Understanding Uncertainties in Tropical Cyclone Rainfall Hazard Modeling Using Synthetic Storms

Dazhi Xi, Ning Lin

Department of Civil and Environmental Engineering, Princeton University

Corresponding Author: Dazhi Xi

Email: dxi@princeton.edu

Abstract

Tropical cyclone (TC) rainfall hazard assessment is subject to the bias in TC climatology estimation from climate simulations or synthetic downscaling. In this study, we investigate the uncertainty in TC rainfall hazard assessment induced by this bias using both rain gauge and radar observations and synthetic-storm-model-coupled TC rainfall simulations. We identify the storm's maximum intensity, impact duration, and minimal distance to the site to be the three most important storm parameters for TC rainfall hazard, and the relationship between the important storm parameters and TC rainfall can be well captured by a physics-based TC rainfall model. The uncertainty in the synthetic rainfall hazard induced by the bias in TC climatology can be largely explained by the bias in the important storm parameters simulated by the synthetic storm model. Correcting the distribution of the most biased parameter may significantly improve rainfall hazard estimation. Bias correction based on the joint distribution of the important parameters may render more accurate rainfall hazard estimations; however, the general technical difficulties in resampling from high dimensional joint probability distributions prevent more accurate estimations in some cases. The results of the study also support future investigation of the impact of climate change on TC rainfall hazards through the lens of future changes in the identified important storm parameters.

1. Introduction

Extreme rainfall is one of the tropical cyclones (TCs) hazards that have significant impacts on coastal (Gori et al. 2022) and inland (Aryal et al. 2018) areas. Recent events of TC extreme rainfall include Hurricane Harvey in 2017, Hurricane Florence in 2018, and Hurricane Ida in 2021, which induced economical losses of $125 billion, $24.23 billion, and $50.1 billion, respectively. Studies have also warned that TC rainfall hazard may greatly increase in the future (Emanuel 2017, Hall and Kossin 2019), suggesting that better understanding of TC rainfall and the associated hazards is urgently needed.

One way to study TC rainfall is from the physics perspective, namely examining the different physical mechanisms that contribute to TC rainfall. The understanding of TC rainfall mechanisms has advanced over the past few decades. The rainfall in eyewall regions and primary rainbands is dominated by convective rainfall, while the rainfall in outer rainbands is mostly caused by stratiform rainfall, which is weaker but covers larger areas (Houze 2010). Studies have shown that the frictional effect (Shapiro 1983), topographic forcing (Cheung et al. 2008, Yang et al. 2011), vertical wind shear (Braun and Wu 2006, Willoughby et al. 1984), and vortex stretching related to TC intensity evolution (Lu et al. 2018) are four important mechanisms for TC rainfall generation.

Another way to study TC rainfall is from the hazard perspective, that is, to understand the characteristics of TCs that are likely to produce extreme rainfall. For example, one would expect stronger TCs to produce more rainfall than weaker TCs, and the locations that are closer to a TC's center to receive more rainfall than more distant locations (Rodgers and Adler 1981). Another example is that Hurricane Harvey stalled around Houston for several days, which is one reason Harvey caused extreme flooding in Houston (Hall and Kossin 2019). These examples point out that several TC-related parameters may influence the possibility of extreme TC rainfall hazard. In this study, we aim to better understand which TC-related parameters may control TC rainfall hazard.

Such an improved understanding of the relationships between the important TC parameters and rainfall hazard will be useful for assessing TC rainfall hazard. First, it will help us understand and reduce the bias in estimated TC rainfall hazard. One approach to simulating TC rainfall hazard involves coupling a synthetic storm model and a physics-based TC rainfall model (Emanuel et al. 2017), neither of which is free of bias. Previous studies focused on the uncertainties of the TC rainfall model (Lu et al. 2018, Xi et al. 2020, Feldmann et al. 2019). However, the synthetic storm models are known to have bias in storm parameters including locations, translation speed, and intensity (Emanuel et al. 2008, Lee et al. 2017, Jing and Lin 2020), which will induce bias in the simulated rainfall hazard. Thus, it is worth also exploring whether the bias of simulated rainfall hazard could be significantly reduced by bias-correcting the probability distributions of storm parameters that are important for rainfall. This understanding can also support rainfall hazard assessment using parametric rainfall models (e.g., Tuleya et al. 2007, Villarini et al. 2021). Second, knowledge of important parameters for TC rainfall hazard will be useful for understanding the change of TC rainfall hazard due to climate change. While a number of studies have discussed the macroscopic changes of TC rainfall under climate change, such as changes of averaged TC rain rate and rainfall area (e.g., Knuston et al. 2010, Liu et al. 2018), very few studies have quantified the effects of climate change on TC rainfall hazards at landfall (e.g., Emanuel 2017). Knowing the connection between TC parameters and rainfall hazard improves understanding of the response of TC rainfall hazards to the changes in TC climatology characteristics, such as the increased intensity (Knuston et al. 2010, Emanuel 2021), reduced translation speed (Kossin 2018, Hall and Kossin, 2019), and poleward migration of TC tracks (Yin 2005, Tamarin-Brodsky and Kaspi 2017), and helps understanding climate-model-based projections on TC rainfall hazard (Wright et al. 2015).

In this study, we apply both observations and a TC rainfall model (TCR) (Emanuel 2017, Zhu et al. 2013, Lu et al. 2018) to study the important TC parameters for rainfall. TCR is a physics-based TC rainfall model that simulates TC precipitation by calculating vertical vapor transportation within TCs caused by the main rainfall generation mechanisms, including frictional effect, vortex stretching, baroclinic effect, and topographic effect. The model has been proven to be capable of reproducing climatology features of TC total rainfall (Feldmann et al. 2019, Xi et al. 2020) although it is less capable of reproducing climatology features of rainfall time series (Xi et al. 2020). A modeling-based investigation of TC rainfall is first performed as the observed TC rainfall is influenced by non-TC factors (such as surface roughness, atmospheric water vapor content, etc.), which are beyond the scope of this study. Also, TCR has been used to perform TC rainfall hazard analysis (Feldmann et al. 2019); thus, we can directly use the diagnosed important parameters to understand bias in TC rainfall hazard estimated by TCR. After selecting the key parameters based on the simulation, we use observations to examine the relationship between the selected parameters and the rainfall hazard and to validate the representation of these relationships in the TCR modeling.

Next, we couple TCR with two synthetic storm models to assess TC rainfall hazard along the U.S. East and Gulf Coasts and to investigate the bias in the estimated rainfall hazard as connected to the bias in the important parameters. The synthetic storm models we use include the statistical-deterministic model of Emanuel et al. (2008), which is based on the Coupled Hurricane Intensity Prediction System (hereafter CHIPS), and the Princeton environment-dependent Probabilistic tropical Cyclone (PepC) model (Jing and Lin 2020). To deal with the bias in TC rainfall hazard estimation caused by the bias in the important parameters in the synthetic storm models, we investigate whether the bias of the simulated rainfall hazard can be significantly reduced by bias-correcting the probability distribution of the important parameters simulated in these synthetic storm models. Though TCR itself also have intrinsic biases (Feldmann et al. 2019, Xi et al. 2020), in this study, we focus on the bias in storm simulations and how it propagates into landfalling TC rainfall hazard estimation.

The structure of this paper is the following. Section 2 revisits the synthetic storm models and TCR, summarizes the data used in this study, and introduces the bias-correction method we propose. Section 3 shows the result in storm parameter identification, TCR simulated rainfall hazard evaluation, synthetic storm model and rainfall comparison, and storm and rainfall correction. In the discussion (Section 4), we explore the potential applications of this study. We also compare the dependence of the bias in simulated rainfall hazard on storm parameters and the dependence of the bias in simulated wind hazard on storm parameters, given that the simulated wind drives TCR simulation of the rainfall. The major part of this study focuses on the total rainfall caused by TCs at a point of interest. However, the hourly maximum rain rate may also be of interest; Section 4 also briefly discusses the important parameters for the maximum rain rate. Section 5 summarizes the conclusions of this study.

2. Models, simulations, and analysis method

In this study, we first identified the important TC parameters for TC rainfall hazard by analyzing both simulated and observed TC rainfall and investigating how these parameters control the statistical properties of TC rainfall hazard. Then we coupled the synthetic storm models with TCR to understand how the spatial distributions of the bias in important parameters influence the bias in the estimated rainfall hazard. We further selected 10 rain gauge stations to test whether a bias correction on the important parameters based on resampling can improve the estimation of TC rainfall hazard.

2.1 Data

The historical record of TCs from 1979-2018 was obtained from the International Best Track Archive for Climate Stewardship (IBTrACS). The best track data contains 6-hourly information on TC intensity and location. We used the data to investigate the important parameters for TC rainfall, to evaluate synthetic storm models, and to create the observed probability distributions of TC parameters.

Coupling the historical TCs with the TCR model requires several environmental parameters, including the 900 hPa specific humidity, deep layer (200 hPa – 850 hPa) vertical wind shear, and surface drag coefficient. Following Xi et al. (2020), we obtained these parameters from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis data and linearly interpolated them from 6 hourly to 1 hourly for TC rainfall simulation.

We investigated how TC important parameters modulate TC rainfall hazard using the TCR model; the results were evaluated by radar rainfall observations that have a large spatial coverage. We used the Stage IV quantitative precipitation estimation (Lin and Mitchell 2005) from 2002-2018. Stage IV is a radar rainfall observation product that provides rainfall observation in 4-km horizontal resolution and 1-h temporal resolution. The fine resolution makes Stage IV observations preferable for analyzing landfalling TC rainfall. Stage IV has been shown to match well with observations for several historical TC events (Villarini et al. 2011, Luitel et al. 2018).

To evaluate the probability distribution of TC rainfall hazard simulated from IBTrACS-TCR and synthetic storm-coupled TCR and perform bias correction analysis, we obtained the historical record of TC rainfall at 10 coastal and inland locations across the United States from 1979-2018. The 10 rain gauge observations, obtained from National Centers for Environmental Information (NCEI), are in daily time resolution and are interpolated to hourly time steps. We used gauge observations instead of Stage IV radar observations because accurately estimating the probability distributions of TC rainfall requires long observation histories while Stage IV observation started in 2002.

2.3 Synthetic storm models

To generate synthetic storms for TC rainfall hazard assessment, we used two synthetic storm models: CHIPS and PepC. First, we used PepC-generated U.S. landfalling storms to analyze the storm parameters that control TC rainfall. We use only one synthetic model to generate data for this analysis as the relationship between storm parameters and simulated rainfall that we seek is determined solely by the TCR model. Then, CHIPS and PepC models were used to generate two large sets of US landfalling storms to understand and compare their performances when coupled with TCR. The detailed formulation of CHIPS and PepC can be found in previous research (Emanuel et al. 2008, Jing and Lin 2020); here we briefly summarize the models.

CHIPS consists of three parts: a genesis model, a TC track model, and an intensity model. The genesis of TC is modeled by a random seeding process. After generation, the TC is moved by the beta drift and steering wind. The steering wind is modeled by the linear combination of synthetic winds at 850 hPa and 200 hPa levels. The intensity of the storm is modeled with an air-sea coupled TC dynamic model. This study used 5018 US landfalling TCs under the historical climate (1980-2005) generated from CHIPS (Marsooli et al. 2019).

PepC follows a structure similar to that of CHIPS but is a statistical model. The genesis component is environment-dependent, and the number of TC seeds in each grid is predicted by a cluster-based Poisson regression model. Instead of using a beta-advection model for TC motion, PepC employs an analog-wind model that involves the steering wind as well as observed track patterns. The intensity evolution is modeled as a 3-state environment-dependent hidden Markov chain. This study uses 3013 U.S. landfalling synthetic storms generated by this model (Jing and Lin 2020).

