AGU PUBLICATIONS

Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE

10.1002/2017MS001186

An Environmentally Forced Tropical Cyclone Hazard Model

Chia-Ying Lee[1], Michael K. Tippett[2,3], Adam H. Sobel[2,4], and Suzana J. Camargo[4]

[1]International Research Institute for Climate and Society, Columbia University, Palisades, NY, USA, [2]Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA, [3]Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia, [4]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA

Key Points:
- The description and assessment of a tropical cyclone (TC) hazard model
- A tool for understanding changes in the global climate to the regional TC hazard
- TC climatology, including the statistics of rapid intensification, and regional TC hazard are well simulated

Correspondence to:
C.-Y. Lee,
clee@iri.columbia.edu

Citation:
Lee, C.-Y., Tippett, M. K., Sobel, A. H., & Camargo, S. J. (2018). An environmentally forced tropical cyclone hazard model. *Journal of Advances in Modeling Earth Systems*, *10*, 223–241. https://doi.org/10.1002/2017MS001186

Received 3 OCT 2017
Accepted 2 JAN 2018
Accepted article online 5 JAN 2018
Published online 20 JAN 2018

Abstract A new statistical-dynamical model is developed for estimating the long-term hazard of rare, high impact tropical cyclones events globally. There are three components representing the complete storm lifetime: an environmental index-based genesis model, a beta-advection track model, and an autoregressive intensity model. All three components depend upon the local environmental conditions, including potential intensity, relative sea surface temperature, 850 and 250 hPa steering flow, deep-layer mean vertical shear, 850 hPa vorticity, and midlevel relative humidity. The hazard model, using 400 realizations of a 32 year period (approximately 3,000 storms per realization), captures many aspects of tropical cyclone statistics, such as genesis and track density distribution. Of particular note, it simulates the observed number of rapidly intensifying storms, a challenging issue in tropical cyclone modeling and prediction. Using the return period curve of landfall intensity as a measure of local tropical cyclone hazard, the model reasonably simulates the hazard in the western north Pacific (coastal regions of the Philippines, China, Taiwan, and Japan) and the Caribbean islands. In other regions, the observed return period curve can be captured after a local landfall frequency adjustment that forces the total number of landfalls to be the same as that observed while allowing the model to freely simulate the distribution of intensities at landfall.

1. Introduction

From 1963 to 2012, tropical cyclones (TCs) were responsible for more than 50% of all meteorologically induced financial losses (Geiger et al., 2016). TC hazard assessment is important to government, industry, financial institutions, NGOs, and even individual households in the context of individual events, seasonal predictions, and climate adaptation. Accurate risk assessment depends on the hazard—the probability of a TC of a given magnitude in a given location—in addition to vulnerability factors, such as the growth of wealth and population (Estrada et al., 2015). We focus on hazard in this study. Because of the limited historical record, a common approach for estimating TC hazard is to compute statistics from simulated as well as observed storms (e.g., Emanuel et al., 2008). In this approach, the complete lifetime of each simulated storm, including its genesis, track, intensity, and landfall, are simulated. Alternatively, one can statistically model the landfall rate alone (Tolwinski-Ward, 2015). Most industry catastrophe models (models which represent TC hazard as well as vulnerability and financial losses to insured assets) use statistical methods to generate synthetic storms that are similar to those in historical data (e.g., AIR WORLDWIDE, 2015). Some of them include the dependence of storm activity on a few environmental parameters, such as basin sea surface temperature (SST) or measures of the El Niño-Southern Oscillation (ENSO) (e.g., Hall & Jewson, 2007; Yonekura & Hall, 2011, 2014). These models, while they generally perform well in the current climate, are strongly constrained to the historical record and are not designed to consider the effects of climate change. To understand the impact of climate change on TC hazard, global climate models or dynamical downscaling methods are the most straightforward approaches. Dynamical models calculate individual TC evolution based on the laws of physics, and can provide information globally (whereas many statistical models are developed for individual basins). However, at the high spatial resolutions necessary for TC simulation, it is computationally expensive to generate a sufficient number of synthetic storms for hazard assessment, where one is particularly interested in very rare and extreme events. Thus, Emanuel et al. (2006) proposed a novel statistical-dynamical downscaling method. In this method, each TC's evolution is calculated using a combination of statistical and simplified dynamical models that are forced by environmental conditions taken from global models. The model of Emanuel et al. (2008) randomly seeds storms globally, moves them

using a beta-advection model (Marks, 1992), and calculates intensity evolution using a simple coupled ocean-atmosphere tropical cyclone model (CHIPS, Emanuel et al., 2004). Emanuel's model has been broadly used for understanding the impact of a changing climate on TC climatology (Emanuel, 2013, 2015), storm surge hazard (Lin et al., 2012), and TC-induced economic losses (Geiger et al., 2016).

In previous work, we focused on developing a model for TC intensity, which is a challenging issue even for short-term forecasting. Lee et al. (2015, 2016a) describe a global autoregressive (AR) TC intensity model. The AR model contains a deterministic component, derived empirically, which advances the TC intensity in time and accounts for the surrounding large-scale environment. The stochastic forcing of the AR model represents the component of TC intensification that is not linearly related to the storm's ambient conditions. Simulating the intensity evolution along the observed tracks, the AR model captures the observed TC intensity climatology well, except for the bimodal distribution in the lifetime maximum intensity (LMI). This bimodality, which is important for simulating the frequency of the most intense storms, might be associated with a combination of two factors. The first one is physical, namely the distinction between those storms that undergo rapid intensification (RI) and those that do not (Lee et al., 2016b), while the other one is arguably an artifact associated with a weakness of the Dvorak technique (Kossin et al., 2013; Landsea & Franklin, 2013). In this study, we will show that the AR intensity model is capable of simulating the observed LMI distribution when the simulated storm lifetime is determined consistently with the intensity model, rather than by the lifetime of the prescribed tracks. The consistency between AR and the simulated lifetime improves the model's ability to simulate RI, providing further evidence on the physical cause of the bimodality in the LMI distribution. Other intensification models include that of Lin et al. (2017), who used a multiple linear regression model, found that the dependence of TC intensification to environment is nonhomogeneous, and suggested a mixture modeling approach as a solution. Recently, Emanuel (2017) reduced the complexity of his intensity model to a set of two prognostic equations for storm intensity and innercore moisture and further increased the efficiency of his hazard model.

In the present study, we develop and assess a complete statistical-dynamical downscaling TC hazard model. We develop genesis and track components and couple them to the existing AR intensity model described in detail in Lee et al. (2015, 2016a). Both the genesis and track components depend on the local environment. Thus, the whole system is environmentally forced with no explicit spatially dependent component. (Spatial dependence is captured implicitly, through the dependence on environmental variables which themselves have spatial structure.) The model is developed for the current climate with all the environmental parameters downscaled from European Centre for Medium-Range Weather Forecasts interim reanalysis (ERA-Interim). The data and methods used for the model development and evaluation are described in section 2. We introduce the individual model components (genesis, track, and intensity), respectively, in section 3. The TC hazard model performance is first evaluated by its ability to capture the observed TC climatology, including frequency, intensity, landfall, and interannual variability (section 4). Next, we compare the observed and simulated hazard in various places across the globe (section 5). Throughout this study, we define "hazard" as the probability (or equivalently the return period) of the storm intensity at landfall exceeding a given threshold at a particular location. The summary and discussion are given in section 6.

2. Data and Methods

2.1. Observational and Reanalysis Data Sets

The best-track data set, HURDAT2, produced by the National Hurricane Center (NHC) is used for the North Atlantic (ATL) and Eastern North Pacific (ENP) (Landsea & Franklin, 2013). For TCs in the Western North Pacific (WNP), Indian Ocean (IO), and Southern Hemisphere Ocean (SH), we use the best-track data from Joint Typhoon Warning Center (JTWC, Chu et al., 2002). Both data sets include 1 min maximum sustained wind, minimum sea level pressure, and storm location every 6 h. Large-scale environmental variables are calculated from the European Centre for Medium-Range Weather Forecasts interim reanalysis (ERA-Interim, Dee et al., 2011. We use monthly data for all three TC hazard model components—the genesis, track, and intensity models. In the track model, daily 250 and 850 hPa steering flow winds are used as well. In this study, data from 1981 to 2012 are used for evaluation. Data from 1981 to 1999 are used as training data for the intensity model. (The training of the intensity model was done in previous study (Lee et al., 2016a), and therefore was using data from different period.)

Throughout this study, the Saffir-Simpson scale is used to categorize storm strength in all basins. The ranges used are 64–82 kt for category 1 (Cat 1) storms, and 83–95, 96–112, 113–136, >137 kt for categories 2–5 (Cat 2–5) storms, respectively. The threshold for tropical storm (TS) is 34 kt. Storm lifetime maximum intensity (LMI) is defined as the maximum sustained wind speed during the storm's life cycle.

2.2. Identifying Landfall Locations

For risk assessment, it is important to calculate the landfall probability at a given location, and thus to identify landfall. We first linearly interpolate track data (for both observations and simulations) to a 15 min resolution. Surface type (land or ocean) is assigned to each interpolated point using 0.5° resolution topography data from NASA (https://neo.sci.gsfc.nasa.gov/view.php?datasetId=SRTM_RAMP2_TOPO). Then, landfall is defined when a storm center moves from a ocean point to a land point. To avoid counting landfalls multiple times in the situation when a storm moves over archipelago regions, such as the Philippines, landfalls need to be at least 100 km and 6 h apart to be considered as independent landfalls.

2.3. Experimental Design

Simulations from the TC hazard model will be called GTI here, in which "G," "T," and "I" stand for Genesis, Track, and Intensity models, respectively. In order to isolate the influence of the individual components on the estimated TC statistics and hazard, we design two additional experiments: $\hat{G}\hat{T}I$ uses only the intensity model while the genesis and tracks, represented as $\hat{}$, are taken directly from the best-track data set; $\hat{G}TI$ uses both track and intensity models, but observed genesis locations and times. As we will discuss in the next section, each of three components in the TC hazard model contains a stochastic parameter. Thus, the hazard model is a stochastic system. We construct 400 realizations of a 32 year period (1981–2012) in every experiment. In $\hat{G}\hat{T}I$, the 400 realizations differ in only in the component due to the intensity model. In $\hat{G}TI$, there are 10 sets of tracks (with the same observed genesis locations) and each set has 40 intensity realizations. Realizations with the same underlying tracks but different intensities can still differ in their lifetimes (due to the different realizations of the intensity model solution), and thus in how much of each track is actually covered by a storm. A similar design is used for GTI but the genesis locations in each set are calculated from the genesis model separately.

2.4. Evaluation Measures

To evaluate a stochastic model performance, we use two statistical measures:

The *t statistic* of a variable is defined as the simulated ensemble mean ($\bar{m}$) minus the observed mean ($\bar{o}$) divided by a weighted standard deviation ($\sigma_{weighted}$):

$$t\ \ statistic = \frac{\bar{m} - \bar{o}}{\sigma_{weighted}}, \tag{1}$$

$$\sigma_{weighted} = \sqrt{\frac{(n_o - 1)\sigma_o^2 + (n_m - 1)\sigma_m^2}{n_o + n_m - 2}} \tag{2}$$

where n_o and n_m are the number of samples for observations and simulations (i.e., number of years) while σ_o and σ_m are the respective standard deviations. The *t* statistic shows the difference between simulations and observations compared to the *natural variability*, defined in equation (2). A caveat is that the natural variability defined here tilts toward the simulated variability, because n_m is significantly larger than n_o. However, the real natural variability is likely to be under-sampled in the observations. Since there is no obviously good choice, we choose to use $\sigma_{weighted}$ and treat observation as one of the realizations. The distribution of *t* statistic also tells us whether the differences between observations and simulations are systematic (i.e., has a pattern) or nonsystematic (the positive and negative values are randomly distributed).

