all intensities, with the fractional change increasing gradually with intensity. By contrast, elevated GSST induces a sizeable increase only at Cat5, with smaller or no increases at lower intensities. The combined GSST–RSST increase, representing 2030 (RCP8.5) and 2038 (RCP4.5), results in an approximate factor-2.5 increase in Cat5 occurrence and a 30%–60% increase at lower intensity thresholds.

An important point here is that large-scale changes in TC characteristics (formation, propagation, intensity) in response to these global and basin SST covariates can result in regional coastal changes in TC activity that are highly heterogeneous, even in sign. Figure 9 is an idealized illustration of this point that helps to understand the increased occurrence rate in the

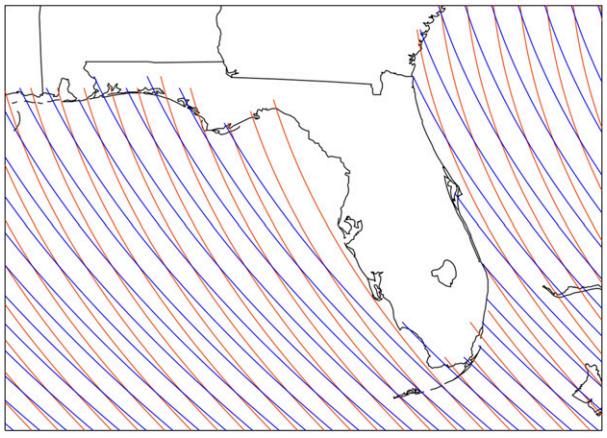

FIG. 9. An idealized set of tracks illustrating the regional impacts of basinwide propagation changes. Two sets of mean tracks start from the same origin points. Red differs from blue only in having a constant uniform propagation anomaly of 5 km eastward for every 25 km northward. Florida's Gulf Coast is more exposed to the red tracks, whereas its Atlantic coast is less exposed.

northeast Gulf. Consider two sets of idealized mean tracks both originating from the same fixed set of formation points at 15°N every 0.5° in longitude. One set differs from the other only in a constant uniform eastward propagation anomaly of 5 km eastward for every 25 km northward. Despite the uniformity of the anomaly, the coastal impact varies regionally. Figure 9 shows the example of Florida. Without the eastward anomaly, Florida's eastern Panhandle, Big Bend region, and much of its Gulf Coast are shielded from direct landfalls. With the eastward anomaly, parts of these regions now get direct landfalls. Thus, there is an increase in activity on western Florida. By contrast, the same track anomaly results in decreased activity on Florida's east coast, especially the northeast, because the mean tracks now curve roughly parallel to the coast without making landfalls. Clearly, with realistic tracks (stochastic in NASHM), any region can be struck directly. However, the idealization illustrates that mean track shifts, even if uniform, can increase probability in some regions and decrease it in other regions. The NASHM mean track changes of Fig. 6 have similar character to these idealized changes and result in the increased occurrence rates in the northeast Gulf.

NASHM is trained on HURDAT2 1900–2017. However, there are known biases in the earlier presatellite record, approximately prior to 1970 (Landsea 2007; Landsea et al. 2010; Vecchi and Knutson 2011), especially for weak storms short that do not landfall. We have attempted to buffer against these biases in the model training by only using HURDAT2 tracks that reach at least tropical storm (TS) status and starting the tracks at the point where TS status is reached. Nonetheless, to test sensitivity to the data record period, we have reconstructed NASHM using 1970–2017 HURDAT2. Vecchi and Knutson (2011) estimate that after 1965 undercounts in annual TC frequency due to sparce sampling are negligible, hence 1970–2017 is inside the well-sampled satellite era. Figure 10 repeats Fig. 8's illustration of formation, mean track, and LMI sensitivity to GSST and RSST, but now for the 1970–2017-trained model. The GSST and RSST anomalies for 2030/2038 (RCP8.5/RCP4.5) are here reduced to 0.57° and 0.19°C from 0.72° and 0.20°C because they are now computed with respect to a 1970–2017 baseline.