2.4 A physics-based TC rainfall model (TCR)

Lu et al. (2018) provided a detailed description of TCR; here we briefly summarize the model. TCR generates rainfall by computing the vertical vapor flux through the top of the boundary layer. The vertical vapor flux is computed as the product of the vertical velocity across the top of the boundary layer and specific humidity. The model computes the vertical velocity at the top of the TC boundary layer as the sum of the vertical velocities generated by five mechanisms: topographic forcing, frictional effect, vortex stretching, baroclinic effect, and radiative cooling. The model requires a wind profile as input; we use the wind profile proposed by Chavas et al. (2015) to drive TCR, as Xi et al. (2020) found that this wind profile works better than three other wind profiles. The horizontal resolution of the model was set to be 0.05°×0.05°, following previous research (Lu et al. 2018, Xi et al. 2020). The drag coefficient in the model is a function of surface roughness, independent of storm features. The horizontal resolution of drag coefficient parameter was set to be 0.25°x0.25°, following Feldmann et al. (2019).

2.5 Simulations and analysis method

In the main results, we focused on ETR as the rainfall hazard. ETR is defined as the total rainfall precipitated by the TC to a point of interest (POI) while the TC is within 300 km to the location, following Feldmann et al. (2019). (A distance threshold is used here to avoid the inclusion of rainfall from other weather systems.) Another reason we focused on ETR is that recently Hurricanes Harvey and Florence produced dramatic total rainfall that caused destructive inundation and flooding in Houston and North Carolina, respectively. To determine the important TC-related parameters for ETR, we first examined the TCR simulation and then evaluate the findings with observations. TCR proved to be reliable in previous studies in reproducing the statistics of ETR (Feldmann et al. 2019, Xi et al. 2020), and TCR provides us with both a larger amount of data than the Stage IV observations and the flexibility of excluding the influence of non-TC parameters (detailed later). We coupled the 3013 PepC synthetic storms with TCR to identify storm parameters that are important for ETR and the favorable conditions (in terms of the selected parameters) for extreme rainfall. The drag coefficient and specific humidity were set to be constant, 0.002 and 0.012, respectively, to represent conditions on land. We set the drag coefficient and specific humidity as constant in this experiment to help us focus on the parameters that relate only to storm track and intensity. We divided the continental United States into 0.5°x0.5° grid points, and each point is a POI in this study. We interpolated the simulated ETR for each storm into these POIs and computed the values of the tentative important storm parameters for each POI in each storm event. The 0.5°x0.5° analysis grid is sufficient for resolving the rainfall variability caused by the spatial variability of TC characteristics, and the rainfall variability caused by spatial difference of surface future is largely removed in this experiment by setting a constant drag coefficient. We then performed least absolute shrinkage and selection operator (LASSO) regression (Tibshirani 1996) with the tentative parameters and varied the penalty coefficient to reduce the number of important parameters. After identifying subset of important parameters, we examined how the probability distribution of the important parameters influenced the chance of extreme rainfall events.

We simulated synthetic storms and assessed rainfall hazard from the two synthetic storm models: CHIPS and PepCs. To understand and reduce the bias of estimated rainfall hazard induced by the bias in storm climatology, we propose a method for bias-correcting the rainfall hazard through bias-correcting the probability distribution of the important TC parameters, as also a way to further investigate the relationship between the important TC parameters and rainfall hazard. We employ the importance sampling method. That is, while the synthetic storms originally have an equal probability of occurrence in the simulated dataset, we adjust the probability of occurrence for each storm according to joint probability distributions of selected parameters in the simulation relative to the observations.

Specifically, the probability of exceedance of a certain rainfall level η_r is

$$\mathbb{P}(\eta > \eta_r) = \int \mathbf{1}_{\eta(\boldsymbol{x})>\eta_r}\, p(\boldsymbol{x}) d\boldsymbol{x} = \int \mathbf{1}_{\eta(\boldsymbol{x})>\eta_r} \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} q(\boldsymbol{x}) d\boldsymbol{x} = \mathbb{E}_q\left(\frac{\mathbf{1}_{\eta(\boldsymbol{x})>\eta_r} p(\boldsymbol{x})}{q(\boldsymbol{x})}\right) \qquad (1)$$

where $\mathbf{1}_{\eta(\boldsymbol{x})>\eta_r}$ is the indicator function for whether a storm with parameter $\boldsymbol{x}$ will generate rainfall larger than η_r. The nominal distribution (or the target distribution), $p(\boldsymbol{x})$, is fitted by the observed joint distributions of the important parameters. The importance distribution, $q(x)$, is the distribution which we can draw samples from. In this study, the importance distribution is the joint distributions of the important parameters from the synthetic storm simulation. $\mathbb{E}_q(\boldsymbol{x})$ denotes the expectation of x over the sample space of the synthetic storm parameters.

The expectation in (1) can be estimated statistically, as in (2):

$$\mathbb{E}_q\left(\frac{\mathbf{1}_{\eta(\boldsymbol{x})>\eta_r} p(\boldsymbol{x})}{q(\boldsymbol{x})}\right) \approx \frac{1}{n}\sum_{i=1}^{n} \frac{1_{\hat{\eta}(\boldsymbol{x}_i)>\eta_r} \hat{p}(\boldsymbol{x}_i)}{\hat{q}(\boldsymbol{x}_i)} = \frac{1}{n}\sum_{i=1}^{n} 1_{\hat{\eta}(\boldsymbol{x}_i)>\eta_r} w(\boldsymbol{x}_i)\,, \boldsymbol{x}_i \sim \hat{q} \qquad (2)$$

where n is the number of simulated storms that influence the POI. $\boldsymbol{x}_i$ stands for the vector of parameters for the i-th storm in the synthetic dataset. $\hat{p}(\boldsymbol{x}_i)$ is the estimated joint probability density function for the important parameters evaluated with the observations; $\hat{q}(\boldsymbol{x}_i)$ is the same but evaluated with the synthetic storm dataset. $\hat{\eta}(\boldsymbol{x}_i)$ represents the TCR simulated rainfall level, as an approximation for the real rainfall level, for given storm parameters. In practice, one can compute the rainfall exceedance probability using equation (2) or, equivalently, assign each storm a weight $w(\boldsymbol{x}_i) = \frac{\hat{p}(\boldsymbol{x}_i)}{\hat{q}(\boldsymbol{x}_i)}$ and resample the storm events to generate a new storm dataset with equal occurrence probabilities to calculate the rainfall exceedance probability directly. In theory, the joint probability distribution of the important parameters in the new storm dataset should match the observations, but sampling errors exist, especially if $\boldsymbol{x}_i$ has high dimensions. In this study, we used the resampling method to calculate the rainfall exceedance probability and check the probability distribution of the important parameters after resampling.

The above formulations show that if the selected parameters control the estimated rainfall probability distribution, the rainfall hazard estimation would be largely influenced by the adjustment of probability distribution of the parameters. Further, in theory, the estimated rainfall hazard will have no bias if three conditions comply: the TCR has no bias, the selected parameters fully control TC rainfall estimation, and in practice the resampling process matches the probability distributions of the parameters exactly with observations. In practice, however, the important sampling with irregular probability distributions, especially with high dimensional probability distributions, can be challenging. Nevertheless, as a first attempt, we explore the potential of such a statistical method in bias-correcting rainfall hazard estimation. To simplify the resampling process, we assumed the correlations of the important parameters can be modeled using the Gaussian copula. We used the Gaussian copula to model the joint feature for two reasons. First, the high-dimensional Gaussian copula is generally implemented in statistical packages and can be easily fitted. Second, we tested several Vine copulas, and they have performances similar to the simpler Gaussian copula in modeling the joint probability distribution of important parameters. To avoid the uncertainty introduced by selecting different forms of copulas for different locations, we thus use the Gaussian copula for each POI in the United States.

To understand the spatial distribution of rainfall hazard and how the important parameters affect the bias, we examined the differences between CHIPS-TCR simulation and PepC-TCR simulation based on the differences of the spatial distributions of the important parameters. We also examined the rainfall return levels and their spatial distribution corresponding to 10-

year, 50-year, and 100-year return periods in coastal states from CHIPS-TCR and PepC-TCR simulations. To compute the return levels L, we used Equation (3)

$$L = F^{-1}\left(1 + \frac{\ln\left(1 - \frac{1}{T_L}\right)}{\lambda}\right) \quad (3)$$

where T_L is the return period desired, λ is the storm frequency for the POI, F is the cumulative density function of the simulated ETR, and storms are assumed to arrive as a Poisson process (Lin et al. 2012). To focus on the differences of simulated return levels caused by the differences in TC characteristic parameters, we excluded the differences in storm frequency in the two synthetic storm models by using the observed historical frequency.

3. Results

3.1 Storm Parameter Selection and Favorable Conditions for Extreme ETR

We selected important parameters for TC rainfall hazard from the following tentative parameters evaluated when the storm is within 300 km of the POI: mean intensity of the storm; maximum intensity of the storm; mean distance of the storm to POI; minimal distance of the storm to POI; mean translation speed of the storm; heading of the storm when it is nearest to the POI; duration of the storm; and mean radius of the maximum wind. The 8 parameters described here are fed into a LASSO regression between the 8 parameters and the simulated ETR for the POI. Data from all POIs in the continental United States are grouped together to perform the LASSO regression. After gradually increasing the regularization coefficient, we reduced the number of responding parameters to 3. The selected parameters are maximum intensity, duration, and minimal distance within 300 km. As shown in Figure 1, the highest correlation coefficient of the LASSO-predicted ETR and TCR-simulated ETR when using only the 3 parameters is 0.6618, compared to the correlation coefficient around

0.6686 when the regression uses all 8 of the tentative parameters. The result shows that these three parameters are particularly important for the ETR in the TCR simulation. The regression coefficients for the maximum intensity and duration are positive while the coefficient for minimal distance is negative; as expected, intense storms with long impact times will cause heavy precipitation if they move close to the POI.

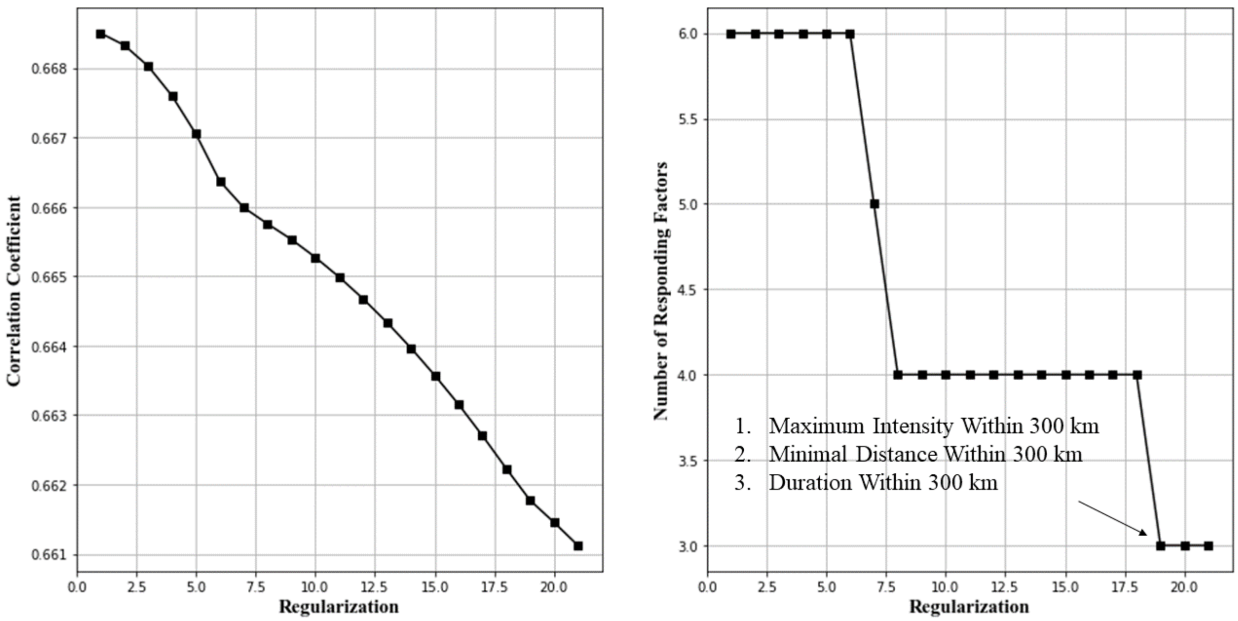


Figure 1. Lasso parameter selection process for important storm parameters for ETR. a). Correlation coefficient between Lasso-predicted ETR and TCR-simulated ETR against the regularization coefficient; b). Number of responding factors against the regularization coefficient. The regularization stands for the weight that modulates the importance of the L2 regularization term in Lasso. The increase of the regularization coefficient reduces the number of variables selected in the simulation.