The *Rank histogram* of a variable is defined as the distribution of the rank (in percentage) of the observations with respect to the simulations. If the ensemble members and the observations are drawn from the same probability distribution, the rank of observations with respect to the simulations will be uniformly distributed. When the simulation is biased, under-dispersed or over-dispersed, the shape of the rank histogram will be tilted, bimodal with peaks at two ends, or monomodal.

3. Development of Individual Model Components

The key hypothesis of our model is that storm properties can be represented using model components that are functions of small numbers of key local environmental variables. First, the genesis model determines the rate at which weak vortices are formed throughout the domain, which are then passed to the intensity and track models to determine the rest of the storms' life cycles.

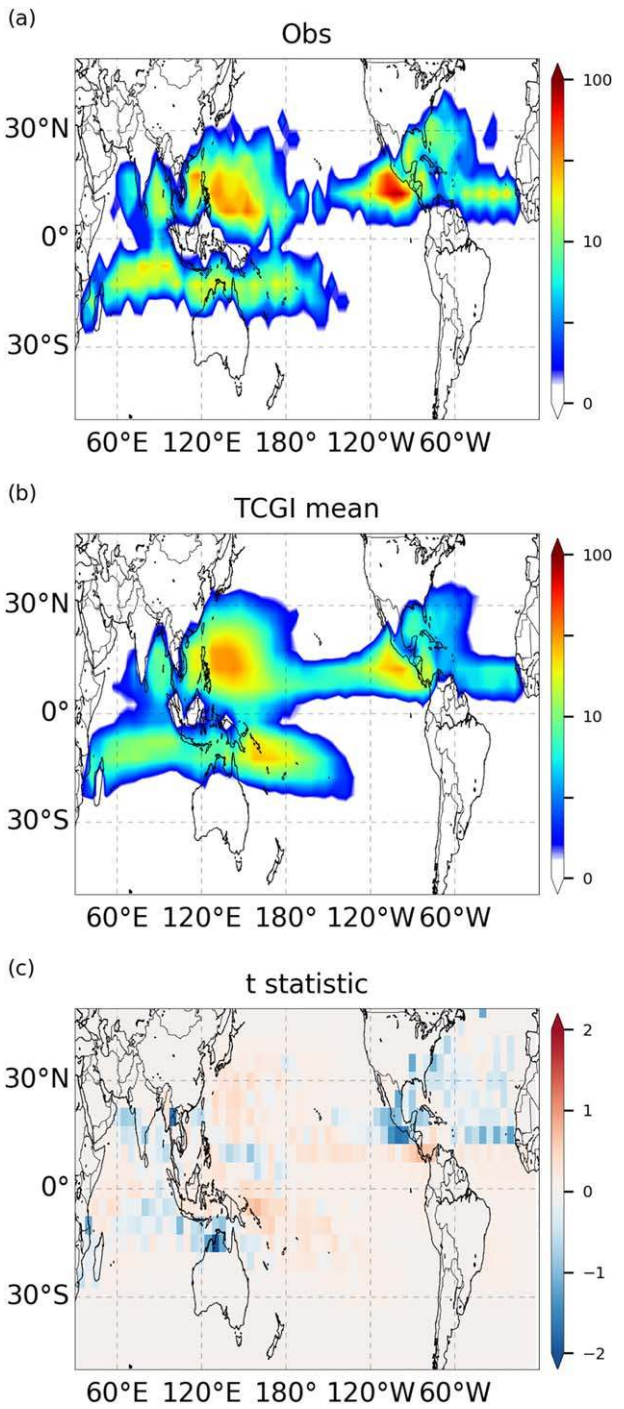

Figure 1. Number of TC genesis per 5° × 5° box from 1981 to 2012 (a) from observations, and (b) averaged from 40 TCGI simulations. (c) *t* statistic of TCGI simulations. Note that the scales in Figures 1a and 1b are logarithmic, while the scale is linear in Figure 1c.

3.1. Genesis—Tropical Cyclone Genesis Index (*TCGI*)

An essential element in the genesis model is the seeding rate. Previous studies have shown that with only a few crucial environmental parameters, various TC genesis (potential) indices can capture the location, frequency, and the seasonality of TC formation, including ENSO-induced variability (Bruyère et al., 2012; Camargo et al., 2007a, 2007b; Emanuel, 2010; Emanuel & Nolan, 2004; McGauley & Nolan, 2011; Tippett et al., 2011). Menkes et al. (2012) compared the existing indices, and found that all have similar performance in genesis climatology. The Tropical Cyclone Genesis Index (TCGI, developed by Tippett et al., 2011), however, has the least bias and the best simulated seasonality. Thus, we calculate the seeding rate based on TCGI:

$$TCGI = \exp\left(b + b_\eta \eta_{850} + b_{RH} RH_{600} + b_{SST} SST_r + b_{SHRD} SHRD + \log\left(\cos\phi\right)\right). \quad (3)$$

The TCGI is the expected number of genesis events. η_{850}, RH_{600}, SST_r, $SHRD$ are the absolute vorticity at 850 hPa, the relative humidity at 600 hPa, relative SST (SST relative to tropical mean SST), and vertical shear between the 850 and 200 hPa levels. b is the intercept term and b_x is the coefficient corresponding to variable x. After fitting equation (3) with 32 years of interannually varying data, we obtain a climatological relationship (b, and b_x) between observed genesis rate and the predictors. We then apply the same relationship to monthly data from 1981 to 2012 at spatial resolution of 200 km to obtain monthly TCGI.

For each seed, a genesis location and date are then chosen randomly on a 1 km resolution within the selected month. This seeding method allows the hazard model to form more than one vortex on the same day at the same location, but this situation never occurs in our simulations. By construction, the TCGI is always positive, and thus predicts a nonzero (though hopefully very small) probability of storm formation even in locations where no TC genesis events have been observed.

To evaluate the genesis model, we construct 40 (32 years) simulations for the period from 1981 to 2012. Globally, there are on average 95 storms per year and 11, 29, 23, 26, and 5 are in the ATL, WNP, ENP, SH, and IO, respectively. In the simulations, on average there are 94 storms per year with 8, 33, 18, 32, and 4 in each basin. The TCGI systematically underestimates the genesis frequency in the ENP and ATL, and overestimates it in the WNP and SH.

The spatial distributions of genesis counts in observations (Figure 1a) and in the simulations based on the TCGI (Figure 1b) are in a good agreement. The TCGI has local maxima in approximately the right locations, but with lower peak values and a smoother distribution. The observed highest TC formation rate occurs in the ENP in observations, but in the WNP in the TCGI. The simulated distribution spreads further equatorward in the WNP and IO than in the observations. The simulated formation rate in the central Pacific is higher than observed. These differences are shown quantitatively in Figure 1c using *t*

statistic (section 2). The negative genesis bias in the ENP is comparable to the natural variability with *t* statistic around 1. The biases in the tropical Atlantic and southern Indian Ocean (negative) and in the southern Pacific and subtropical WNP (positive) are both small compared to the natural variability. With the *t* statistic close to zero, the Central Pacific bias and those at the Equator are not significant, either. Additionally, Figure 1c suggests that the TCGI errors are systematic, i.e., could be corrected.

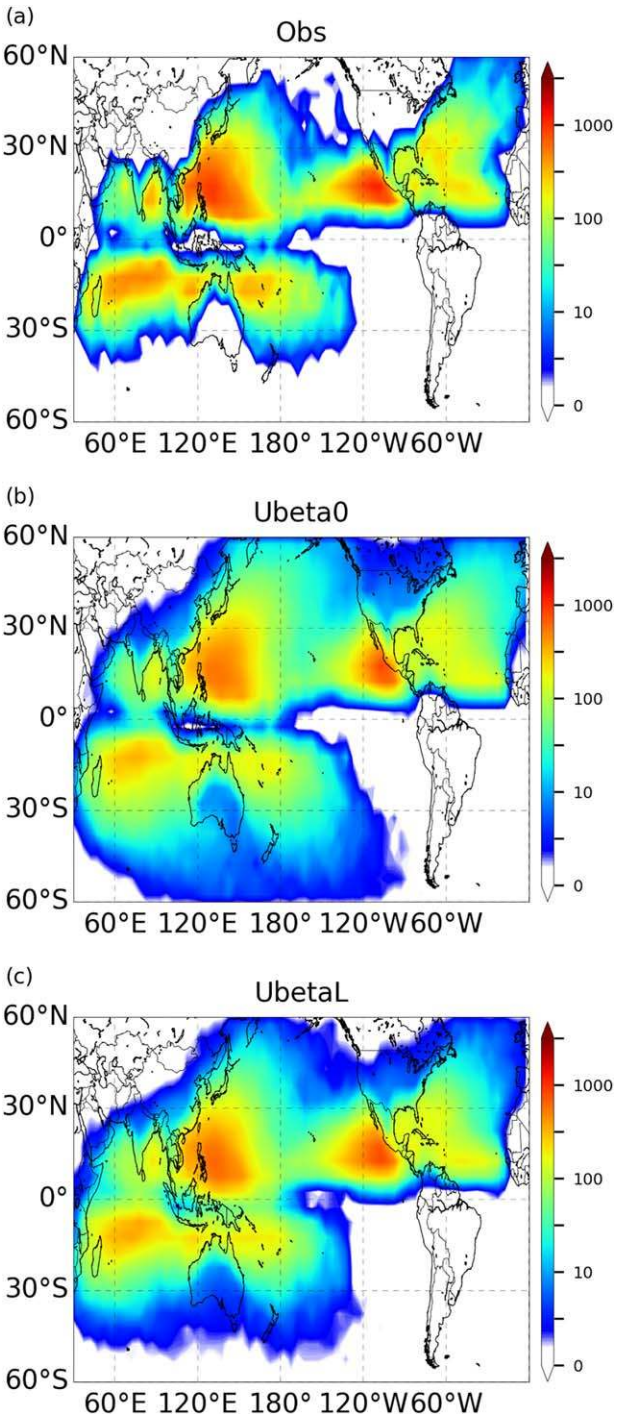

Figure 2. Track counts in every 5° × 5° box from 1981 to 2012 from (a) observations, (b) averaged from 20 BAM simulations with zero zonal beta component (Ubeta0), and (c) averaged from 20 BAM simulations with latitude-dependent beta drift (betaLat). The color scale is logarithmic. In Figures 2b and 2c, the storms' genesis locations and lifetimes are from observations.