The qualitative behavior in TC sensitivity is similar for all three model components: a shift eastward in formation toward the eastern MDR, an eastward propagation anomaly, and an increase in the LMI distribution at the highest intensities. The resulting changes in TC passage rates, shown in Fig. 11, are also similar to those of the 1900–2017 training shown in Fig. 6. However, there are quantitative differences: the basinwide increase with respect to the zero-anomaly state for GSDST and RSST in annual TC count is reduced from 4.8 to 3.0, and the eastward track anomaly is enhanced. The change in baseline period explains much of the change in basinwide formation; the 1970–2017 period has 1.5 more TC yr^{-1} than 1900–2017 (11.9 vs 10.4). Undercounts in the early HURDAT2 record (Landsea 2007) may be responsible for some of this difference. The increased eastward track anomaly could be an indication of an acceleration in the signal or simply statistical variability, which is exaggerated in the shorter data training record. These quantitative differences have an impact on occurrence rates in certain coastal regions. Figure 6 shows rates based on 1970–2017 analysis, as well as 1900–2017. The uncertainty ranges for the shorter period are significantly larger, and the western Gulf and U.S. Southeast do not change significantly. However, the increase in rates on Florida's Gulf Coast is significantly amplified, likely due to the increased eastward track anomaly. In addition, the 1970–2017-based analysis projects and increase on Florida's Atlantic coast for Cat1+ and Cat3+. Southern Florida is near a node between decreased and increased rates (Figs. 6 and 11), and in the 1970–2017 period.

Similar SST-dependent formation, track, and intensity features have been reported in other studies, which have examined a range of time periods and driving mechanisms. The NASHM intensity variations are consistent with a number of modeling and observational studies finding that positive trends in TC activity are fractionally greatest at the highest intensities (Sobel et al. 2016; Gutmann et al. 2018; Knutson et al. 2020; Elsner et al. 2008; Kossin et al. 2013, 2020). Meanwhile, an eastward shift in formation has been seen by Kossin et al. (2010) during the warm phase of the Atlantic meridional mode (AMM), due not only to increased subtropical SST, but also to covariation with vertical wind shear. Eastward formation shifts have also been reported by Wang et al. (2011) in the context of expanded Atlantic warm pools, by Murakami and Wang (2010) in a 20-km-mesh atmospheric model driven by SST projections, and by Colbert et al. (2013) in a CMIP3 downscaling study. Chauvin et al. (2019) has also seen an eastward formation shift using high-resolution global modeling.

A subtropical eastward track anomaly similar to that shown here (Figs. 8 and 10, top row) has also been reported in modeling and observational studies. Explanations of mechanisms vary, but all the studies have in common an association between increased SST and changes in tropospheric wind patterns that steer TCs. Wang et al. (2011) argue that an expanded Atlantic warm pool, which could be associated with changes in