The parameter selection is relatively robust. The selection of 300 km as a threshold (for both ETR and storm parameters) was based on previous research on TC rainfall hazard simulations (Feldmann et al. 2019) although other thresholds (e.g., 500 km, 600 km) were also used in previous studies (Xi et al. 2020). We briefly test the results using a 500-km threshold. We find the duration and minimal distance are still the important parameters for ETR; however, the other important parameter becomes the mean intensity, rather than the maximum intensity, when the TC is within 500 km of the POI. For the 300-km radius, the corresponding duration of storm is relatively short, so the rain rate when the storm is most intense is relatively important to ETR. However, for the 500-km radius, the corresponding duration of the storm is relatively long, so the rain rate when the storm is most intense is less important. Also, to test whether the selection of parameters is location-dependent, we also performed the LASSO regression using data for individual POIs. It is found that the duration and minimal distance are always selected. For some locations the maximum intensity is selected but in other locations the mean intensity is selected. However, it is noted that for most cases, larger maximum intensity corresponds to larger mean intensity ($R = 0.87$). In each case, storm heading, translation speed, and radius of maximum wind are not selected. The asymmetric feature of TC rainfall caused by storm translation could be important for the rainfall hazard produced by single cases, but such asymmetry was smoothed out when we sampled a large number of TCs to estimate TC rainfall hazard, as TCs can approach from any direction to a POI. The mean radius of maximum wind is correlated with storm intensity, and its impact on the ETR for a specific POI may also depend on distance.

Applications of identifying the important parameters for TC rainfall include investigating the criteria for a storm to produce extreme rainfall and evaluating whether a rainfall model such as TCR can generate extreme rainfall under the same conditions that favor extreme rainfall in reality. To explore the favorable conditions (regarding the 3 selected parameters) to generate extreme ETR, we examined the joint and marginal probability distributions of the 3 selected parameters under two conditions, extreme rainfall events and ordinary rainfall events. We defined extreme rainfall events as the events whose ETR exceed 99-th percentile of all the observed or simulated ETR. Both observations and IBTrACS-TCR simulation (rainfall simulation driven by observed storm parameters) show that ordinary rainfall events (<99-th percentile) usually happen when maximum intensity is less than 30 m/s, as shown in Figure 2a & b. Both simulation and observations show that short duration favors a weak rainfall event, and the duration for weak rainfall is usually less than 50 hours. For extreme rainfall events (Figure 2c & d), we found that the intensity and duration in both observations and simulation extends to larger values, showing that intense TCs and/or the TCs lingering around the POI induce extreme rainfall. For extreme rainfall events, the minimal distance in both simulation and observations concentrates at small values while in ordinary events the minimal distance is uniformly distributed. This result implies that shorter minimal distance favors extreme rainfall events. For the joint distributions of pairs of parameters, we found that extreme rainfall events in both the simulation and observations show the following shared features: (1) for intensity and distance, when intensity is low, the distance is also short so that extreme rainfall can be produced by relatively weak TCs; (2) for intensity and duration, when intensity is small, there are long duration cases to favor extreme rainfall generation while when duration is short, intense TCs generate extreme rainfall; and (3) the joint probability distribution of distance and duration concentrates in short distances and longer durations. The simulation shares similar features with observations in the abovementioned characteristics of marginal and joint probability distribution, implying that the simplified physics-based TC rainfall model generates extreme rainfall events provided the conditions that favor extreme rainfall events in the observations.

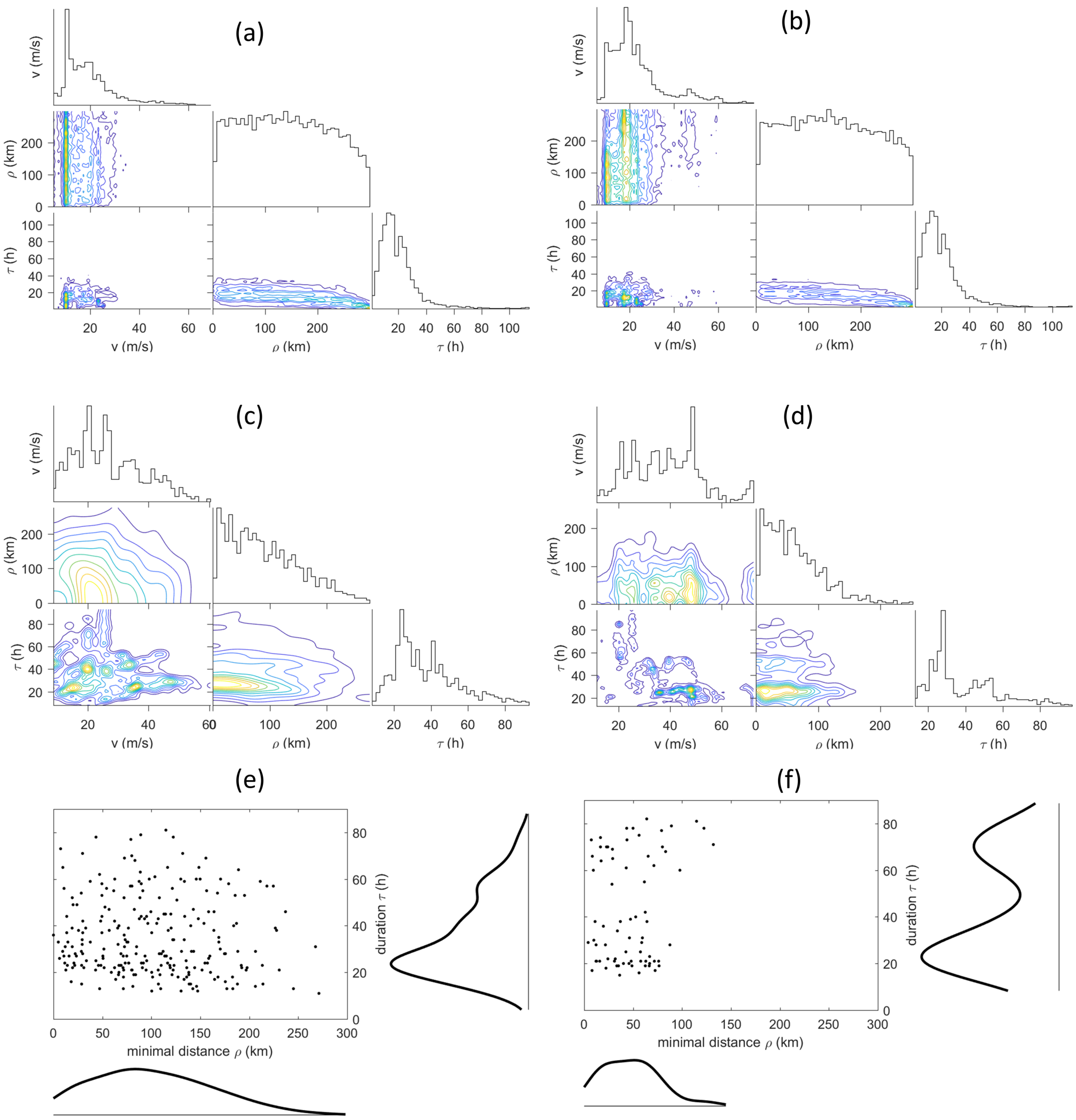


Figure 2. Probability distributions of the important parameters for ETR in different rainfall groups. (a) (b). Marginal and joint distribution of the three important parameters for ordinary rainfall event (<99% percentile) for Stage IV observations (<196.55 mm) and TCR simulation (<152.2 mm), respectively; (c) (d). Same as (a) (b), but for extreme rainfall (>=99% percentile) events for Stage IV observations (>= 196.55 mm) and TCR simulation (>=152.2 mm). (e) (f) marginal and joint distribution of minimal distance and duration conditioned on ETR>=100 mm and maximum storm intensity lower than 17.5 m/s (tropical storm intensity) for Stage IV observations and TCR simulation, respectively.

Previous research found that on the storm level, TCR may not reproduce the spatial distribution of the TC rainfall for specific storms (e.g., weak storms or those affected by other synoptic systems) as in the observations (Xi et al. 2020); however, the TCR is able to reproduce the correct averaged rainfall climatology, e.g., spatial distribution of annual average TC rainfall (Xi et al. 2020) and TC rainfall probability distribution (Feldmann et al. 2019). The findings in this study provide another perspective from which to understand the performance of TCR. The ETR of TC can be viewed as a dependent random process that is largely controlled by important parameters (maximum intensity, duration, and minimal distance) but that can be influenced by other parameters and larger-scale synoptic systems (Xi et al. 2020), which can be viewed as random forcing. At the storm-level, the random forcing can be an important factor to determine ETR, but for a large sample of storms, the probability distribution of ETR may be controlled mainly by the probability distribution of the important parameters. The finding that the TCR responds to the important parameters in the same way as the observed rainfall statistically, especially for the extreme events, indicates that TCR is a capable model for TC rainfall hazard assessment.

We also found that weak TCs can sometimes generate significant amounts of rainfall in the observations while TCR simulations have a very strict condition (minimal distance < 100 km) for producing heavy rainfall by weak TCs. Here we compare the joint and marginal probability distributions of minimal distance and duration of TCs below tropical storm intensity (<17.5m/s) that produce ETR larger than 100 mm. We found that in the observations, these events are usually favored in cases with short minimal distance (Figure 2e), but it is possible for a storm far from the POI (up to 270 km) to produce heavy rain. In the IBTrACS-TCR simulation, though, the cases in which weak TCs generate heavy rainfall can happen only when the distance is smaller than 100 km. Thus, the weak-TC-heavy-rain events are more likely to happen in reality than in the TCR simulation. One possible explanation is that the structure of weak TCs may not be compact and symmetric, which violates the assumption of TCR. The other possible reason is that the heavy rainfall associated with weak TCs is likely to be linked with other synoptic systems that are not taken into consideration in TCR or related to extratropical transitions with baroclinic effects not well captured by TCR. We also note that rainfall produced in TCR can be regarded as convective rainfall while rainfall in a distant rainband (may or may not be within 300 km from the TC center) is stratiform. The above analysis shows that though TCR reproduces extreme rainfall when conditions are favorable for TCs to generate extreme rainfall, TCR misses the unusual cases when TCs are less intense but still produce extreme rainfall.

3.2. Spatial Distribution of Simulated Important Storm Parameters in USA and Their Relationship with Simulated Rainfall Hazard

Section 3.1 identifies three important TC-related parameters that significantly influence TC rainfall hazard. Here we explore how the bias and uncertainty of these parameters influence the simulated TC rainfall hazard by analyzing the rainfall hazard across the United States, estimated by two synthetic storm models (CHIPS and PepC) coupled with TCR. First, we present the spatial distributions of the mean of the three important parameters from both observations and synthetic storm simulations (Figure 3). The mean of maximum intensity from CHIPS and PepC matches well with observations of the Texas Coast and the tip of Florida. However, both CHIPS and PepC underestimate the mean of maximum intensity on the East Coast, especially at high latitudes. PepC also underestimates storm intensity and overestimates minimal distance for locations inland. Both of these biases are related to the oversimplified TC decay model in PepC when storms are over landmasses. The overestimation of minimal distance results from the short on-land lifetime of TCs in the PepC model, resulting in most of the TCs not moving close to inland POIs. For the minimal distance, PepC performs well along the Southern Coast (from Texas to Alabama), but it shows overestimation from coastal Florida to the Northeast. CHIPS underestimates the minimal distance in west Florida and overestimates the minimal distance in other U.S. coastal areas, although the overestimation from CHIPS is less significant than the overestimation from PepC. PepC overestimates the duration of the storms in Florida and in high latitudes on the East Coast, while CHIPS underestimates the duration along the South Coast and East Coast.