3.2. Track—Beta-Advection Model (*BAM*)

After genesis, the track model moves the storm forward with an hourly time step. Following Emanuel et al. (2006), we use a Beta Advection Model (BAM, Marks, 1992). The BAM combines "beta drift" (Li & Wang, 1994) with mean advection based on a linear combination of the large-scale low-level (850 hPa) and upper-level (250 hPa) winds:

$$\mathbf{V}=\alpha\mathbf{V}_{850}+(1-\alpha)\mathbf{V}_{250}+\mathbf{V}_{\beta}, \tag{4}$$

V is the vector of zonal (*u*) and meridional (*v*) wind time series at 850 and 200 hPa. α is a scalar weighting the winds at these two levels, and is set to 0.8 here. $\mathbf{V}_{\beta}$ is the beta drift vector. The wind components are:

$$\begin{aligned}
u_{250}(x,y,\tau,t)&=\bar{u}_{250}(x,y,\tau)+A_{11}F_1(t)\\
v_{250}(x,y,\tau,t)&=\bar{v}_{250}(x,y,\tau)+A_{21}F_1(t)+A_{22}F_2(t)\\
u_{850}(x,y,\tau,t)&=\bar{u}_{850}(x,y,\tau)+A_{31}F_1(t)+A_{32}F_2(t)+A_{33}F_3(t)\\
v_{850}(x,y,\tau,t)&=\bar{v}_{850}(x,y,\tau)+A_{41}F_1(t)+A_{42}F_2(t)+A_{43}F_3(t)+A_{44}F_4(t),
\end{aligned} \tag{5}$$

in which $\bar{u}$ and $\bar{v}$ are daily resolution (τ) winds linearly interpolated from monthly mean fields in a *x* and *y* grid. F_1, a Fourier series variable with a random phase, represents the variability in winds for timescales smaller than monthly (Emanuel et al., 2006):

$$F_1(t)\equiv\sqrt{\frac{2}{\sum_{n=1}^{N}n^{-3}}}\sum_{n=1}^{N}n^{-3/2}\sin[2\pi(nt/T+X_n)]. \tag{6}$$

In F_1, T is the lowest frequency (15 days) in the time series (highest frequency is 1 day), N (15) is the total number of waves retained, and X_n is, for each n, a random number between 0 and 1. F_2, F_3, and F_4 have the same form as equation (6), but with different random phases, X_n. $A_{i,j}$ is the *i*th and *j*th coefficient in a lower triangular matrix **A** that satisfies

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}=\mathrm{COV}, \tag{7}$$

where COV is the covariance matrix of the flow components. **A** is function of *x*, *y*, and τ.

The coefficient $n^{-3/2}$ in equation (6) is chosen to mimic the observed spectrum of geostrophic turbulence. The power spectrum of the kinetic energy of the synthetic winds from equation (6) falls close to the inverse cube of the frequency, and is steeper than that of the steering flow based on daily winds from reanalysis data (not shown). In short, equations (5) and (6) provide synthetic winds at 850 and 250 hPa whose monthly means, variances, and covariances match those in the reanalysis data.

The observed track density is roughly in phase with the observed genesis distribution (comparing Figure 2a to Figure 1a). In order to

separate the BAM's performance from the genesis bias, we conduct 20 track realizations using the 32 years' observed genesis locations, using the simulated tracks with the same lifetimes as the best-track data.

Two experiments are conducted with different values for $\mathbf{V}_\beta$. In the first experiment, we set $\mathbf{V}_\beta = (0.0, 2.5)$ following Emanuel et al. (2006), that is, zero beta drift in the zonal direction and 2.5 m s^{-1} in the meridional direction. This setting is called "Ubeta0." A recent study by Nakamura et al. (2017) shows a systematical north-northeast-ward track bias in Emanuel's data set in the WPC. Such bias might be related to the zero beta drift, which prevents westward moving tracks. Therefore, in the second experiment we choose $\mathbf{V}_\beta$ as a function of the cosine of latitude (ϕ), with maxima of 1.0 and 2.5 m s^{-1} in zonal and meridional directions, i.e., $\mathbf{V}_\beta = (2.5 \cos\phi, 1.0 \cos\phi)$. The cosine function is used because the β-drift changes with Coriolis force (Zhao et al., 2009). We call this second experiment "betaLat."

The spatial distributions of the observed tracks and both experiments are in good agreement. This is primarily because they have the same initial locations. The spatial correlations between observations and both Ubeta0 and betaLat are very high (above 0.9). While there is no clear reason, based on these results alone, to view one as the better than the other, the fact that betaLat is more physics-based makes it more attractive, and we choose it here.

3.3. Intensity—Autoregressive (AR) Model

The AR intensity model:

$$V_{t+12h} - V_t = L(V_t, V_{t-12h}, X_t, X_{t+12h}) + \epsilon_{t+12h} \tag{8}$$

was described in our previous studies, Lee et al. (2015, 2016a). We refer readers to these two studies for details of the intensity model. Here we describe its general structure. V_t is the storm intensity at time t and X are environmental variables related to TC intensification. The deterministic component, $L(V_t, V_{t-12h}, X_t, X_{t+12h})$, has the form of a second-order vector autoregressive linear model with environmental variables as exogenous inputs. To predict intensity at $t + 12h$, L includes storm information, V_t, $V_t - V_{t-12h}$, V_t^2, and the storm translation speed. Three essential environmental variables, potential intensity (PI, Bister & Emanuel, 2002; Camargo et al., 2007a), 800–200 hPa deep-layer mean vertical wind shear (SHR, Chen et al., 2006), 500–300 hPa midlevel relative humidity (midRH), are sufficient to reasonably simulate the storm intensity statistics (Lee et al., 2015). PI enters L in the form of the difference between PI and initial storm intensity ($PI - V_t$), and its square and cubic forms: $(PI - V_t)^2$ and $(PI - V_t)^3$.

The stochastic forcing component (ϵ) accounts, in a statistically representative sense, for the internal storm dynamics or other physical processes that do not depend explicitly on the environment. In other words, ϵ is the forecast error resulting from the linear assumption and the limited variables included in L. Assuming that the forecast error is uncorrelated in time but in conditioned on the initial intensity V_t, we randomly draw ϵ from the training period errors. Lee et al. (2016a) showed that including the white-noise stochastic term improves the simulated LMI distribution as well as the spatial distribution of Cat 3–5 storms. When a storm is close to land or when it makes landfall, we switch the intensity model to the one that is fitted with an additional parameter representing the surface type in L. Gray lines in Figure 3a are the AR simulated,

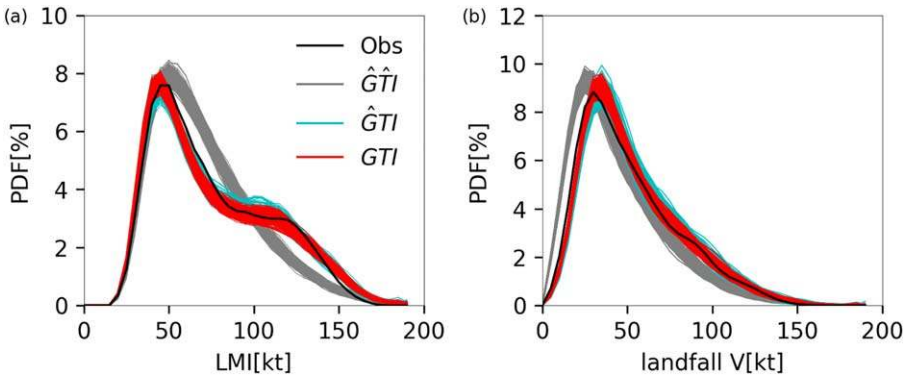

Figure 3. (a) LMI from 1981 to 2012 from observations (black), $\hat{G}\hat{T}I$ (gray), $\hat{G}TI$ (cyan), GTI (red). (b) Similar to (a) but for landfall intensity distributions. Each of the experiments contains 400 realizations.

global LMI distributions using the observed tracks and those in Figure 3b are the global landfall intensity distributions. For LMI, the AR model captures the observed (black line) first peak but not the shoulder feature at the tail part of the distribution, which we will discuss more in section 4.3. In the case of the landfall intensity, there is a small leftward shift representing a low bias in the simulations.

4. A TC Hazard Model

The next step is to integrate all three components together to form a TC hazard model and to evaluate model performance by comparing model output to observations. When all three components are fully interactive, we refer to the solutions with the label, GTI, where "G," "T," and "I" represent the genesis, track, and intensity, respectively. For each synthetic storm, the initial intensity is taken from the observed global distribution, not taking into account the basin-dependent values of initial storm intensity (15–35 kt for the ATL and ENP, 15–30 for the other basins). Dissipation is defined to occur when the intensity drops below 10 kt. We examine the storms' evolution and only keep those which intensify and reach at least tropical storm (TS) strength (LMI larger than 34 kt).

In GTI, only 70 $\pm$ 1% of seeds become TS. This is because TCGI gives a nonnegative chance for storm formation globally, which can result in some initial seeds starting very close to the Equator or in very unfavorable environments. BAM can also move the storm to an unfavorable environment since it only knows the steering flow. Both situations lead to a low bias in global TC frequency because TCGI is trained to match the genesis of tropical storms (whose lifetime maximum intensities are at least 35 kt), not the formation of the tropical disturbances that can potentially become tropical cyclones. In order to maintain a realistic global mean storm number, we revise the GTI simulations by seeding more storms than what the TCGI suggests. The adjustment factor is the multiplicative inverse of the survival rate. Because the hazard model is trained globally, it might not be able to adequately capture some regional physical relationship. As a result, the survival rate varies by basin. Nevertheless, we do not use a basin-dependent seeding rate. We apply an adjustment factor of 1.4 globally instead. We will, however, apply a local frequency adjustment when conducting hazard assessment (in section 5). Another restriction applied on the TCGI is that we do not allow storms to form within 2° from the Equator, globally.

GTI, with the adjustment factor, generates synthetic storms whose climatology is generally good agreement in most areas with the observed one (Figure 4). They both have more intense storms in the WNP and less in the ATL, a westward followed by a north-eastward movement in the northern hemisphere, and almost no

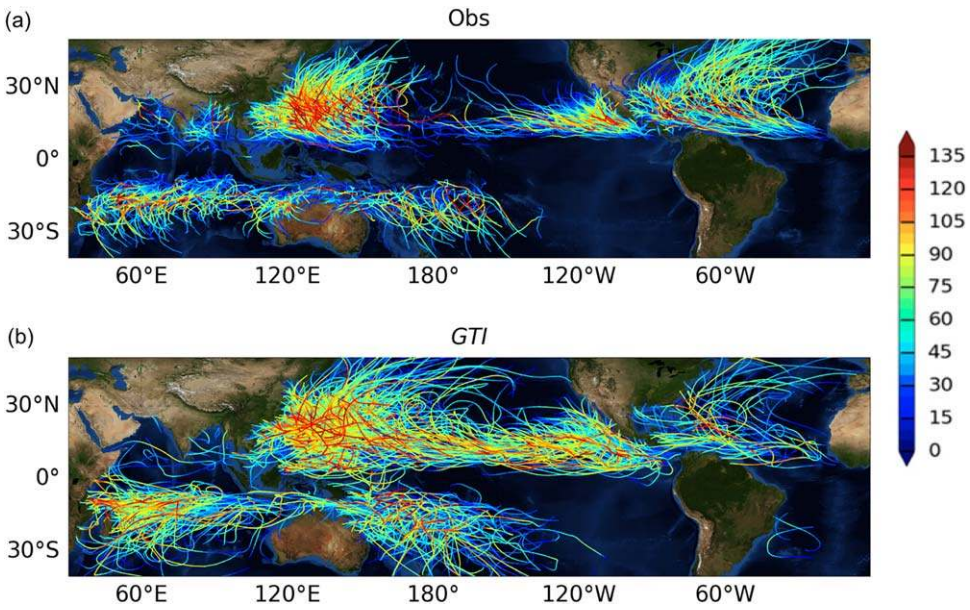

Figure 4. (a) 2000–2012 historical tracks color-coded by intensity. (b) Similar to (a) but from a randomly selected member (out of 400 realizations) from GTI.