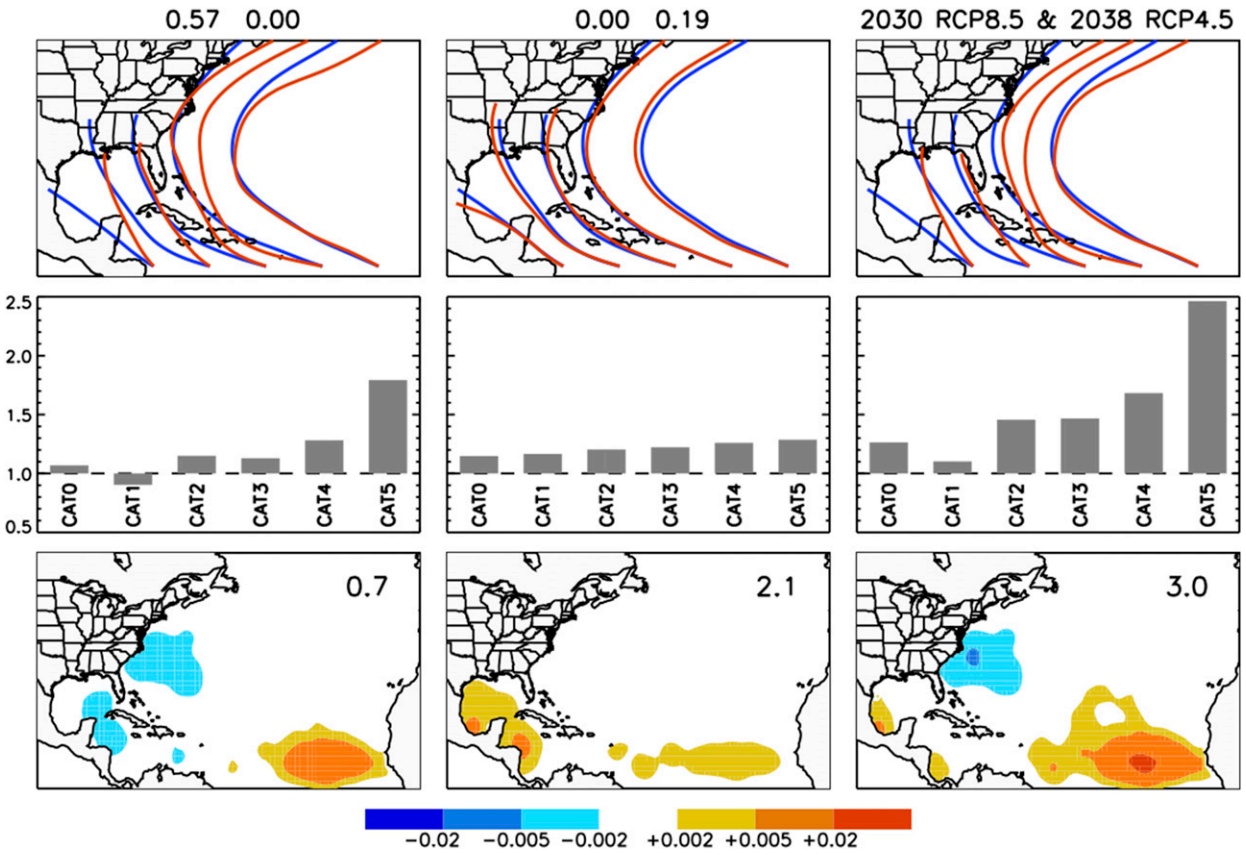

FIG. 10. As in Fig. 8, but here for the model trained only on HURDAT2 from 1970 to 2017.

either GSST or RSST, results in a weakened NA subtropical high pressure system, thereby weakening westward tropical midtropospheric winds that steer TCs. In statistically downscaled CMIP3 climate change simulations, Colbert et al. (2013) observe an eastward TC track anomaly, due largely to a reduction in westward steering winds. A study of Hurricane Sandy under enhanced SST conditions found a tendency for the storm to recurve eastward away from the coast (Lau et al. 2016), similar to the eastward anomaly seen here in midlatitude. Liu et al. (2017), using 4-km-resolution regional modeling, detected a northeastward anomaly in the NA midtropospheric TC steering winds. On the basis of reconstructions from proxy records, Baldini et al. (2016) presented evidence for a northeastward shift over the industrial era in TC tracks that originate in NA MDR. Recently, Hassanzadeh et al. (2020) have found in analysis of CMIP5 climate projections a positive meridional anomaly in steering winds in the eastern Gulf, which would direct tracks northward before reaching the western Gulf, consistent with the eastward anomaly we find in the Gulf. All of these changes project onto U.S. landfall rates.