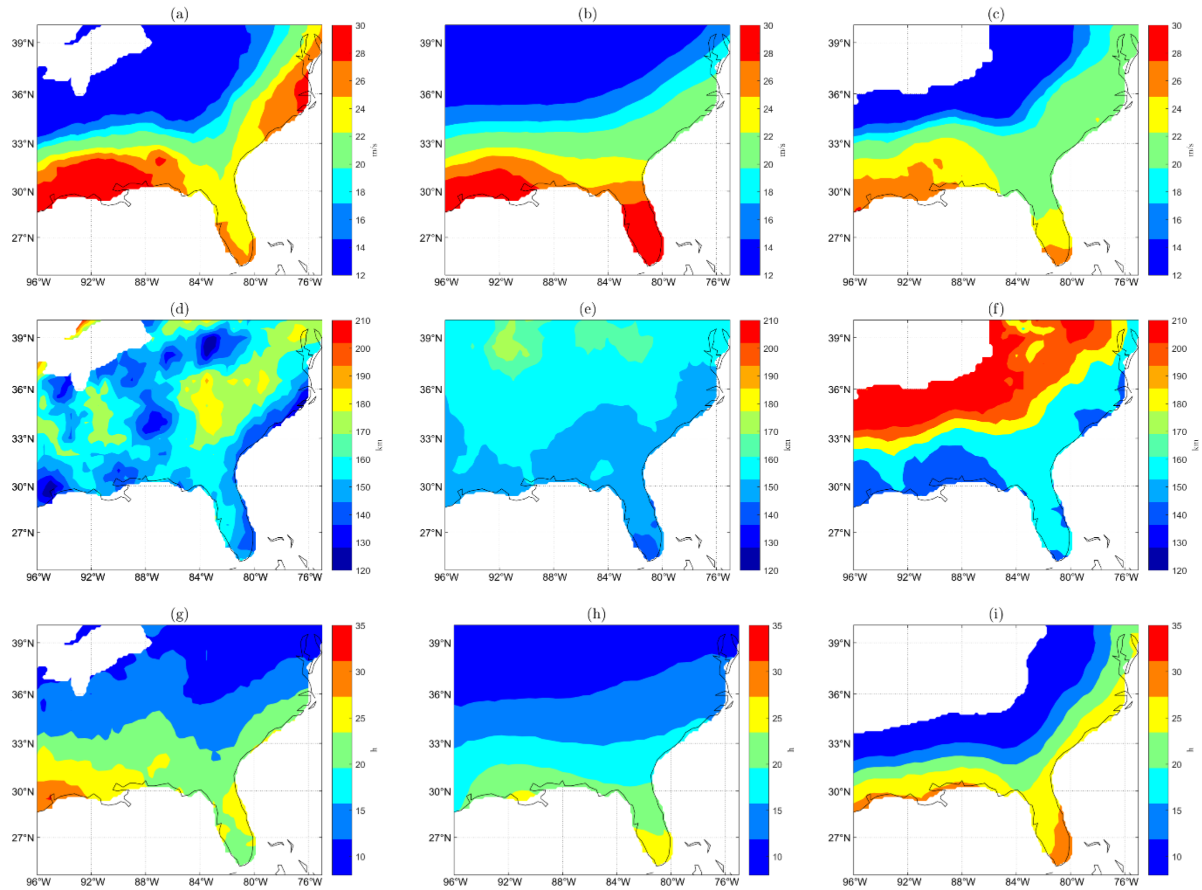


Figure3. Spatial distribution of the mean of the three parameters. (a) (b) (c). Mean of maximum intensity of IBTrACS, CHIPS, and PepC; (d) (e) (f). Mean of minimal distance; (g) (h) (i). Mean of duration.

The differences in the important parameters estimated by the two synthetic storm models can be used to understand the different performances of the synthetic storm models for rainfall hazard estimation. Here we show the return level estimated from CHIPS-TCR and PepC-TCR for return periods of 10-years, 50-years, and 100-years (Figure 4). One difference between the two coupled models is that PepC-TCR predicts higher rainfall hazard on the East Coast. The differences in minimal distance and maximum intensity in the East Coast in PepC and CHIPS are not significant (Fig. 3b-c, 3e-f); however, the duration of a storm simulated by PepC is much longer than the duration simulated by CHIPS (Figure 3h-i), which leads to higher rainfall hazard on the East Coast estimated by the PepC-TCR simulation. The two models also show differences in rainfall hazards in Coastal Texas, where PepC simulates higher rainfall hazard than CHIPS. In Coastal Texas, though the maximum intensity simulated by PepC is around 84% of the maximum intensity simulated by CHIPS, the minimal distance simulated by PepC is 13% lower than that in CHIPS, and the mean duration in PepC is around 50% longer than that in CHIPS in Coastal Texas. The higher rainfall hazard in Texas estimated by PepC is a result of longer duration and shorter minimal distance, despite the storm intensity in this dataset is slightly lower.

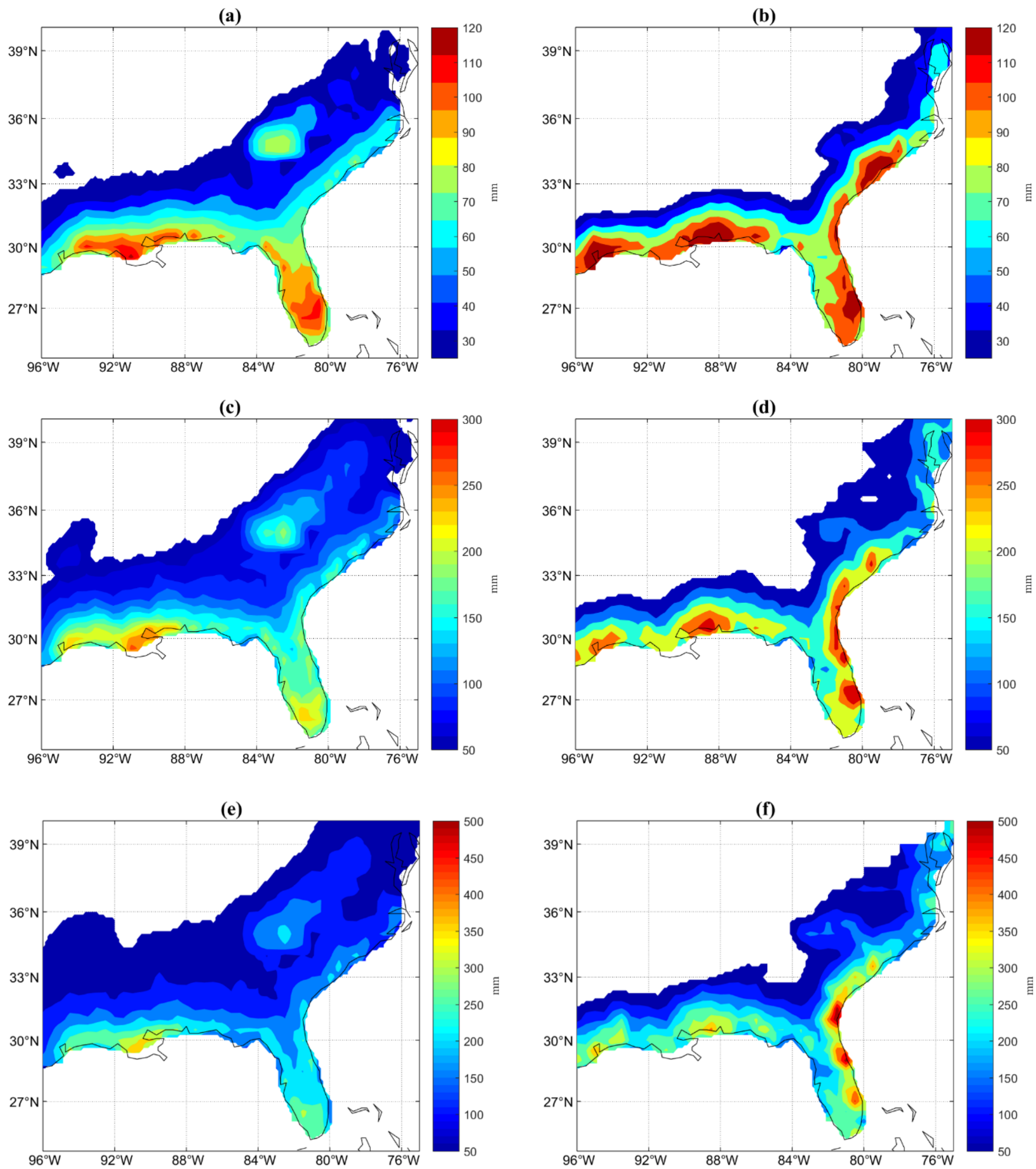


Figure 4. Return level of the simulated rainfall hazard. (a)(b). 10-year return period; (c)(d). 50-year return period; (e)(f). 100-year return period; (a)(c)(e). CHIPS-TCR; (b)(d)(f). PepC-TCR.

We have explored the spatial distribution of the mean of the parameters. Here we briefly discuss the correlations of these parameters (Figure 5). Though there is no physical reason for duration and minimal distance to be related to storm intensity, they may be statistically related. We found that for most coastal areas, the linear correlation between each pair of the three parameters is small. However, if the direction of motion and translation speed of storms near a POI have low variability across the storms, the minimal distance and duration should be negatively correlated. Thus, a significant negative correlation between the minimal distance and duration in Coastal Texas implies that the motion of storms in this area has low variability. In the next section, we model the joint probability distribution of the three important parameters by Gaussian copula, to partly capture the weak correlations between the three parameters.