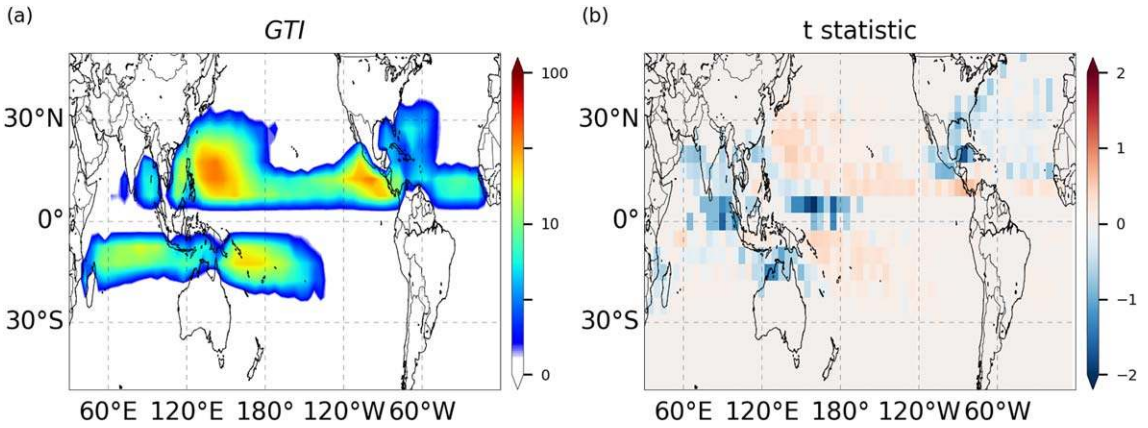

Figure 5. (a) Number of TC genesis per 5° × 5° box averaged over 400 GTI simulations. (b) *t* statistic of GTI simulations. The color scale is logarithmic in Figure 5a and linear in Figure 5b.

storms in the southeastern Pacific and southern Atlantic. There are some differences as well, such as more central Pacific storms and a less pronounced equatorial gap in the simulations. In addition to GTI, we designed two more experiments to isolate the influence of individual components on the total estimated TC statistics: $\hat{G}$TI, and $\hat{G}\hat{T}$I. When ($\hat{}$) is used above these letters, observational data are used instead of simulations. We construct 400 realizations of 32 year global simulations (1981–2012) in each experiment (see section 2 for details). In $\hat{G}$TI and $\hat{G}\hat{T}$I, the observed initial intensities are used for the corresponding formation locations.

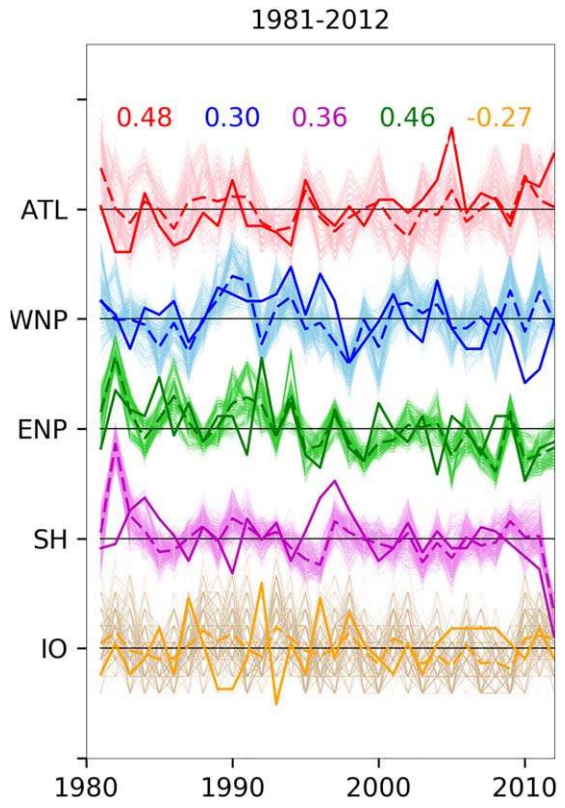

Figure 6. Interannual variability of storm genesis in ATL (red), WNP (blue), ENP (green), SH (purple), IO (yellow). The observed time series are in thick solid lines while the GTI simulated ones are in thin solid lines with the thick dashed lines representing the ensemble mean. Data are normalized by the corresponding mean and standard deviation and the black lines are the reference lines (i.e., zero). The interval between two black horizontal lines is four standard deviations. The correlation coefficient between observations and ensemble means in individual basins are given at the top of the figure.

4.1. Genesis Density and Interannual Variability

$\hat{G}\hat{T}$I and $\hat{G}$TI genesis climatologies (not shown) are similar to the observed one because best-track genesis locations were used in these simulations. Similarly, the spatial distribution of genesis location in GTI (Figure 5a) is close to TCGI (Figure 1b) in section 3. This is because, while the track and intensity models determine the survival of initial vortices, they do not largely alter the genesis climatology. They do, however, slightly enhance the positive bias in the central Pacific (comparing Figure 1c and Figure 5b), which might be due to too many storms surviving in the central Pacific in GTI. On the other hand, GTI has larger negative biases in the IO than TCGI does, perhaps because many seeded storms die too quickly there. Another noticeable difference between Figure 1c and Figure 5b is the large negative bias in the equatorial region (especially over the Pacific and Indian Oceans) where, by design, no simulated storms are allowed to form. Thus, the simulated standard deviation (i.e., σ_m in equation (2)) is very small here, and the small negative biases between observations and simulations are enhanced in the *t* statistic.

The interannual correlation of total storm number in GTI with that in observations is shown in Figure 6. The correlation coefficient for ATL hurricanes in GTI is 0.48, similar to Emanuel et al. (2008) while with the new intensity model, it increases to 0.7 in Emanuel (2017). The correlation coefficient for WNP, ENP, SH, and IO in GTI are 0.30, 0.36, 0.46, and −0.27, respectively. Menkes et al. (2012) found that existing genesis indices, including the TCGI, do not capture the full spectrum of interannual variability in storm frequency well, although they are all able to simulate the impact of ENSO. This deficiency is inherited in our model.

4.2. Track and Landfall Frequency

The track density plots from observations, $\hat{G}$TI and GTI are shown in Figures 2a, 7a, and 7c. The $\hat{G}\hat{T}$I track density is the same as in observations and is not shown here. In both observations and simulations, the highest value of the track density are in the WNP and ENP, followed by the southern Indian Ocean and the western South Pacific. Both simulations show the typical observed recurvature track pattern in the ATL. The relatively high track densities over northwestern Australia and the Bay of Bengal, however, are missing in the simulations. A comparison between *t* statistic from $\hat{G}$TI (Figure 7b) and GTI (Figure 7d) suggests that the negative frequency bias in the ENP is partially due to the TCGI, consistent with the results from section 3.1. The positive frequency bias in the central Pacific, which is also seen in the genesis *t* statistic in Figure 5b, extends further northward in Figure 7d. The large negative values of t-statistic over land in Figures 7b and 7d are from very small simulated standard due to the design of the hazard model—the AR intensity model stops right after storms making landfall and thus leads to almost no variability over land. Similarly to the negative *t* statistic in the equatorial region in Figure 5b discussed in the previous section, very small σ_m in equation (2) leads to large *t* statistic.

The regional landfall frequency in observations and in the GTI simulation are shown in Figures 8 and 9. There are low biases in the coastal regions of the northern Indian Ocean and the north Atlantic from Mexico to New England, where the observed (black) frequency is constantly above the simulated spread (red patches). The rank histograms (section 2) also tilt toward high ranks in these regions (Figure 10). In these regions, we also see negative biases in the track density (Figure 7d). In Taiwan and the Philippines, there are positive landfall frequency biases and the rank histogram distributions tilt toward low ranks. The track density plot, however, shows a negative bias near Taiwan. This inconsistency between biases in track density and landfall frequency occurs because the landfall frequency is calculated at much finer spatial resolution (50 km) than is the track density (about 500 km). Thus, Taiwan covers only part of a large grid box in the track density plot. Another possible reason is that landfall is related to the direction in which a storm is

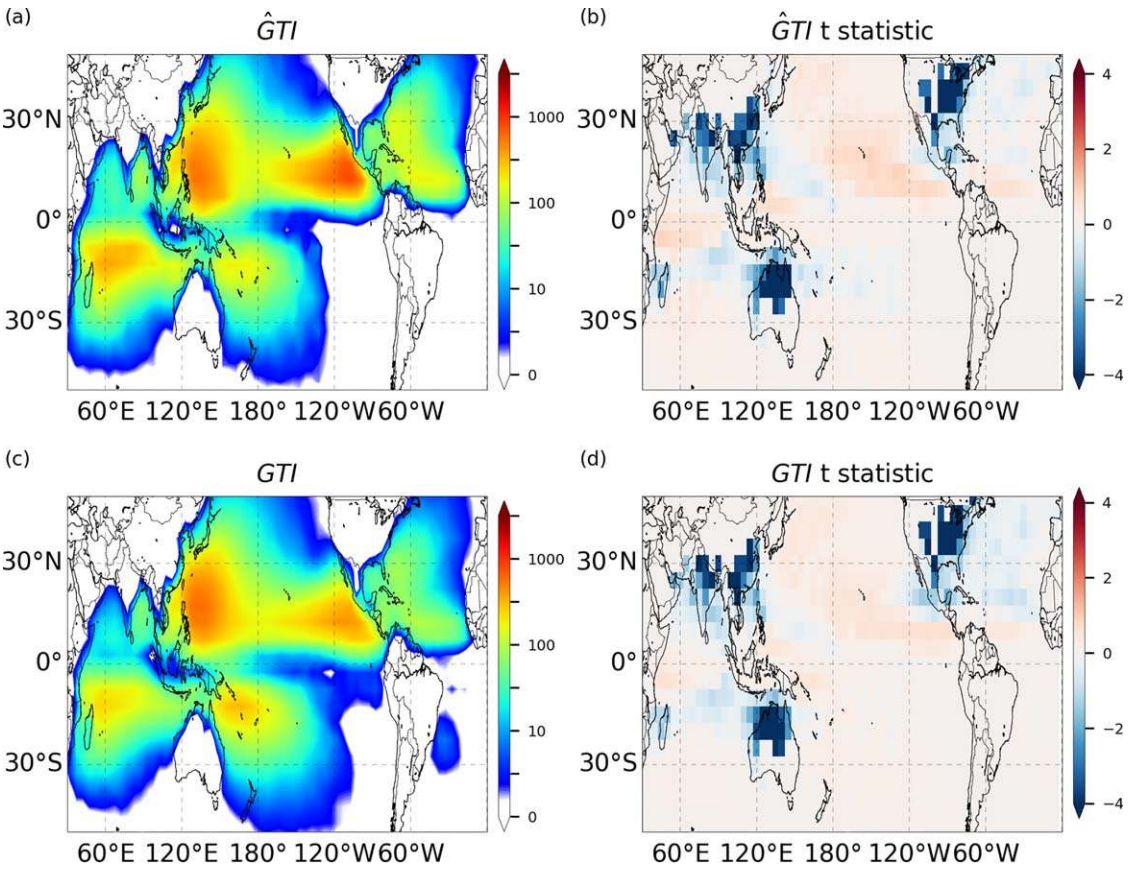

Figure 7. (a) Simulated 1981–2012 TC track counts per 5° × 5° box from 400 ensemble mean from $\hat{G}$TI. (b) *t* statistic of the $\hat{G}$TI simulations. (c) Similar to (a) but from GTI. (d) *t* statistic of the GTI simulations. The scales are logarithmic in Figures 7a and 7c and linear in Figures 7b and 7d.