The projected basinwide increase in annual NA TC frequency shown here runs counter to many, though not all, previous results. Climate model results of Murakami et al. (2020) indicate that increased NA TC frequency of recent decades will wane in future, as greenhouse gas forcing dominates aerosol forcing, and TC frequency will decline in the NA starting around 2030 (their Fig S3). By contrast, Bhatia et al. (2018) show an increase in global TC frequency, including in the NA. Increased frequency has also been seen in CMIP5 models (Villarini and Vecchi 2013). The work of Vecchi et al. (2019) indicates that projected TC frequency change has complex dependence on numerical model resolution and the spatial pattern of warming, with decreases in some experiments and increases in others. Analyzing the distribution of results across a large number of studies, Knutson et al. (2020) show that the median North Atlantic TC frequency change is negative, but that the interquartile range across the studies spans positive change. We deduce an increase of 4.6 yr^{-1} in 2038 (RCP4.5) relative to a 1900–2017 baseline of 10.4 yr^{-1}, which is an increase of 2.6 yr^{-1} relative to the 1986–2005 baseline of 12.4 yr^{-1} used by Knutson et al. (2020). The 21% increase (2.6/12.4) falls within the 5%–95% quantile spread show by Knutson et al. (2020). Similarly, the mean time trend of 0.6 TCs per decade 1995–2038 (RCP4.5) falls within the spread across CMIP5 models shown by Villarini and Vecchi (2013) for the first half of the twenty-first century. For the model trained on 1970–2017 HURDAT2, we deduce an increase by 2038 (RCP4.5) of 2.0 yr^{-1} relative to 1986–2005, a 17% increase