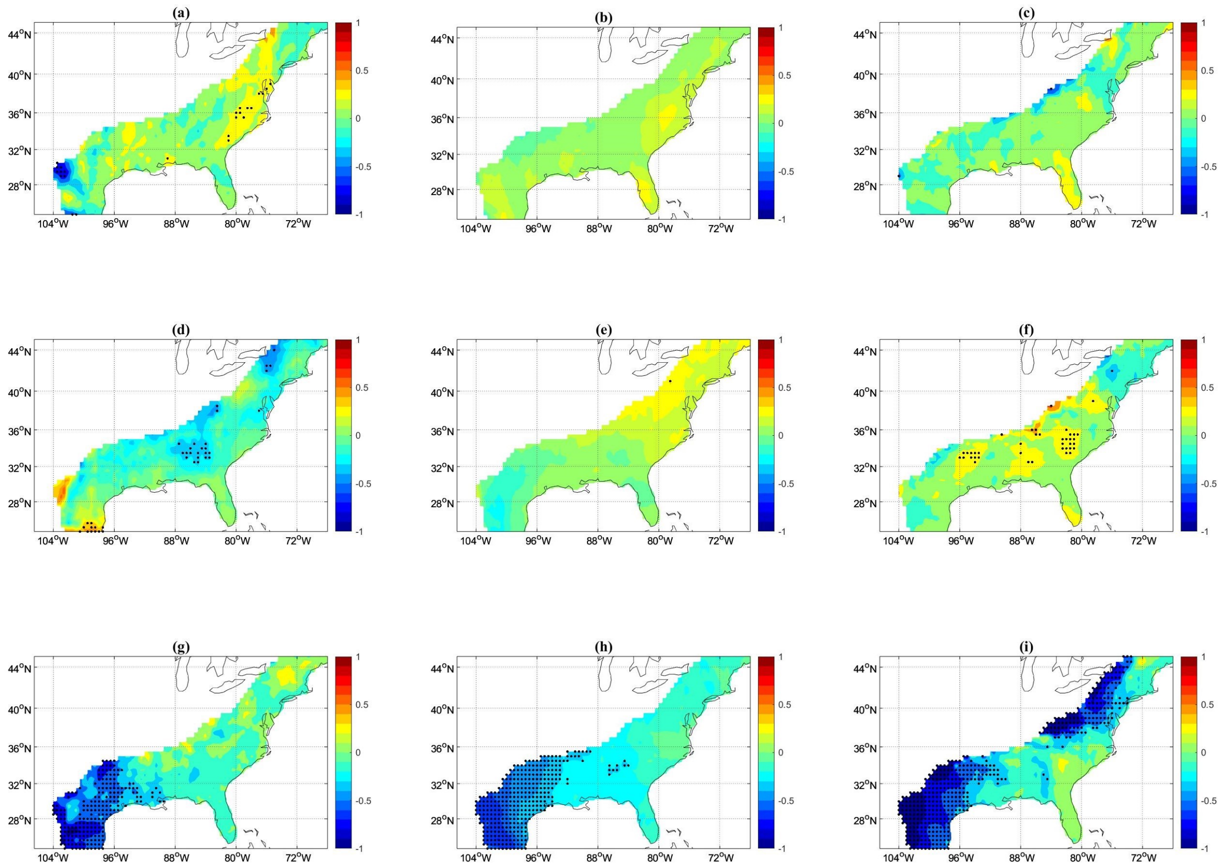


Figure 5. Correlation of the selected parameters in observations, CHIPS synthetic storms, and PepC synthetic storms. Figures a), b), c) Correlation between maximum intensity and minimal distance for IBTrACS, CHIPS, and PepC, respectively; d), e), f) Correlation between maximum intensity and duration for IBTrACS, CHIPS, and PepC, respectively; g), h), i) Correlation between duration and minimal distance for IBTrACS, CHIPS, and PepC, respectively. The locations that are marked by black dots are where the correlation is significant, and the absolute value of the correlation coefficient is larger than 0.3.

3.3. Probability Distribution of Important Parameters and Storm Probability Correction

The findings in Section 3.1 indicate a way to examine the bias of rainfall hazard probability distribution simulated by TCR coupled with synthetic storm models, that is, to examine the differences of the three important parameters in the synthetic storm simulation and observations. Section 3.2 shows the difference of the mean of important parameters in two synthetic storm models and the effects on rainfall hazards. Here we further discuss how features of the probability distribution of the parameters influence the probability distribution of the simulated rainfall. we selected 10 locations (shown in Figure 6) and compared the probability distribution of the ETR in the two model simulations at each location. The 10 locations are selected to cover the coastal locations from Gulf Coast to the Northeast, with Atlanta as a representative for inland regions. We did not select locations in regions with complex terrains (the Appalachian Mountains) or affected significantly by synoptic-system-influenced TC rainfall events (Texas), as the bias of simulated rainfall in these locations are largely induced by the bias in TCR (Xi et al. 2020).

For each location, for ETR exceedance probability distribution, we compare the rain gauge observations (black dot in Figure 7 upper panels) and IBTrACS-TCR simulations (purple dot in Figure 7 upper panels); a relatively small difference between the two confirms that the bias in TCR is small for the selected location. The bias that caused by the synthetic storm simulations from CHIPS (red solid line in Figure 7 upper panels) and PepC (blue solid line in Figure 7 upper panels) are shown by comparing their hazard curves with IBTrACS-TCR. Then we applied the storm probability correction method introduced in Section 2.3 to examine the influence of the bias in parameters on the estimated rainfall hazard and to investigate if bias-correcting the probability distribution of the important parameters can lead to significantly better estimation of the ETR distribution, i.e., becoming closer to the IBTrACS-TCR simulations. We present the bias correction based on marginal distributions (dashed red and blue lines in Figure 7 upper panels for CHIPS and PepC respectively) of the three individual parameters to show how a single parameter influences the estimated ETR exceedance rate and present the bias correction based on joint distributions of three parameters (dotted red and blue lines in Figure 7 upper panels for CHIPS and PepC respectively) to explore the possibility of using joint probability distribution to correct rainfall exceedance probability.

To connect the bias in the rainfall hazard to that in the storm parameters, we also compare the probability distribution of each individual parameter from historical observation (solid black line in Figure 7 lower panels), synthetic storm model simulations (solid red and blue lines in Figure 7 lower panels for CHIPS and PepC respectively), and the resampled results based on the probability distribution of the corresponding individual parameter (dashed red and blue lines in Figure 7 lower panels for CHIPS and PepC respectively) and the joint probability of the parameters (dotted red and blue lines in Figure 7 lower panels for CHIPS and PepC respectively).

Figure 6. Locations of rain gauge observations.

For most of the 10 selected sites, we found that the IBTrACS-TCR has no significant bias compared to gauge observations in the low rainfall region, where the estimation of the exceedance rate from observations has limited uncertainty. We acknowledge that for some other locations not considered here such as Texas and mountainous areas, IBTrACS-TCR has noticeable bias compared to rain gauge observations. However, in this study, we focus on the bias that caused by storm simulations. In the 10 selected sites, we found in Charleston, Jacksonville, Miami, New Orleans, and Tallahassee both CHIPS-TCR and PepC-TCR coupled TCR perform well compared to the IBTrACS-TCR simulation. Thus, for these locations, correcting the bias based on single important parameters shows limited effects on changing the probability distribution of ETR, especially at the low ETR regime, where the uncertainty of ETR from the rain gauge observations is small. However, applying the bias correction based on the joint parameters may worsen the rainfall hazard estimation for these sites, e.g., Charleston and New Orleans, with possible reasons discussed later.

We then examined how the bias of the parameters in the other five locations contributes to the bias of rainfall hazard simulated. PepC simulation in Atlanta significantly underestimates TC rainfall hazards. The underestimation is related to the overestimation of minimal distance and underestimation of duration at this inland location (as shown in the probability distributions), though the intensity is overestimated. As expected, bias correction based on storm intensity has no effect on correcting the bias of simulated rainfall, but the bias correction based on minimal distance and duration improves the models' performance. CHIPS simulated rainfall matches better with the rain gauge observations than PepC does for Atlanta with slight overestimation, but the better match is likely a combined effect of overestimation of intensity and underestimation of duration. Correcting based on the probability of maximum intensity will slightly improve the estimation of rainfall hazard, but correcting based on the probability of duration will worsen the estimation. This result shows that correcting a parameter that biases in the opposite direction of estimated rainfall bias will further worsen the estimation. Applying the bias correction based on joint probability distribution shows improvement on the rainfall hazard estimation for both PepC and CHIPS in Atlanta.

In Baltimore, PepC overestimates rainfall hazard because the simulated duration and intensity probability distributions shift to larger values than in the observations. Bias correction based on these two parameters improves the simulation. For CHIPS, model simulation matches well with observations, which is likely related to the combined effect of underestimation of duration, slight overestimation of mean intensity, and underestimation of minimal distance, as seen in the probability distribution. Bias correction based on each single parameter shows limited effects on the simulation of rainfall hazard. Correction based on the joint probability distribution can improve rainfall estimation over both CHIPS and PepC. In Mobile, both CHIPS and PepC show most bias in the probability of the duration of storm, and correction based on this parameter and based on the joint distribution improves both CHIPS and PepC simulations.

In Norfolk, PepC overestimates the rainfall hazard. The overestimation is mainly due to the significant overestimation of the duration of storms that impact this location. Bias correction based on duration for PepC improves the estimation of rainfall hazards. However, applying the bias correction based on the joint probability distribution does not improve the simulation results, especially for PepC. The unsatisfactory performance of the joint probability correction of PepC occurs probably because the resampled storm intensity does not match well with observations, and the unsatisfactory performance of joint probability correction of CHIPS occurs likely due to the mismatch between resampled duration and observations.

In Savannah, both PepC and CHIPS overestimate rainfall hazards, and the overestimation of PepC is more significant. CHIPS overestimates intensity and chances of short distance but slightly underestimates storm duration. Bias correction based on intensity and distance slightly improves the estimation of rainfall hazard from CHIPS. The bias of PepC is most significant in duration, and the bias correction based on duration improves the rainfall hazard estimation. Bias correction based on the joint parameters shows limited improvement for CHIPS, but it improves the rainfall estimation from PepC in Savannah. The analysis above shows that bias correction based on the most biased parameter can improve the estimation of rainfall hazards while using the joint probability cannot always improve the estimated TC rainfall hazard.