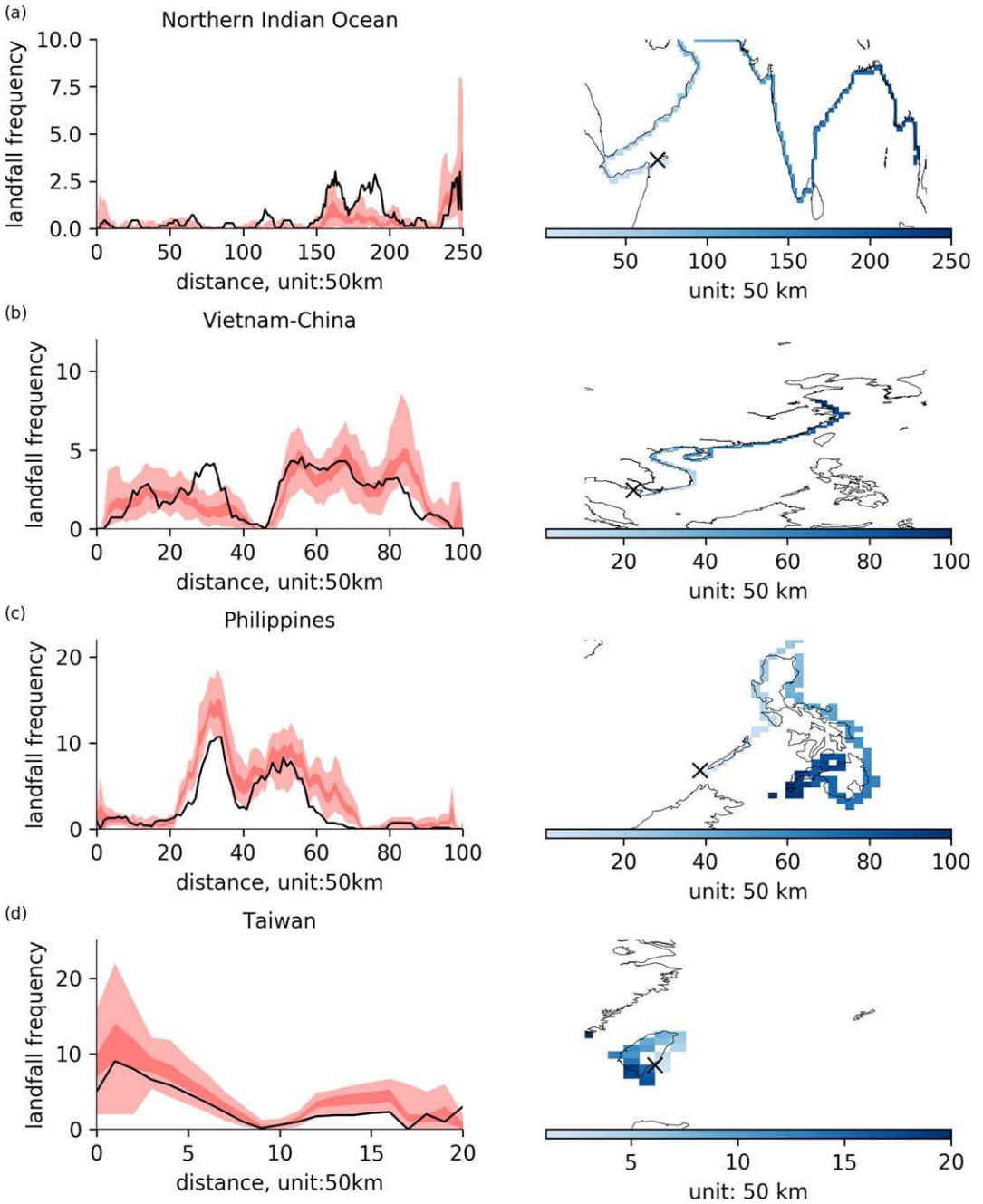

Figure 8. Observed (black line) and GTI simulated (red shading) landfall frequencies in number of occurrences at every 50 km along the coastline of (a) Northern Indian Ocean, (b) Vietnam to China, (c) the Philippines, (d) Taiwan. The simulated landfall frequencies are shown as 0, 25, 75, 100 percentile based on the 400 realizations. *X* axis in each panel matches with colors along the corresponding coastline in the map on the right, starting from "X" symbol. The color is darkker with increasing distance.

moving. A low track frequency does not necessarily result in a low landfall occurrence if the number of westward moving tracks is higher. The simulated landfall frequency is unbiased in the coastal regions from Vietnam to China.

4.3. LMI and Landfall Intensity

In Figure 3a, $\hat{GTI}$ captures the first peak of the global LMI distribution well, but misses the shoulder feature due to an insufficient number of simulated RI storms (TCs that intensify, at least once, more than 35 kt

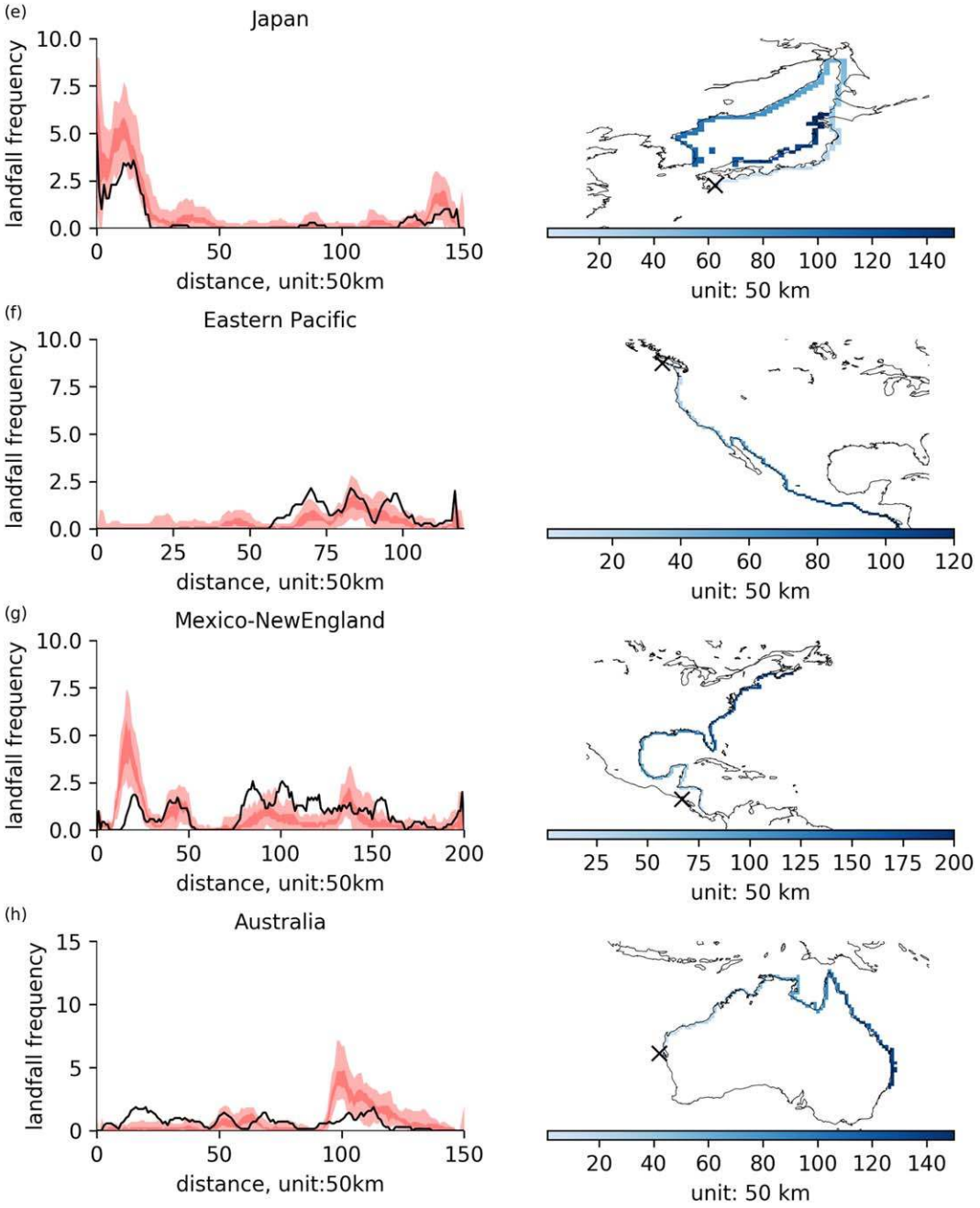

Figure 9. Continuation of Figure 8. (e) Japan, (f) Eastern Pacific, (g) Mexico to New England, (h) northern Australia.

within 24 h in their lifetimes), consistent with the results in Lee et al. (2016a). PDFs of LMI from $\hat{G}TI$ (light blue lines) and GTI (red lines), however, successfully capture the shoulder feature. This improvement has a simple explanation, namely the consistency between track and intensity evolution. In $\hat{G}TI$, each synthetic storm ends when the observed record ends, regardless of the storm's intensity at that time. As a result, some die while they are intensifying, or still at or above TS level, and thus are artificially denied future opportunities to undergo RI. Coupling the intensity model to the track model (in $\hat{G}TI$ and GTI) resolves this artifact by giving each synthetic storm a self-consistent opportunity to undergo RI when the environment permits. Thus, GTI and $\hat{G}TI$ generate numbers of RI storms close to those found in observations (comparing the pink lines to red line in Figure 11) and match the observed LMI distribution. The successful simulation of RI storms shows that the stochastic forcing in the intensity model, as proposed in Lee et al. (2016a), is an

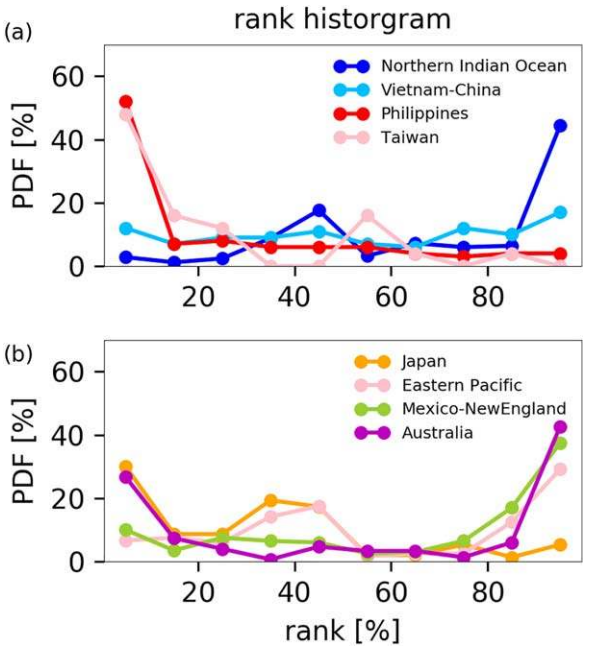

Figure 10. The normalized rank histogram from landfall frequencies from Figures 8 and 9.

effective way to produce RI storms. As we have mentioned in the section 1, the shoulder feature of the global LMI distribution (or the bimodal distribution at individual basins) might be due to a combination two factors: RI and the artifacts of the Dvorak technique. Given that the observed number of RI storms is not built into the TC Hazard model by construction, our results give further evidence that RI and the LMI distribution are related and that the bimodality is not purely an artifact of the Dvorak technique.

PDFs of the landfall intensity (Figure 3b) in GTI and $\hat{GTI}$ are almost indistinguishable from the observed one. Coupling between track and intensity model improves not only the simulation of peak intensities, but the intensity evolution throughout storms' lifetimes as well, including at landfall.

5. Tropical Cyclone Hazard in the Current Climate

In sections 4.3 and 4.2, we discussed the performance of GTI in predicting TC landfall frequency and intensity, respectively. When considering hazard, however, it is essential to use joint measures that contain information about both of them. Thus, here we define TC hazard as the probability of the landfall intensity's exceeding a given threshold at a particular location. TC hazard will be calculated based on the historical record, and synthetic storms from the three simulations, namely $\hat{GTI}$, $\hat{GTI}$, and GTI. We will discuss TC hazard from both global and regional perspectives.