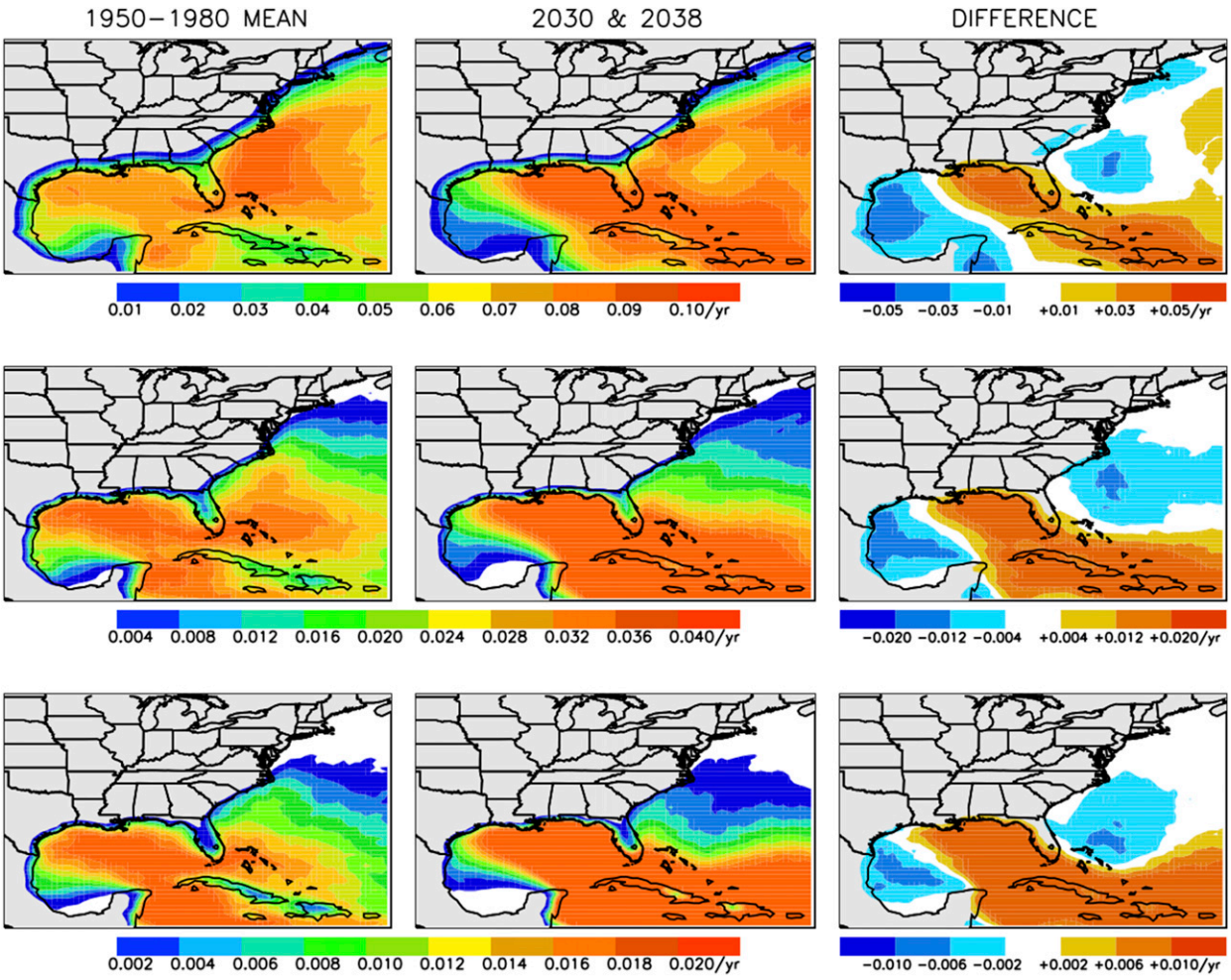

FIG. 11. As in Fig. 6, but here the projected maps are based on NASHM trained on HURDAT2 1970–2017, instead of 1900–2017.

(2.0/12.4) that also falls within the 5%–95% quantile spread shown by Knutson et al. (2020).

5. Conclusions

In summary, we have used a statistical model trained on historical TCs and large-scale SST indices to project U.S. hurricane activity into the 2030s, given ensemble-mean climate model projections of the SST indices under warming scenarios. We find a complex signal of changing hazard along the U.S. coast driven by changes in basinwide TC frequency and changes in track characteristics. Some areas are forecast to experience increased hazard, such as the Gulf Coast of Florida, and other areas are forecast to experience decreased hazard, such as parts of the western Gulf and the U.S. mid-Atlantic.

In our analysis, changes in U.S. TC activity are driven by projected secular increase in tropical SST and, indirectly, changes in other factors impacting TC activity, such as steering winds, wind shear, and spatial SST patterns, via their covariance with the large-scale SST indices. We stress that, in addition to these changes, there is large, well-documented interannual TC variation driven by internal modes of climate variability, in particular the AMO (e.g., Goldenberg et al. 2001) and El Niño–Southern Oscillation (ENSO) (e.g., Gray 1984). Ensemble-mean climate model SST projections provide no prediction for these important modes of variability, and we do not include their effects here. From year to year, variations in TC activity driven by strong ENSO events can be larger than the secular variations discussed here, and on a multiyear time scale, a changing AMO phase may have an even larger impact. It has not been our goal here to estimate the impact on TCs of natural variability, such as the AMO and ENSO, important as they are. There are likely additional variations at small scales in wind shear, steering winds, and SST patterns affecting TC activity that are not correlated with GSST and RSST, and the impact of these variations is also not directly represented here, except through their impact on uncertainty estimates. Instead, our aim has been to estimate the impact on TCs into the 2030s of projected large-scale secular climate trends that exist in addition to the natural variability. Estimation of trends in TC hazard in the coming decades, as well has the impact of natural variability, has important applications to estimating the changing risk to populations and infrastructure in harm's way.

Acknowledgments. This research was supported by the National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), and The Climate Service. Author Hall's effort was funded by NASA Modeling, Analysis, and Prediction (MAP) Program Award 19-MAP19-0005. We thank three anonymous reviewers for comments and ideas that led to an improved paper.

Data availability statement. The HURDAT2 data on which NASHM is trained are freely available for download at the National Hurricane Center website, as are the historical SST data at the Hadley Centre website. CMIP5 model-projected SST data are also freely available on several online platforms.

REFERENCES

Baldini, L. M., and Coauthors, 2016: Persistent northward North Atlantic tropical cyclone migration over the past fine centuries. *Nat. Sci. Rep.*, **6**, 37522, https://doi.org/10.1038%2FSREP37522.

Bhatia, K., G. Vecchi, H. Murakami, S. Underwood, and J. Kossin, 2018: Projected response of tropical cyclone intensity and intensification in a global climate model. *J. Climate*, **31**, 8281–8303, https://doi.org/10.1175/JCLI-D-17-0898.1.

Bister, M., and K. A. Emanuel, 2002: Low frequency variability of tropical cyclone potential intensity, 1. Interannual to interdecadal variability. *J. Geophys. Res.*, **107**, 4801, https://doi.org/10.1029/2001JD000776.