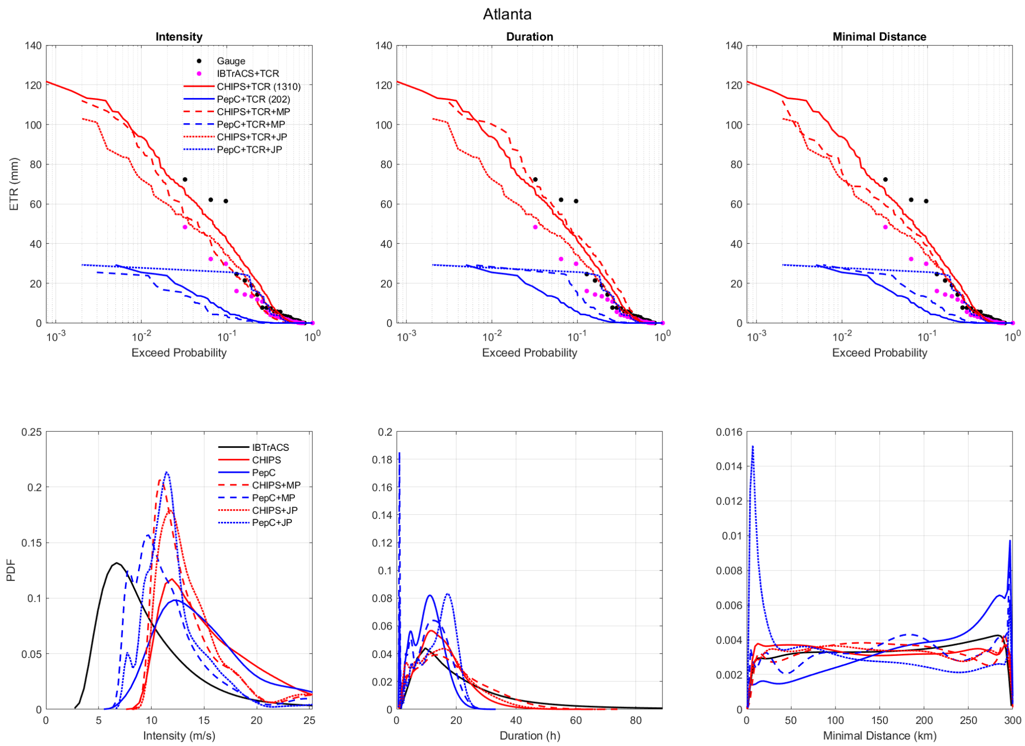
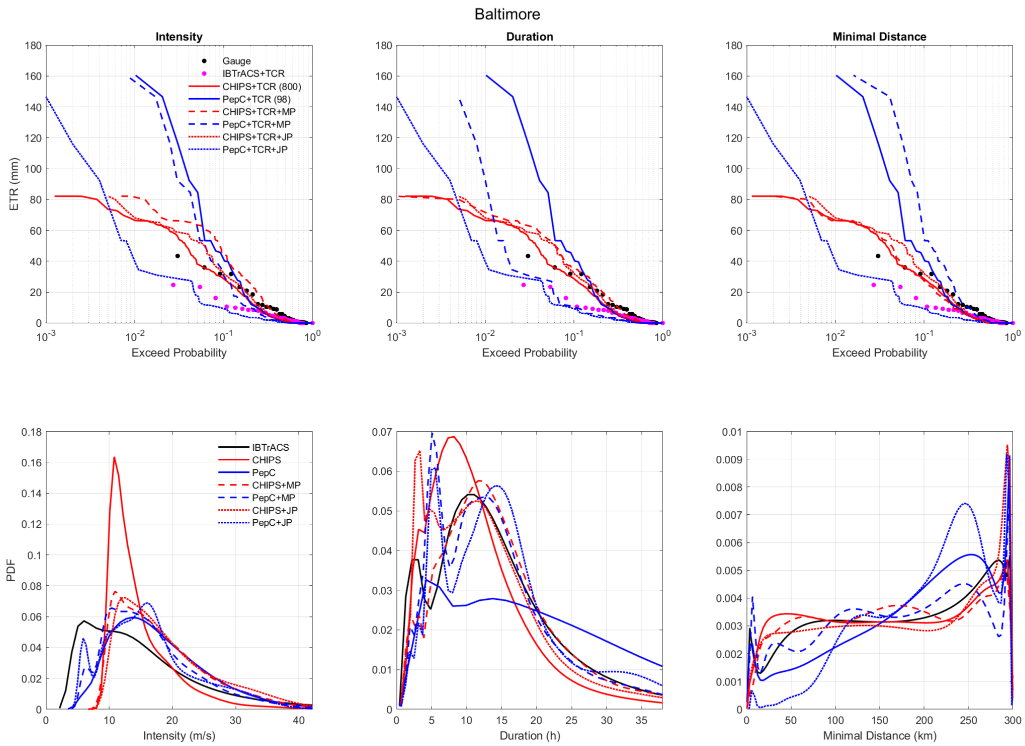
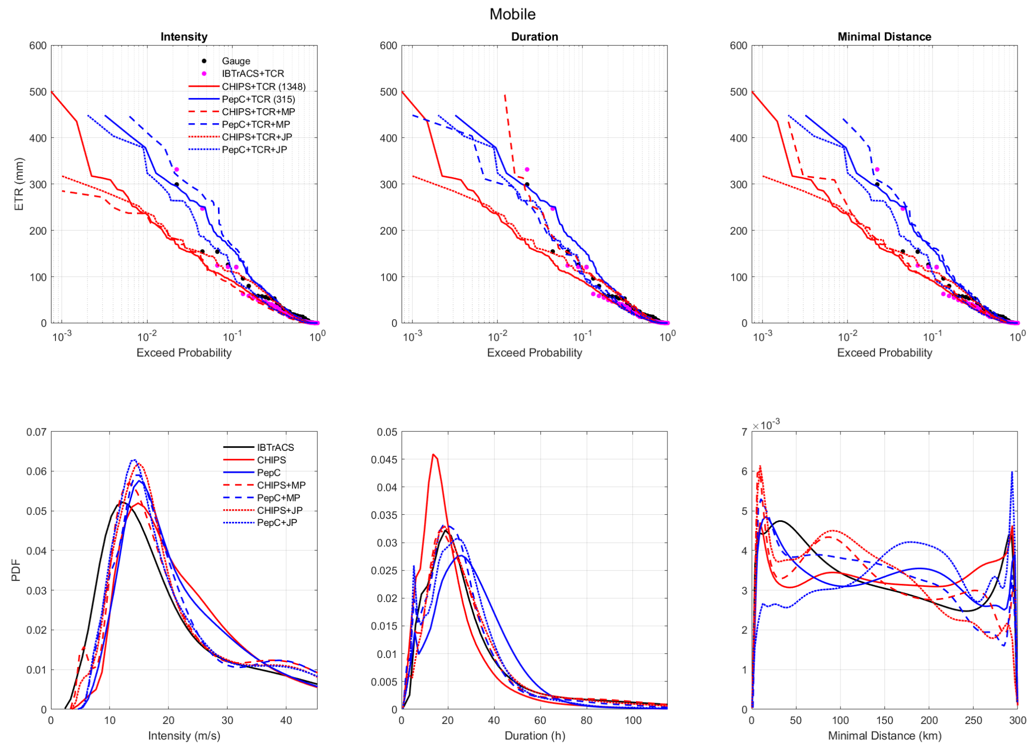


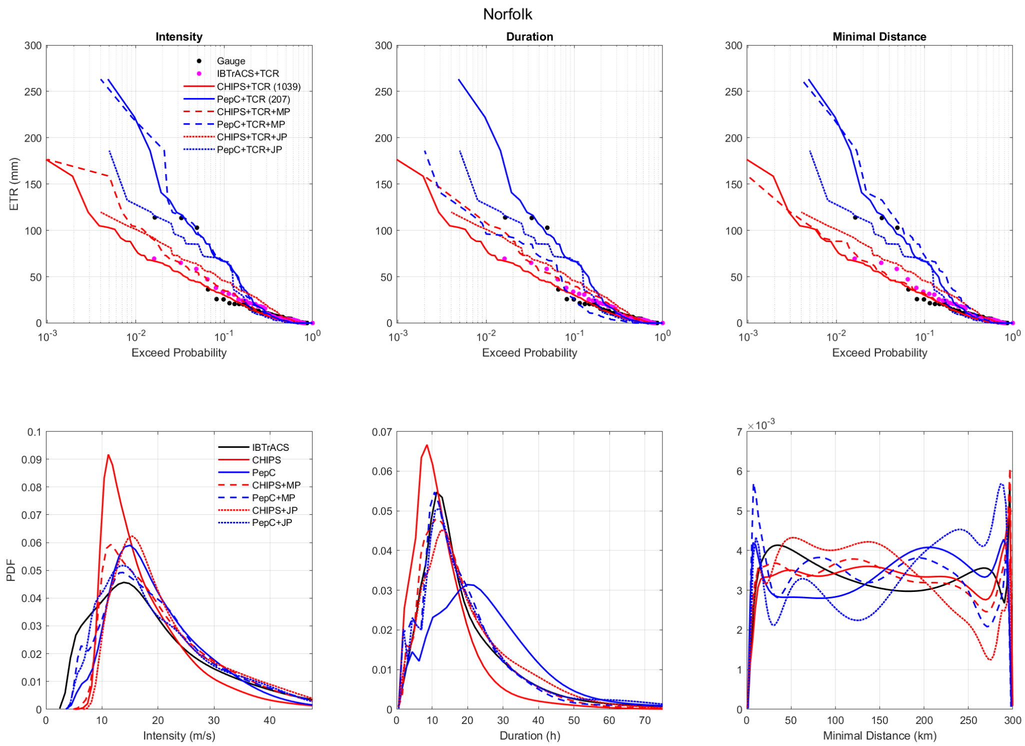
Figure 7. Comparison of the observed and simulated ETR exceedance rate and the probability distribution of parameters in the 10 selected sites. Upper panel: ETR exceedance rate. Black dot: exceedance probability of rain gauge observations; Green dot: exceedance probability of IBTrACS-TCR simulation; Red solid line: exceedance probability of CHIPS-TCR; Blue solid line: exceedance probability of PepC-TCR; Red dashed line: exceedance probability of CHIPS-TCR-MP; Blue dashed line: exceedance probability of PepC-TCR-MP; MP and JP stand for performing resampling based on the marginal probability distribution and joint probability distribution, respectively. Bottom panel: probability distribution of the parameters. The numbers behind CHIPS+TCR and PepC+TCR are numbers of TCs hit the POI in each dataset. CHIPS-MP (PepC-MP) stands for the probability distribution of the important parameters in CHIPS (PepC) model based on the marginal probability distribution. CHIPS-JP (PepC-JP) stands for the probability distribution of the important parameters in CHIPS (PepC) model based on the joint probability distribution



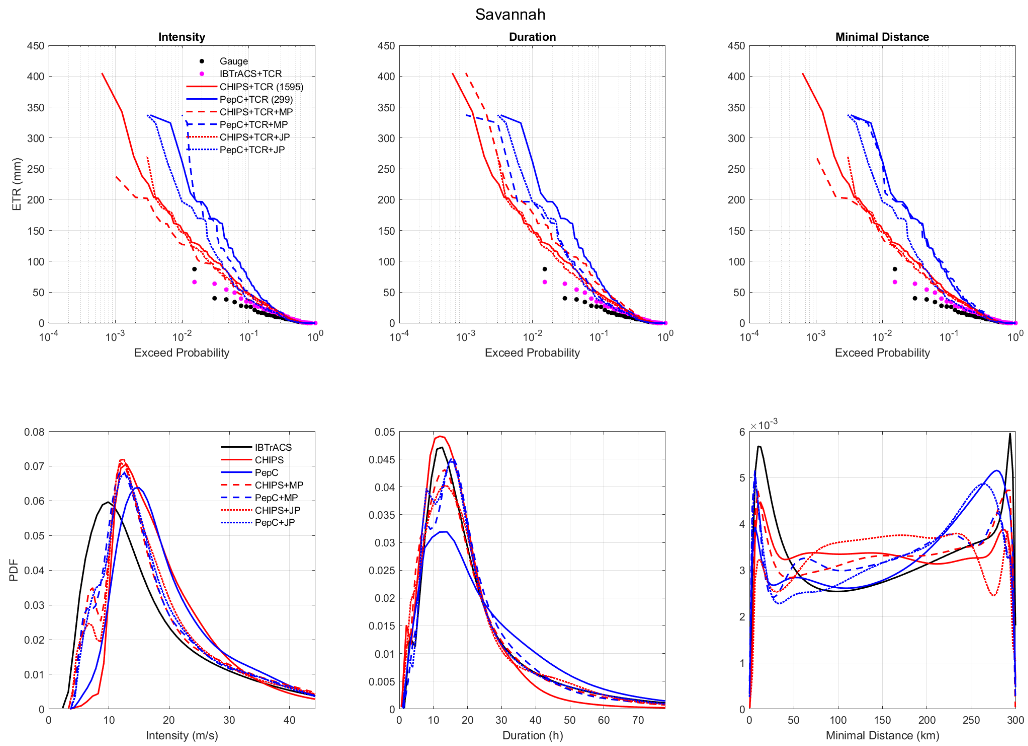
Continuation of Figure 7, for Baltimore.



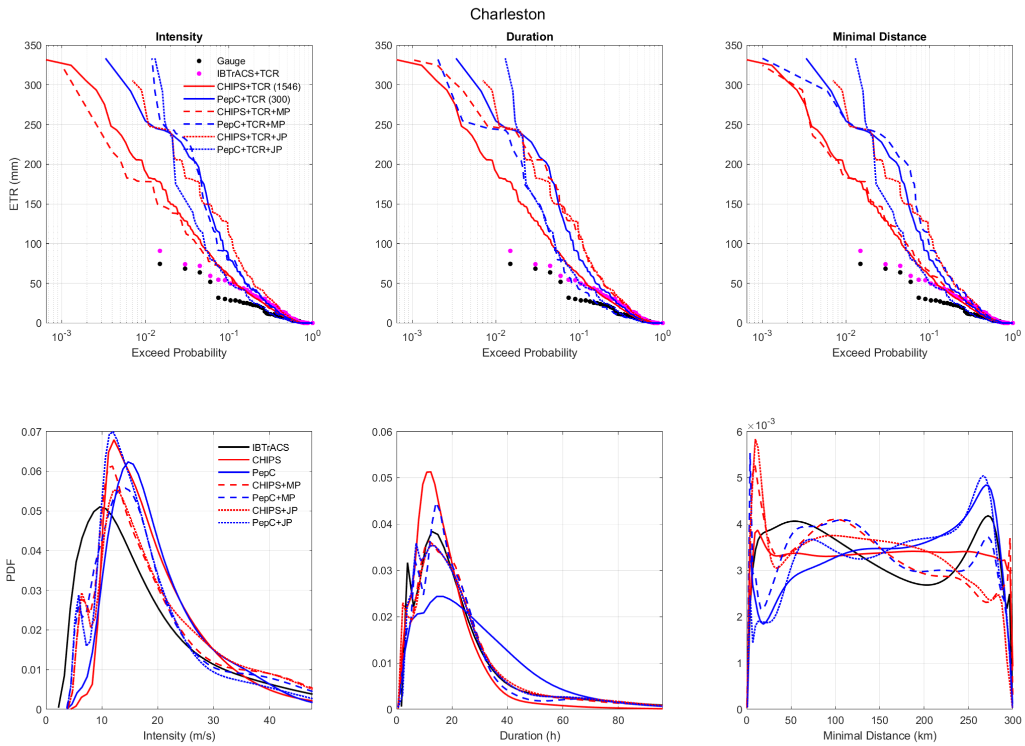
Continuation of Figure 7, for Mobile.