5.1. Global Map of Return Period

Figures 12 and 13 show global maps of return period for hurricanes (Cat1+ storms) and major hurricanes (Cat3+) in observations and simulations. To maintain consistency in the observations and simulations sample size, we select only one realization (32 years of data) for each experiment in Figures 12b–12d and 13b–13d. The features we discuss here, however, are also true for the ensemble mean of our simulations (not shown).

At the coastal regions in the south western WNP (southeastern China, Taiwan, and the Philippines), the observed return period of hurricanes is less than 10 years; it is close to 2–3 years near Taiwan and the Philippines. Another distinct area with a low return period (high hazard) is the ENP. In the southern hemisphere, the 10 year return period contours reach eastern Madagascar. In the Northern Australia, Bay of Bengal and most US coastal regions, the return period for Cat1+ storms is on the order of decades. Because TCs are rare events, the "observed hazard" does not actually represent the true hazard, but is what we estimate based on the 32 years of the reliable historical observations.

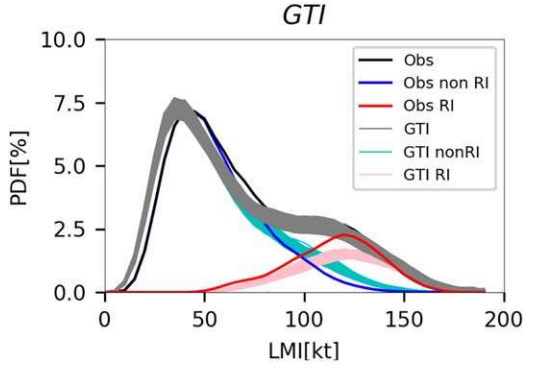

Figure 11. PDF of LMI from 1981 to 2012 global historical record (black) and from 400 GTI realizations (gray). Blue and red lines are PDFs using subsets of non-RI and RI storms from observations. The cyan and pink lines show the same quantities but from simulations.

Comparing the simulated return period maps of hurricane strength in Figure 12 shows the advantages of using observed tracks and formation locations. Figure 12b is much closer to Figure 12a (observations), than Figures 12c ($\hat{GTI}$) and 12d (GTI) are. Compared to observations, some of the biases in the TC climatology discussed earlier are reflected in the return period map. For example, the difference map between GTI and observations (Figure 12e) shows that GTI estimates a higher hazard (shorter return period) in the central Pacific than do the observations. This difference is related to the overestimation of storm activity in that area shown in Figure 7d. In the IO, the simulated hazard is smaller (longer return period) than the observed one, due to the low frequency and the low intensity biases. Despite these differences in detail, GTI captures the high hazard regions for hurricane strength storms in most of the global TC basins, such as WNP, ENP, and Southern Indian Ocean.

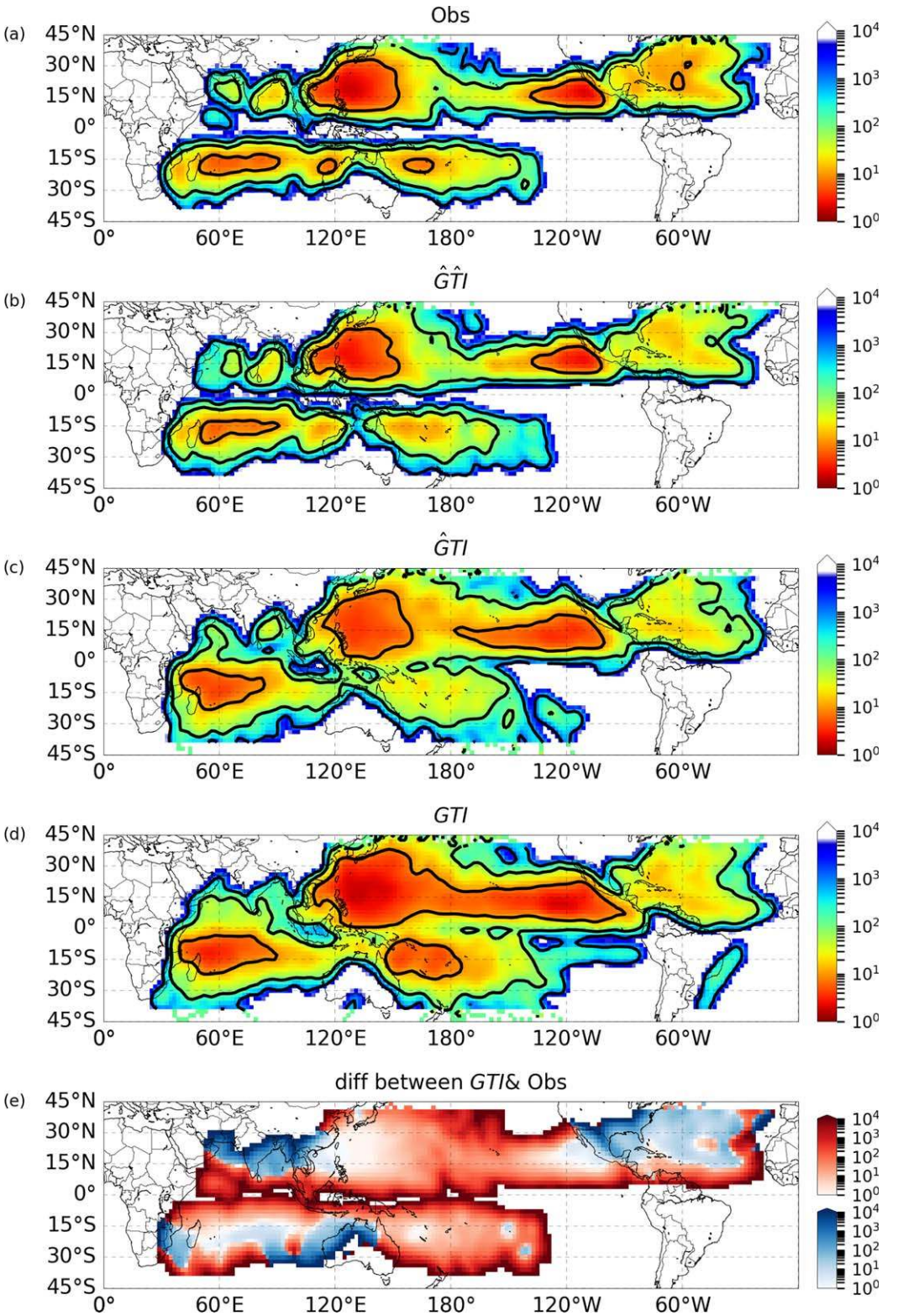

Figure 12. Return period map for storms exceeding Category 1 hurricane strength from (a) 1981–2012 observations, and a 32 year simulations from (b) $\hat{GTI}$, (c) $\hat{GTI}$, and (d) GTI. Data are calculated in $2° \times 2°$, and a Gaussian smoothing is applied with length scale of 3 grid points. The contours represent return periods of 10, 100, and 1,000 years. (e) The difference maps between (a) and (d), where blue (red) tones represent areas in which the hazard is underestimated (overestimated) in GTI.

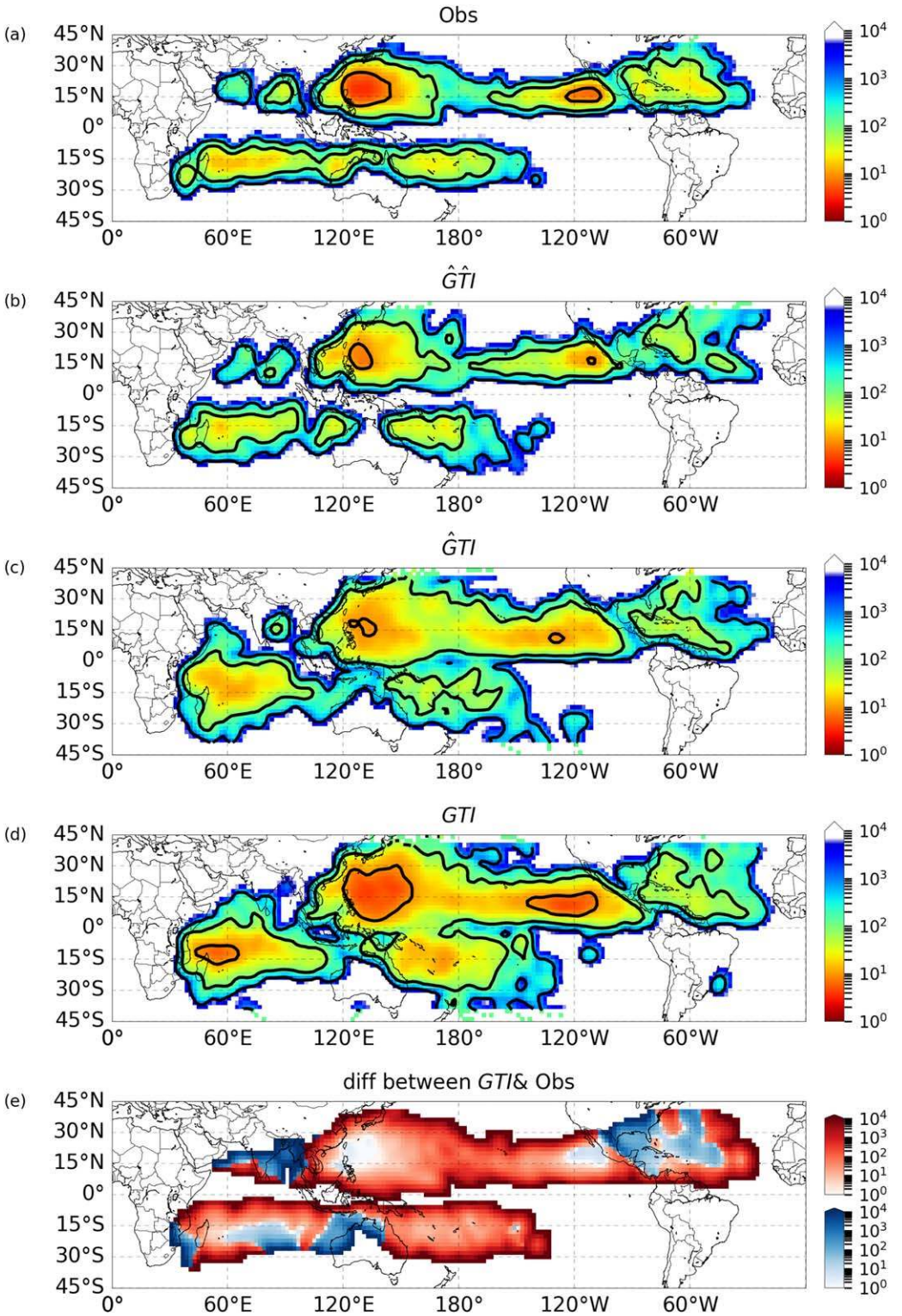

Figure 13. Similar to Figure 12 but for storms exceeding Category 3 hurricane strength.

The return period map of Cat3+ storms (Figure 13) shows the advantages of calculating storm evolution in a consistent environment, i.e., in GTI, for more rare events. $\hat{GTI}$ underestimates the Cat3+ storm hazard, especially in the WNP. GTI, on the other hand, reasonably captures the global hazard of Cat3+. This is again because GTI is able to simulate sufficient numbers of RI storms.