Bloemendaal, N., I. D. Haigh, H. de Moel, S. Muis, R. J. Haarsma, and J. C. J. H. Aerts, 2020: Generation of a global synthetic tropical cyclone hazard dataset using STORM. *Sci. Data*, **7**, 40, https://doi.org/10.1038/s41597-020-0381-2.

Brown, J. N., C. Langlais, and A. S. Gupta, 2015: Projected sea surface temperature changes in the equatorial Pacific relative to the warm pool edge. *Deep-Sea Res. II*, **113**, 47–58, https://doi.org/10.1016/j.dsr2.2014.10.022.

Chauvin, F., P. Romain, P. Palany, and A. Belmadani, 2019: Future changes in Atlantic hurricanes with the rotated-stretched ARPEGE-Climat at very high resolution. *Climate Dyn.*, **54**, 47–58, https://doi.org/10.1007/S00382-019-05040-4.

Colbert, A. J., B. J. Soden, G. A. Vecchi, and B. P. Kirtman, 2013: The impact of anthropogenic climate change on North Atlantic tropical cyclone tracks. *J. Climate*, **26**, 4088–4095, https://doi.org/10.1175/JCLI-D-12-00342.1.

Coumou, D., and Coauthors, 2015: The weakening summer circulation in the Northern Hemisphere mid-latitudes. *Science*, **348**, 324–327, https://doi.org/10.1126/science.1261768.

Efron, B., 1979: Bootstrap methods: Another look at the jackknife. *Ann. Stat.*, **7** (1), 1–26, https://doi.org/10.1214/aos/1176344552.

Elsner, J. B., J. P. Kossin, and T. H. Jagger, 2008: The increasing intensity of the strongest tropical cyclones. *Nature*, **455**, 92–95, https://doi.org/10.1038/nature07234.

Emanuel, K., and A. H. Sobel, 2013: Response of tropical sea surface temperature, precipitation, and tropical cyclone-related variables to changes in global and local forcing. *J. Adv. Model. Earth Syst.*, **5**, 447–458, https://doi.org/10.1002/jame.20032.

——, R. Sundarajan, and J. Williams, 2008: Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bull. Amer. Meteor. Soc.*, **89**, 347–368, https://doi.org/10.1175/BAMS-89-3-347.

Gillett, N. P., P. A. Stott, and B. D. Santer, 2008: Attribution of cyclogenesis region sea surface temperature change to anthropogenic influence. *Geophys. Res. Lett.*, **35**, L09707, https://doi.org/10.1029/2008GL033670.

Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nuñez, and W. M. Gray, 2001: The recent increases in Atlantic hurricane activity. *Science*, **293**, 474–479, https://doi.org/10.1126/science.1060040.

Gray, W. M., 1984: Atlantic seasonal hurricane frequency. Part I: El Niño and the 30-mb quasi-biennial oscillation influences. *Mon. Wea. Rev.*, **112**, 1649–1668, https://doi.org/10.1175/1520-0493(1984)112<1649:ASHFPI>2.0.CO;2.

Gutmann, D., and Coauthors, 2018: Changes in hurricanes from a 13-yr convection-permitting pseudo-global warming simulation. *J. Climate*, **31**, 3643–3657, https://doi.org/10.1175/JCLI-D-17-0391.1.

Hall, T. M., and S. Jewson, 2007: Statistical modeling of North Atlantic tropical cyclone tracks. *Tellus*, **59A**, 486–498, https://doi.org/10.1111/j.1600-0870.2007.00240.x.

——, and E. Yonekura, 2013: North American tropical cyclone landfall and SST: A statistical model study. *J. Climate*, **26**, 8422–8439, https://doi.org/10.1175/JCLI-D-12-00756.1.

——, and K. Hureid, 2015: The frequency and duration of U.S. hurricane droughts. *Geophys. Res. Lett.*, **42**, 3482–3485, https://doi.org/10.1002/2015GL063652.

——, and M. Tippett, 2017: Pacific hurricane landfalls on Mexico and SST. *J. Appl. Meteor. Climatol.*, **56**, 667–676, https://doi.org/10.1175/JAMC-D-16-0194.1.