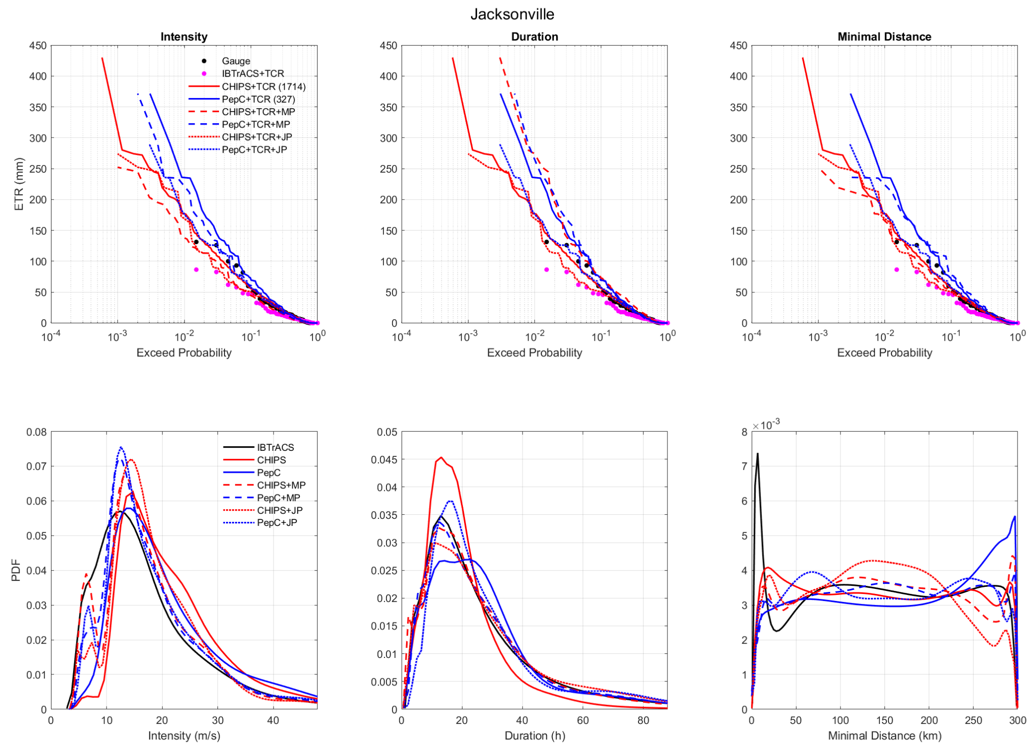
Continuation of Figure 7, for Norfolk.



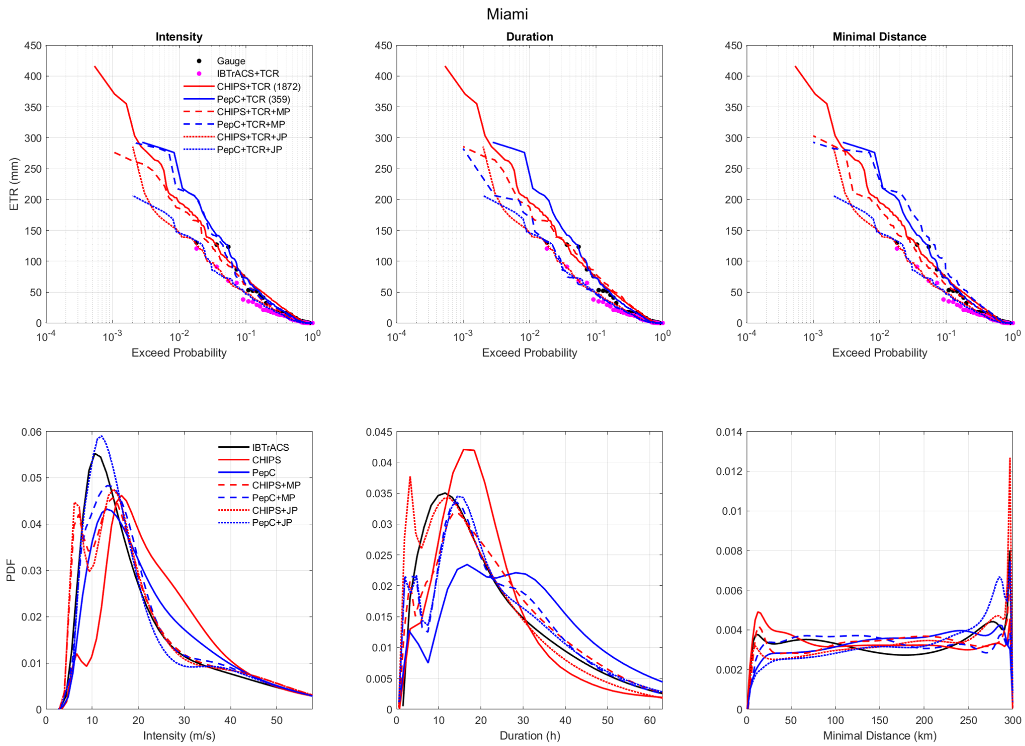
Continuation of Figure 7, for Savannah.



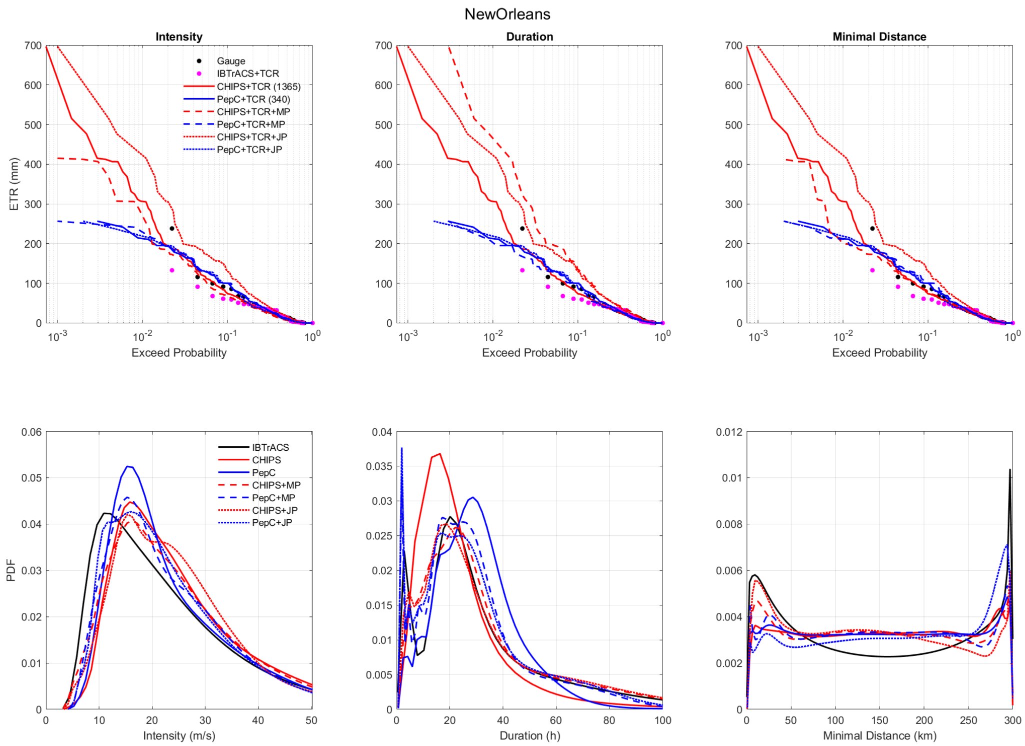
Continuation of Figure 7, for Charleston.



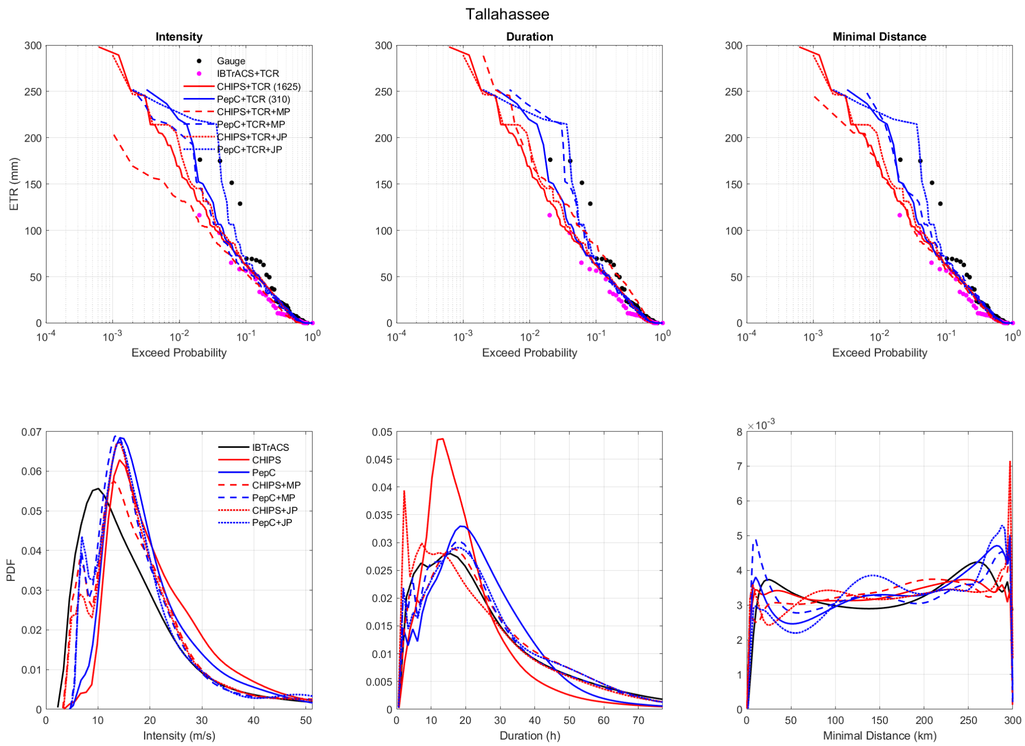
Continuation of Figure 7, for Jacksonville.



Continuation of Figure 7, for Miami.



Continuation of Figure 7, for New Orleans.



Continuation of Figure 7, for Tallahassee.

Bias correction based on the joint distribution of important storm parameters may not always improve the simulation results due to the mismatch between probability distribution of the important parameters after resampling and observation. For locations where joint parameter bias correction improves the simulation results, it is still not guaranteed that the probability distribution of the resampled important parameters matches the observations. One important reason that relates to the unsatisfactory matches with the parameters is the difficulty in resampling using the importance sampling method for high dimensional probability distribution. Resampling for higher (>2) dimensional probability distribution is harder to perform than for lower dimensions. It is mathematically provable that the uncertainty of importance sampling grows with the dimension of the probability distribution (Kroese and Rubinstein, 2016). For those locations where bias correction based on the joint parameter improves the rainfall estimation, it is likely that although the corrected probability distribution does not match the observations well for each single parameter individually, the Gaussian Copula model captures the three-dimensional joint features of the three important parameters well, and thus the corrected rainfall distribution predicted by PepC or CHIPS matches well with IBTrACS estimated rainfall distribution. We even found that in some locations (CHIPS for Charleston and New Orleans), the bias correction based on the joint parameters worsens the simulation result. Both in Charleston and New Orleans, CHIPS estimated rainfall has satisfactory performance, but the good performance is a result of the combined effects of overestimation in intensity and underestimation in duration.