5.2. Regional Return Period

To discuss TC hazard at regional scales, we select 13 sub-basin areas and calculate the return period curves as a function of landfall intensity. The 13 chosen areas are the coastal regions of Madagascar, Bay of Bengal, Vietnam, China, the Philippines, Taiwan, Japan, western Mexico, Caribbean islands, Gulf of Mexico, eastern US, Pacific islands (Papua New Guinea and eastern Indonesia) and northern Australia (Figures 14 and 15). The observed return period curves, especially those for the strongest landfall intensity thresholds, are limited by the available observations. In simulations, the spread of simulated hazard (in color patches) increases with intensity because the low-intensity landfall hazard is mostly controlled by the tracks. $\hat{G}\hat{T}I$ has only one set of tracks by construction—i.e., the observed tracks—and therefore has almost no spread. In

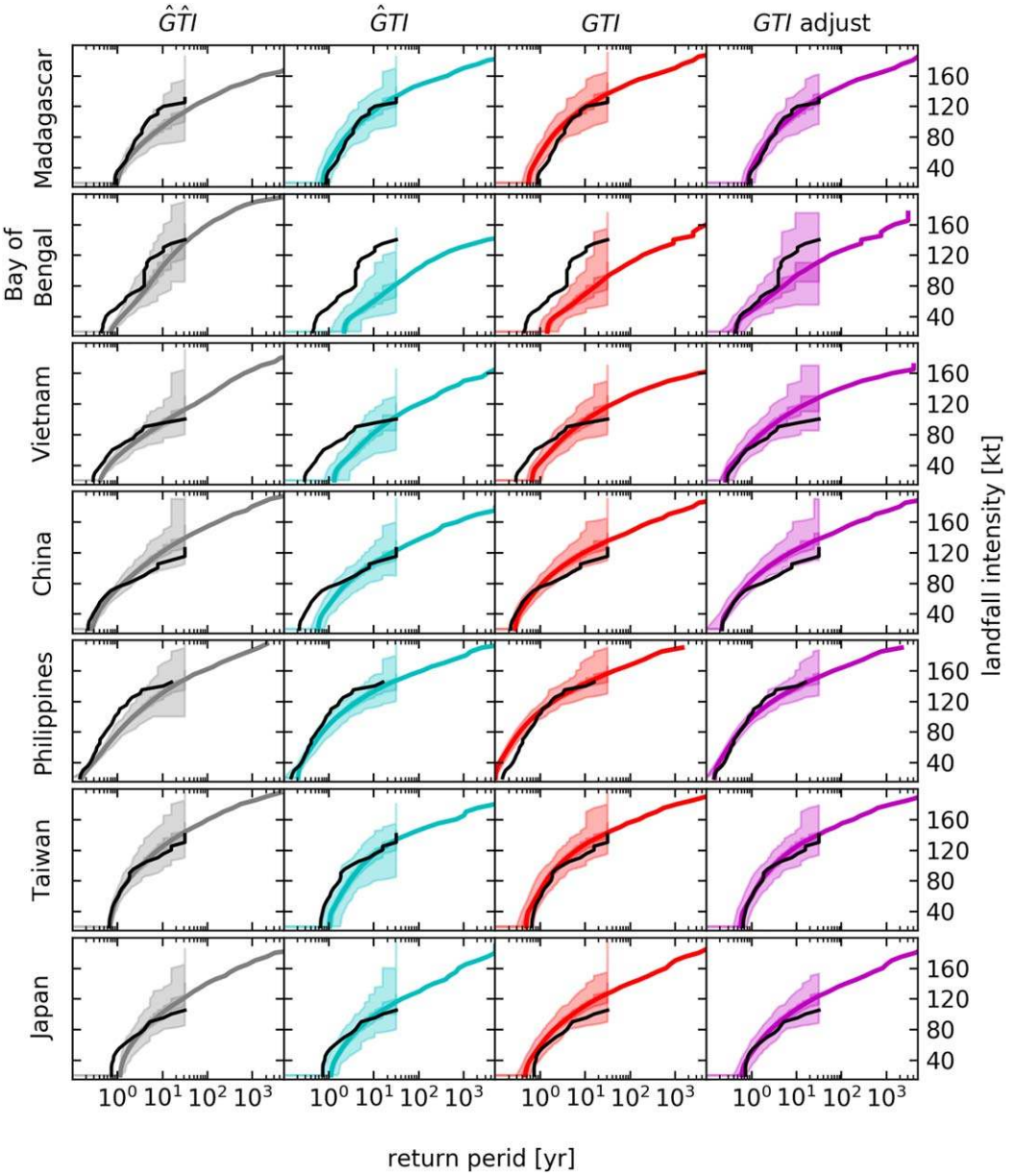

Figure 14. Return period curves of landfall intensity in (a) Madagascar, (b) Bay of Bengal, (c) Vietnam, (d) China, (e) the Philippines, (f) Taiwan, and (g) Japan from observations (black), and 400 ensemble members from simulations (from left to right): $\hat{G}\hat{T}I$ (gray shading), $\hat{G}TI$ (cyan shading), GTI (red shading), and regional frequency-adjusted GTI (purple shading). The lighter (darker) patches represent 0 to 100 (25–75) percentiles. The solid lines are return period curves calculated using all 12,800 year simulations.

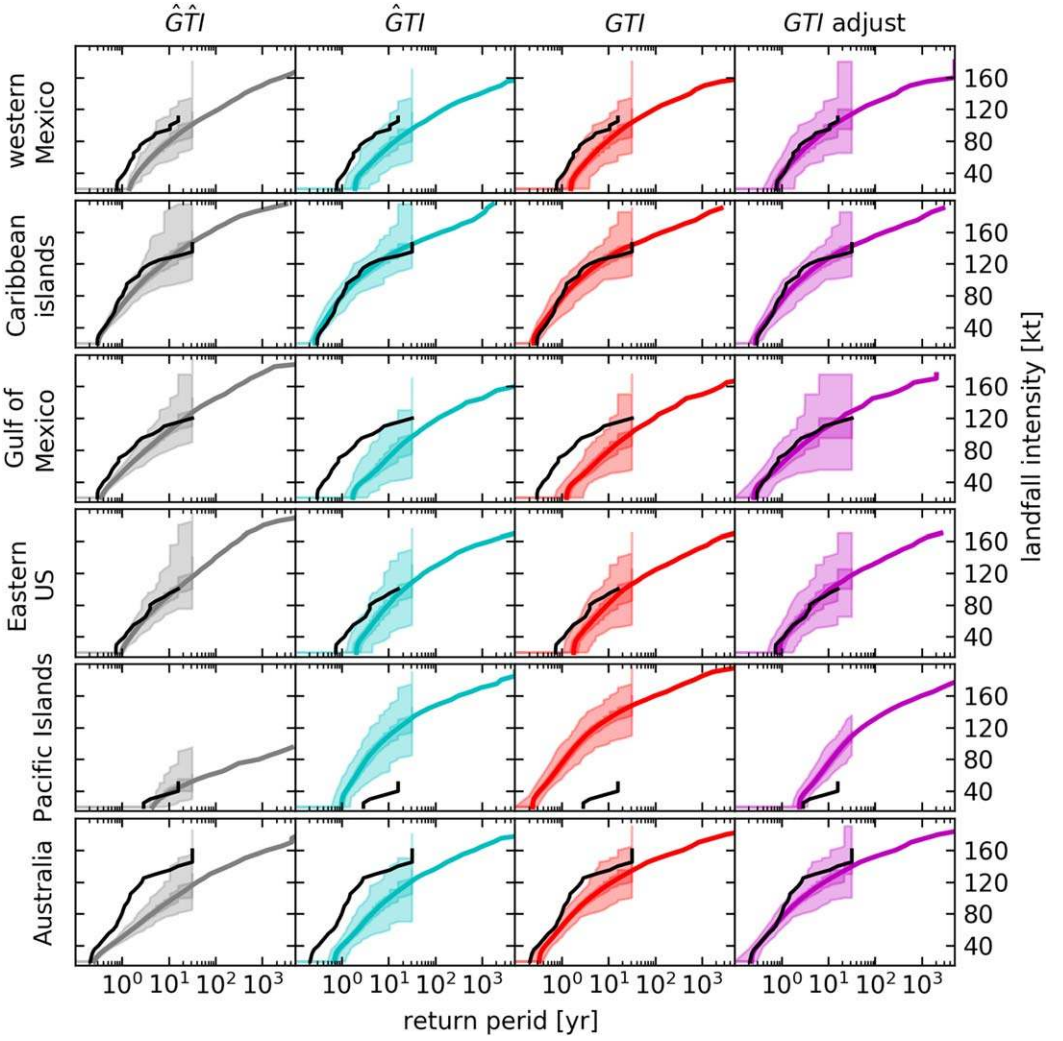

Figure 15. Continuation of Figure 14. (h) Western Mexico, (i) Caribbean islands, (j) Gulf of Mexico, (k) Eastern US, (l) Pacific islands, and (m) Australia.

the $\hat{GTI}$ and GTI, there are 10 sets of tracks that contribute to the spread. At higher intensity thresholds, the intensity model ensemble contributes to the spread of the hazard estimation.

Ideally, the historic return period curve falls within the ensemble spread of the simulated curves, an indication of an unbiased model. Biases in the return period curves have at least two general sources: landfall frequency (the location of the curves) and intensity (the shape of the curves). Model curves shifted toward the right (left) with steeper (lower) slope can be interpreted as underestimation (overestimation) of TC hazard. The observed return period curves (black) lay in the simulated spreads of $\hat{G}\hat{T}I$ (gray patches) in most places. Although the observed tracks are used in $\hat{G}\hat{T}I$, there are still shifts toward the right in the simulated return period curves in the Bay of Bengal, Vietnam, Japan, western Mexico, indicating that some of the observed landfalling storms dissipate in simulations before making landfall. In Australia and the Philippines, $\hat{G}\hat{T}I$ underestimates landfall intensities. Including the track model ($\hat{G}TI$, cyan patches) results in underestimations in most places, except in the Pacific islands where $\hat{G}TI$ has more landfalls than the observations do. This is consistent with the equatorial bias discussed in section 4.

Using the same environmental conditions for genesis, track, and intensity (GTI, red patches) improves the estimated return period curves. There is a small landfall frequency bias in the coastal regions of China, Taiwan, Japan, and Caribbean islands. GTI underestimates the landfall frequency in Madagascar, Vietnam, western Mexico, Gulf of Mexico, and eastern US. The bias is largest in Gulf of Mexico, followed by Bay of Bengal.

Furthermore, GTI results in too many landfall events in the Pacific islands due to the equatorward bias in the southern Pacific.

In order to bias-correct the frequency locally, we shift the return period curves (GTI adjust, purple patches) to match the observed return periods at the lowest threshold (35 kt), which is the threshold with the most observations, and is more reliable than the higher intensity thresholds. After shifting the simulated curves, the observed curves fall within the spread in the simulations for most of the regions, staying within the 25–75 percentiles (darker purple patches), except in China and the Pacific islands. In China, the observed return period curve for landfall intensities larger than 80 kt is at the low edge of the simulated spread, i.e., the hazard is overestimated. The overestimation is much more severe in the Pacific islands.

6. Summary and Discussion

This study describes a new, environmentally forced tropical cyclone (TC) hazard model. It is composed of three model components that, together, represent the complete storm lifetime: a genesis model (TCGI), a beta-advection track model (BAM) and an auto-regressive (AR) intensity model. The TCGI and BAM are developed following Tippett et al. (2011) and Emanuel et al. (2006), respectively, while the AR intensity model is from our previous work (Lee et al., 2016a). The TCGI defines the spatial and temporal formation rate (i.e., the numbers of storms that should form at a given location within a given period) using the observed climatological relationship between storm formation and absolute vorticity, relative humidity, relative sea surface temperature, and vertical shear (section 3.1). After the initial seeding, the BAM moves vortices following the synthetic steering flow (section 3.2). The synthetic wind has the statistics of the monthly averaged winds but also contains high-frequency perturbations calculated from the daily variance and covariance. The intensity model predicts the storm's evolution using a deterministic multiple linear regression plus a stochastic component (section 3.3). In the deterministic component of the intensity model, the TC intensity change is a function of potential intensity, deep-layer mean vertical wind shear, midlevel relative humidity, and storm intensity persistence. The stochastic component represents the physical processes that are not considered in the deterministic model and is necessary in order for the intensity model to simulate the observed distribution of TC intensity.