Hassanzadeh, P., C.-Y. Lee, E. Nabizadeh, S. J. Camargo, D. Ma, and L. Y. Yeung, 2020: Effects of climate change on the movement of future landfalling Texas tropical cyclones. *Nat. Commun.*, **11**, 3319, https://doi.org/10.1038/s41467-020-17130-7.

Held, I. M., and B. J. Soden, 2006: Robust changes of the hydrological cycle to global warming. *J. Climate*, **19**, 5686–5699, https://doi.org/10.1175/JCLI3990.1.

James, M. K., and L. B. Mason, 2005: Synthetic tropical cyclone database. *J. Waterw. Port Coastal Ocean Eng.*, **131**, 181–192, https://doi.org/10.1061/(ASCE)0733-950X(2005)131:4(181).

Knutson, T. R., F. Zeng, and A. T. Wittenberg, 2013: Multimodel assessment of regional surface temperature trends: CMIP3 and CMIP5 twentieth-century simulations. *J. Climate*, **26**, 8709–8743, https://doi.org/10.1175/JCLI-D-12-00567.1.

——, and Coauthors, 2019: Tropical cyclones and climate change assessment: Part I: Detection and attribution. *Bull. Amer. Meteor. Soc.*, **100**, 1987–2007, https://doi.org/10.1175/BAMS-D-18-0189.1.

——, and Coauthors, 2020: Tropical cyclones and climate change assessment: Part II. Projections. *Bull. Amer. Meteor. Soc.*, **101**, E303–E322, https://doi.org/10.1175/BAMS-D-18-0194.1.

Kossin, J. P., 2015: Validating atmospheric reanalysis data using tropical cyclones as thermometers. *Bull. Amer. Meteor. Soc.*, **96**, 1089–1096, https://doi.org/10.1175/BAMS-D-14-00180.1.

——, S. J. Camargo, and M. Sitkowski, 2010: Climate modulation of North Atlantic hurricane tracks. *J. Climate*, **23**, 3057–3076, https://doi.org/10.1175/2010JCLI3497.1.

——, T. L. Olander, and K. R. Knapp, 2013: Trend analysis with a new global record of tropical cyclone intensity. *J. Climate*, **26**, 9960–9976, https://doi.org/10.1175/JCLI-D-13-00262.1.

——, T. M. Hall, T. Knutson, K. E. Kunkel, R. J. Trapp, D. E. Waliser, and M. F. Wehner, 2017: Extreme storms. *Climate Science Special Report: Fourth Annual National Climate Assessment*, Vol. I, D. J. Weubbles et al., Eds., U.S. Global Climate Change Research Program, 256–276.

——, K. R. Knapp, T. L. Olander, and C. S. Velden, 2020: Global increase in major tropical cyclone exceedance probability over

the past four decades. *Proc. Natl. Acad. Sci. USA*, **117**, 11 975–11 980, https://doi.org/10.1073/pnas.1920849117.

Landsea, C. W., 2007: Counting Atlantic tropical cyclones back to 1900. *Eos, Trans. Amer. Geophys. Union*, **88**, 197–202, https://doi.org/10.1029/2007EO180001.

——, and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, https://doi.org/10.1175/MWR-D-12-00254.1.

——, G. A. Vecchi, L. Bengtsson, and T. R. Knutson, 2010: Impact of duration thresholds on Atlantic tropical cyclone counts. *J. Climate*, **23**, 2508–2519, https://doi.org/10.1175/2009JCLI3034.1.

Lau, W. K. M., J. J. Shi, W. K. Tao, and K. M. Kim, 2016: What would happen to Superstorm Sandy under the influence of a substantially warmer Atlantic. *Geophys. Res. Lett.*, **43**, 802–811, https://doi.org/10.1002/2015GL067050.

Lee, C.-Y., M. K. Tippett, A. H. Sobel, and S. J. Camargo, 2018: An environmentally forced tropical cyclone hazard model. *J. Adv. Earth Syst.*, **10**, 223–241, https://doi.org/10.1002/2017MS001186.