The model captures many aspects of TC genesis, track, intensity, and landfall statistics, including their density distributions, probability density function (PDF) of storms' lifetime maximum intensity (LMI) and landfall intensity, as well as the landfall frequency. The model has a positive frequency bias in the central Pacific and in the equatorial region. A particularly interesting result is that it captures the observed LMI PDF, which has a main peak and a" shoulder" at higher intensities. This finding is different from those in our previous study, Lee et al. (2016a), in which the realizations were conducted using the AR intensity model and observed tracks. The observed shoulder feature in the global LMI PDF (the regional LMI PDF are bimodal) appears to be due to the separation in two monomodal PDFs, one from storms which undergo rapid intensification (RI, intensity change larger than 35 kt per 24 h) and the other one from those which do not (Lee et al., 2016b). While the AR intensity model running along the observed tracks is able to simulate RI storms, it does not generate as many RI storms as are found in observations. The reason for this underestimation is that some of the synthetic storms end when the observed track ends regardless of their intensities, which artificially reduces the probability of RI. Combining the AR intensity model and the BAM track model resolves the inconsistency, and gives the synthetic storms opportunity to undergo RI when the environment permits (Figure 11). Self-consistent tracks and intensities improve not only the LMI distribution but the storms' lifetime intensities, and therefore also landfall intensities.

With the well-simulated TC climatology, the model can estimate regional TC hazard reasonably well. However, it predicts more landfalls in the western North Pacific and Pacific islands, and fewer landfalls in the northern Atlantic and Indian Ocean than are found in observations. These landfall biases lead to biases in the estimated TC hazards. A possible causes of the landfall biases is the design of the BAM, especially the beta drift, which does not take into account the storm strength. Further investigation into this issue will be carried out in the future. For now, we have corrected these biases during the post-processing with a local frequency adjustment. The large positive hazard bias for the Pacific islands, however, remains, because the model generates too many strong landfalling storms there. These and other biases in the TC hazard can be corrected to some extent, so that the TC hazard model can generate estimates of the probability of landfall

at a given intensity that are in agreement with observations at shorter return periods, while also giving estimates at longer return periods where such estimates cannot be directly generated from observations.

While the environmental parameters used here are obtained from reanalysis, they can potentially be obtained instead from a global climate model. However, when assessing hazard in a changing climate, it may be appropriate to choose somewhat different predictors. For example, Camargo et al. (2014) showed that using saturation deficit and potential humidity allows for a better representation of the response to mean climate warming than using relative humidity, although both indices have similar behavior in the current climate. Similarly, we might want to use PI instead of relative SST in the genesis model in the future. Parameters used in the intensity model might need some adjustments as well. Preliminary results using one of the CMIP5 models (not shown) suggest that the TC hazard model is able to produce reasonable TC climatologies in both current and future climates. One of the challenging issues will be how to make appropriate bias corrections in the required predictors obtained from different climate models. Application of our model in such a climate change context, forced by a range of global climate models, will be presented in a future publication.

Acknowledgments
The sources of data supporting the model development are described in the section 2.1: Observational and reanalysis data sets and can be download from: ERA-Interim (http://apps.ecmwf.int/datasets/data/interimfull-moda/levtype5pl/); HURDAT2 (http://www.nhc. http://www.nhc.noaa.gov/data/#hurdat), JTWF best-track (http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks). Comments and suggestions from the Reviewers, Drs. Jim Kossin and Paul Della-Marta, were appreciated. The research was supported by the the New York State Energy Research and Development Authority under the research grant NYSERDA 103862. Authors declare no conflict of interest.

References

AIR WORLDWIDE (2015). *The AIR hurricane model: AIR Atlantic tropical cyclone model V15.0.1 as implemented in Touchstone V3.0.0* (technical report). AIR Worldwide. Retrieved from https://www.sbafla.com/method/portals/methodology/ModelSubmissions/2015/20150520_AIR_2014_FCHLPM_submission.pdf

Bister, M., & Emanuel, K. A. (2002). Low frequency variability of tropical cyclone potential intensity 1. Interannual to interdecadal variability. *Journal of Geophysical Research: Atmospheres*, *107*(D22), 4621. https:/doi.org/10.1029/2001JD000780

Bruyère, C. L., Holland, G. J., & Towler, E. (2012). Investigating the use of a genesis potential index for tropical cyclones in the North Atlantic basin. *Journal of Climate*, *25*, 8611–8626.

Camargo, S. J., Emanuel, K. A., & Sobel, A. H. (2007b). Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. *Journal of Climate*, *20*, 4819–4834.

Camargo, S. J., Sobel, A. H., Barnston, A. G., & Emanuel, K. A. (2007a). Tropical cyclone genesis potential index in climate models. *Tellus, Series A*, *59*, 428–443.

Camargo, S. J., Tippett, M. K., Sobel, A. H., Vecchi, G. A., & Zhao, M. (2014). Testing the performance of tropical cyclone genesis indices in future climates using the HiRAM model. *Journal of Climate*, *27*, 9171–9196.

Chen, S. S., Knaff, J. A., & Marks, F. D. (2006). Effects of vertical wind shear and storm motion on tropical cyclone rainfall asymmetries deduced from TRMM. *Monthly Weather Review*, *134*, 3190–3208.

Chu, J.-H., Sampson, C. R., Lavine, A., & Fukada, E. (2002). *The Joint Typhoon Warning Center tropical cyclone best-tracks, 1945–2000* (Tech. Rep. NRL/MR/7540-02-16, 22 p.). Washington, DC: Naval Research Laboratory.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of Royal Meteorological Society*, *137*, 553–597.

Emanuel, K. A. (2010). Tropical cyclone activity downscaled from NOAA-CIRES reanalysis, 1908–1958. *Journal of Advances in Modeling Earth Systems*, *2*, 1. https://doi.org/10.3894/JAMES.2010.2.1

Emanuel, K. A. (2015). Effect of upper-ocean evolution on projected trends in tropical cyclone activity. *Journal of Climate*, *28*, 8165–8170.

Emanuel, K. A. (2017). A fast intensity simulator for tropical cyclone risk analysis. *Natural Hazards*, *88*, 779–796.

Emanuel, K. A., DesAutels, C., Holloway, C., & Korty, R. (2004). Environmental control of tropical cyclone intensity. *Journal of Atmospheric Sciences*, *61*, 843–858.

Emanuel, K. A., Sundararajan, R., & Williams, J. (2008). Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bulletin of the American Meteorological Society*, *89*, 347–367.

Emanuel, K. A. (2013). Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 12219–12224.

Emanuel, K. A., & Nolan, D. S. (2004). Tropical cyclone activity and global climate. *Bulletin of the American Meteorological Society*, *85*, 666–667.

Emanuel, K. A., Ravela, S., Vivant, E., & Risi, C. (2006). A statistical determinatic approach to hurricane risk assessment. *Bulletin of the American Meteorological Society*, *87*, 299–314.

Estrada, F., Botzen, W. J. W., & Tol, R. S. J. (2015). Economic losses from US hurricanes consistent with an influence from climate change. *Nature Geoscience*, *8*, 880–884.

Geiger, T., Frieler, K., & Levermann, A. (2016). High-income does not protect against hurricane losses. *Environmental Research Letters*, *11*, 084012. https://doi.org/10.1088/1748-9326/11/8/084012

Hall, T., & Jewson, S. (2007). Statistical modelling of North Atlantic tropical cyclone tracks. *Tellus, Series A*, *59*, 486–498.

Kossin, J. P., Olander, T. L., & Knapp, K. R. (2013). Trend analysis with a new global record of tropical cyclone intensity. *Journal of Climate*, *26*, 9960–9976.

Landsea, C. W., & Franklin, J. L. (2013). Atlantic hurricane database uncertainty and presentation of a new database format. *Monthly Weather Review*, *141*, 3576–3592.

Lee, C.-Y., Tippett, M. K., Camargo, S. J., & Sobel, A. H. (2015). Probabilistic multiple linear regression modeling for tropical cyclone intensity. *Monthly Weather Review*, *143*, 933–954.

Lee, C.-Y., Tippett, M. K., Sobel, A. H., & Camargo, S. J. (2016a). Autoregressive modeling for tropical cyclone intensity climatology. *Journal of Climate*, *29*, 7815–7830.

Lee, C.-Y., Tippett, M. K., Sobel, A. H., & Camargo, S. J. (2016b). Rapid intensification and the bimodal distribution of tropical cyclone intensity. *Nature Communication*, *7*, 10625. https://doi.org/10.1038/ncomms10625

Li, X., & Wang, B. (1994). Barotropic dynamics of the beta gyres and beta drift. *Journal of Atmospheric Sciences*, *51*, 746–756.

Lin, N., Emanuel, K., Oppenheimer, M., & Vanmarcke, E. (2012). Physically based assessment of hurricane surge threat under climate change. *Nature Climate Change*, *2*, 462–467.

Lin, N., Jing, R., Wang, Y., Yonekura, E., Fan, J., & Xue, L. (2017). A statistical investigation of the dependence of tropical cyclone intensity change on the surrounding environment. *Monthly Weather Review*, *145*, 2813–2831.

Marks, D. G. (1992). *The beta and advection model for hurricane track forecasting* (NOAA Tech. Memo. NWS NMC 70, 89 pp.). Camp Springs, MD: National Meteorological Center.

McGauley, M. G., & Nolan, D. S. (2011). Measuring environmental favorability for tropical cyclogenesis by statistical analysis of threshold parameters. *Journal of Climate*, *24*, 5968–5997.

Menkes, C. E., Lengaigne, M., & Marchesiello, P. (2012). Comparison of tropical cyclogenesis indices on seasonal to interannual timescales. *Climate Dynamics*, *38*, 301–321.

Nakamura, J., Camargo, S. J., Sobel, A. H., Emanuel, K. A., Kumar, A., LaRow, T. E., et al. (2017). Western North Pacific tropical cyclone model tracks in present and future climates. *Journal of Geophysical Research: Atmospheres*, *122*, 9721–9744. https:/doi.org/10.1002/2017JD027007

Tippett, M. K., Camargo, S. J., & Sobel, A. H. (2011). A Possion regression index for tropical cyclone genesis and the role of large-scale vorticity in genesis. *Journal of Climate*, *21*, 2335–2357.

Tolwinski-Ward, S. E. (2015). Uncertainty quantification for a climatology of the frequency and spatial distribution of North Atlantic tropical cyclone landfalls. *Journal of Advances in Modeling Earth Systems*, *7*, 305–319.

Yonekura, E., & Hall, T. M. (2011). A statistical model of tropical cyclone tracks in the western north Pacific with ENSO-dependent cyclogenesis. *Journal of Applied Meteorology and Climatology*, *50*(8), 1725–1739.

Yonekura, E., & Hall, T. M. (2014). ENSO effect on East Asian tropical cyclone landfall via changes in tracks and genesis in a statistical model. *Journal of Applied Meteorology and Climatology*, *53*, 406–420.

Zhao, H., Wu, L., & Zhou, W. (2009). Observational relationship of climatologic beta drift with large-scale environmental flows. *Geophysical Research Letters*, *36*, L18809. https:/doi.org/10.1029/2009GL040126