Liu, C., and Coauthors, 2017: Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dyn.*, **49**, 71–95, https://doi.org/10.1007/s00382-016-3327-9.

Mann, M. E., B. A. Steinman, and S. K. Miller, 2020: Absence of internal multidecadal and interdecadal oscillations in climate model simulations. *Nat. Commun.*, **11**, 49, https://doi.org/10.1038/s41467-019-13823-w.

MoftakharI, H. R., G. Salvadori, A. AghaKouchak, B. F. Sanders, and R. A. Matthew, 2017: Compounding effects of sea level rise and fluvial flooding. *Proc. Natl. Acad. Sci. USA*, **114**, 9785–9790, https://doi.org/10.1073/pnas.1620325114.

Murakami, H., and B. Wang, 2010: Future change of North Atlantic tropical cyclone tracks: Projection by a 20-km-mesh global atmospheric model. *J. Climate*, **23**, 2699–2721, https://doi.org/10.1175/2010JCLI3338.1.

——, E. Levin, T. L. Delworth, R. Gudgel, and P.-C. Hsu, 2018: Dominant effect of relative tropical Atlantic warming on major hurricane occurrence. *Science*, **362**, 794–799, https://doi.org/10.1126/science.aat6711.

——, and Coauthors, 2020: Detected climatic change in global distribution of tropical cyclones. *Proc. Natl. Acad. Sci. USA*, **117**, 10 706–10 714, https://doi.org/10.1073/pnas.1922500117.

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 analysis of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, https://doi.org/10.1029/2002JD002670.

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

Smith, A., and R. W. Katz, 2013: US billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. *Nat. Hazards*, **67**, 387–410, https://doi.org/10.1007/s11069-013-0566-5.

Sobel, A. H., S. J. Camargo, T. M. Hall, C.-Y. Lee, M. K. Tippett, and A. A. Wing, 2016: Human influence on tropical cyclone intensity. *Science*, **353**, 242–246, https://doi.org/10.1126/science.aaf6574.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experimental design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1.

Tebaldi, C., B. H. Strauss, and C. E. Zervas, 2012: Modelling sea level rise impacts on storm surges along US coasts. *Environ. Res. Lett.*, **7**, 014032, https://doi.org/10.1088/1748-9326/7/1/014032.

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

——, and T. R. Knutson, 2011: Estimating annual numbers of Atlantic hurricanes missing from the HURDAT database (1878–1965) using ship track density. *J. Climate*, **24**, 1736–1746, https://doi.org/10.1175/2010JCLI3810.1.

——, K. L. Swansom, and B. J. Soden, 2008: Whither hurricane activity. *Science*, **322**, 687–689, https://doi.org/10.1126/science.1164396.

——, and Coauthors, 2019: Tropical cyclone sensitivity to CO_2 doubling: Roles of atmospheric resolution, synoptic variability, and background climate changes. *Climate Dyn.*, **53**, 5999–6033, https://doi.org/10.1007/s00382-019-04913-y.

Vickery, P. J., P. F. Skerlj, and L. A. Twisdale, 2000: Simulation of hurricane risk in the U.S. using empirical track model. *J. Struct. Eng.*, **126**, 1222–1237, https://doi.org/10.1061/(ASCE)0733-9445(2000)126:10(1222).

Villarini, G., and G. A. Vecchi, 2013: Projected increases in North Atlantic tropical cyclone intensity from CMIP5 models. *J. Climate*, **26**, 3231–3240, https://doi.org/10.1175/JCLI-D-12-00441.1.

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

Wang, C., H. Liu, S.-K. Lee, and R. Atlas, 2011: Impact of the Atlantic warm pool on United States landfalling hurricanes. *Geophys. Res. Lett.*, **38**, L19702, https://doi.org/10.1029/2011GL049265.

Yoshida, K., M. Sugi, R. Mizuta, H. Murakami, and M. Ishii, 2017: Future changes in tropical cyclone activity in high-resolution large-ensemble simulations. *Geophys. Res. Lett.*, **44**, 9910–9917, https://doi.org/10.1002/2017GL075058.