The above analysis shows that the distributions of the selected parameters have profound influences on the simulated TC rainfall hazard probability, and the differences of the distributions of the parameters can be used to understand the different performances of synthetic rainfall hazard assessments. We also found that bias-correction based on the single parameter that is mostly biased can usually improve the TC rainfall hazard estimation. Using the joint probability cannot always improve the estimated TC rainfall hazard due to the limited ability to match the marginal and joint distributions of all individual parameters. It is worthwhile for future research to continue developing improved statistical methods for bias correction of TC rainfall hazard.

4 Discussion

4.1 Wind Input and Wind Hazard

As the wind input is important to the simulation of rainfall in TCR, and the strong wind itself is also an important hazard associated with TCs, here we briefly explore the simulated return period of the event hourly maximum wind. We used the wind profile model proposed by Chavas et al. (2015) to perform the wind simulation and to prepare wind profile inputs for TCR. First, we briefly discuss what TC-related parameters may be important for simulated maximum wind. As the inputs for the C15 wind profile model are only the radii of maximum wind and the maximum wind, and given the wind profile, the wind a POI experiences depends only on the distance between the POI and the TC center. Thus, three parameters are important for event maximum wind: distance, storm intensity and radius of maximum wind. The 10-year return period of the maximum wind simulated from PepC-C15 compares well with the simulation based on historical TCs while CHIPS-C15 tends to underestimate event maximum wind (Figure 8). The differences in simulated wind from the two models explain the lower rainfall hazards estimated from CHIPS-TCR than from PepC-TCR in the Southern and East Coasts (Figure 4). The discrepancy is due to the bias of the synthetic storm models. As the minimal distance simulated from CHIPS is larger than that in the observations (Figure 3d-e), the underestimation of event maximum wind is likely due to the longer distance from TCs to the POI in CHIPS. The bias of simulated wind will be fed into TCR and influence the accuracy of the TC rainfall simulation. We further explore the return levels of event maximum wind according to 50-year and 100-year return periods using synthetic storms and the C15 wind profile model (Figure 9). We find that CHIPS-C15 and PepC-C15 show agreement on the return levels in areas that are farther than 100 km from the coast lines. PepC-C15 shows more intense wind for 50-year and 100-year return periods than CHIPS-C15 in coastal areas.

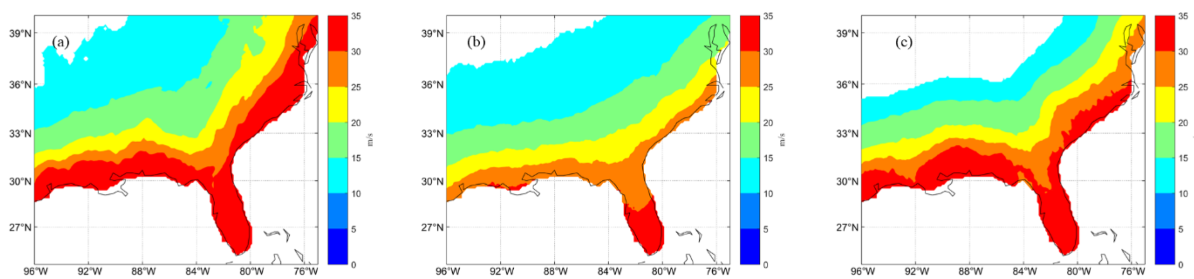


Figure 8. Simulated return levels for 10-year return level of event maximum wind from a). Historic-C15 b). CHIPS-C15 c). PepC-C1.5

We notice that in general the discrepancies of event maximum wind are smaller than the ETR simulated by the two synthetic storm models. For example, the most significant difference between the 10-year return level for ETR from PepC-TCR and CHIPS-TCR is found in Coastal North Carolina, where the PepC-TCR simulated rainfall is two times larger than CHIPS-TCR (Fig. 4). In the wind hazard simulation, though, the maximum difference is found in Georgia, where the difference is less than 20% (Fig. 8). This finding implies that compared to ETR, event maximum wind is less sensitive to the discrepancies in the features of synthetic storms between different datasets. The reason is likely related to the different parameters that are important to the different hazards. As we have mentioned, the important parameters for rainfall are distance, duration, and intensity of the storm. They are three physically unrelated parameters. However, the important parameters for event maximum wind are radius of maximum wind, storm intensity, and distance, and the first two parameters are found to be negatively related. Thus, there are more degrees of freedom in variables that control TC rainfall hazard than in those for wind hazard, so the simulated ETR by different models shows more discrepancies than the simulated event maximum wind. Given that the C15 wind profile itself has uncertainties and we lack observations of hourly wind covering the United States, we hereby only acknowledge the uncertainties of estimation of wind hazard from synthetic storm models without suggesting which synthetic storm model is better for wind hazard assessment.

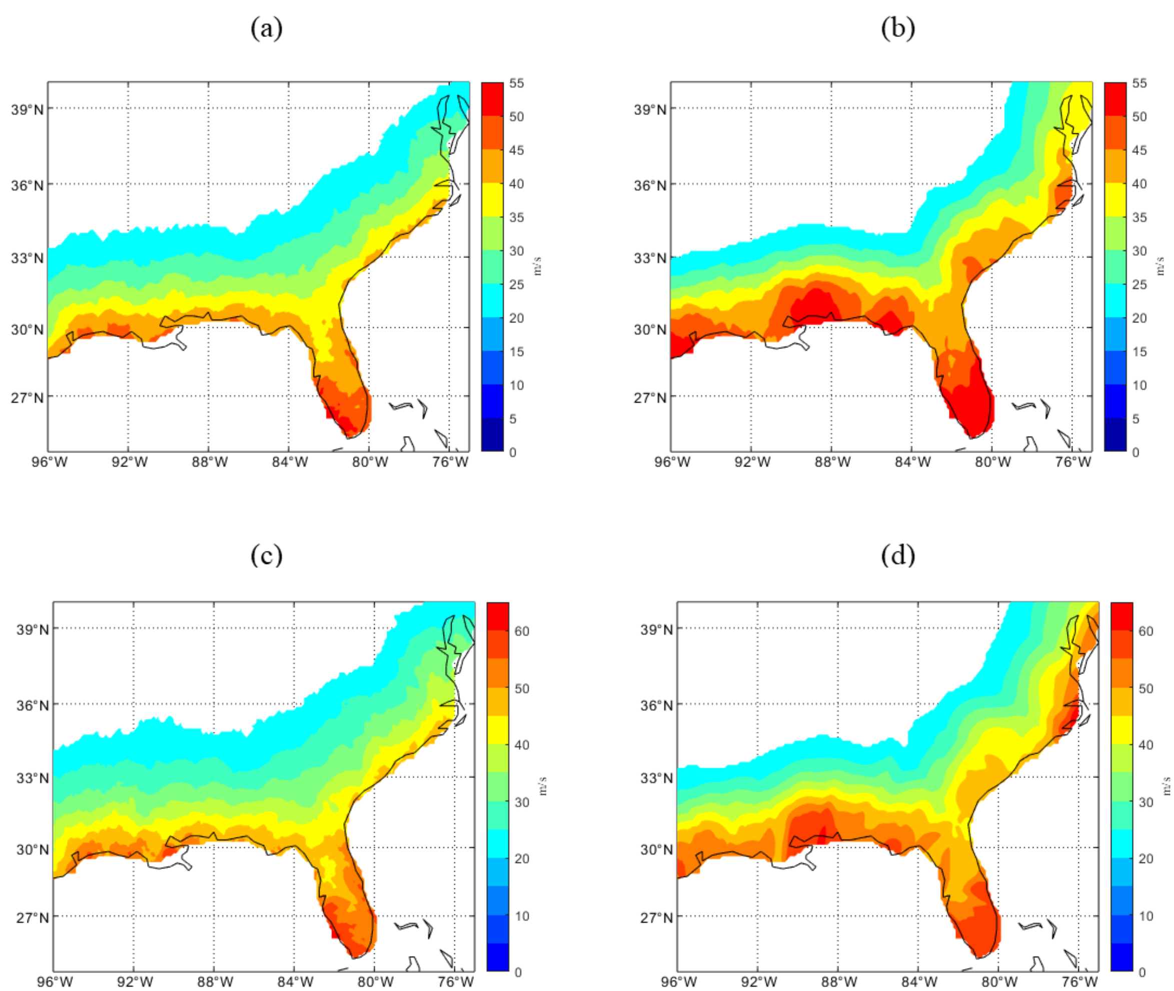


Figure 9. Simulated return levels of event maximum wind from CHIPS a) c) and PepC b) d). 50-year return level: a), b); 100-year return level: c), d).

4.2 Important Parameters for Maximum Rain Rate

In the main results, we focused on the hazard of ETR. Besides ETR, hourly maximum rain rate is also of interest. Hourly maximum rain rate is associated with various hazards including flash flooding and compound flooding (Gori et al. 2019), and previous research has not addressed this variable much. Here we briefly discuss the TC-related parameters that are important for the maximum rain rate from the TCR simulation. We apply the same LASSO analysis as described in Section 3.1 but change the response to hourly maximum rain rate in the TCR simulation. We found the important parameters for maximum rain rate are maximum intensity, minimal distance, and average radius of maximum wind, similar to the case for event maximum wind. The sharing of important parameters between event maximum wind and maximum rain rate indicates that the wind field feature may more critically affect maximum rain rates than ETR. Similar to event maximum wind, maximum rain rate is also a maximum value, rather than a summation value such as ETR, so the hazards of maximum rain rates and event maximum wind share similar important TC-related parameters. As noted in Xi et al. (2020), TCR tends to underestimate the chances of extreme short-term rainfall (even when driven by correct storm parameters); future studies could consider applying a hierarchical bias-correction method to bias-correct both TCR and important parameters for hourly extreme rainfall events.

4.3 Potential Applications of this Study

The results of this study have several potential applications. First, understanding the storm parameters important for TC rainfall can be useful to understand the bias of synthetic storm estimated TC rainfall hazard, as well as to understand the driving force for TC rainfall hazard change in the future. For example, the projected change of storm intensity (Knuston et al. 2010, Emanuel 2020) and lengthened storm stalling time (Kossin et al. 2018) may both favor heavier TC rainfall in the future; applying the important sampling method developed in this study could further examine which factor has the most impact on future TC rainfall hazard change. Second, this study potentially provides a new way to correct TC hazard bias. Previous studies apply the quantile-quantile mapping method directly to bias-correct the simulated hazard probability distribution (Marsooli et al. 2019). The quantile-quantile mapping generates a good match with observations for the estimated hazard curve, but it is not performed at the storm level, which limits the method in capturing the correlations between different hazards (wind, surge, rainfall) caused by the same storm. However, the important parameters for the other hazards may differ from the important parameters for rainfall; future studies may investigate the important storm parameters for joint hazards to support storm-level hazard bias-correction. Finally, the identification of important parameters for TC rainfall is useful to design the Joint Probability Method hazard estimation technique for TC rainfall. The Joint Probability Method generates synthetic storms that have various features (intensity, distance to POIs, translation speed, etc.), assigns probability mass (occurrence rate) for each storm based on the historical joint probability distribution of the important parameters, and applies the generated storms for hazards modeling (Toro et al. 2010). Such a method performs simulations and analysis for local coastal locations and thus may have less bias locally than the basin-wide synthetic storm models such as those used in this study. The current Joint Probability Method is designed based on important parameters identified for TC surge (Toro et al. 2010); similar techniques may be applied to TC rainfall hazard estimation based on the identified important parameters for TC rainfall.

5. Conclusion

This study examines the important parameters that control the TC rainfall hazard. We first identified three important parameters of storms that have significant impacts on ETR and then explored the favorable conditions for extreme ETR events by analyzing both radar observations and TCR simulations. We coupled TCR with two synthetic storm models, CHIPS and PepC, to simulate TC rainfall events for TC rainfall hazard assessments. The bias of rainfall hazards simulated by the synthetic storm coupled TCR models is explained by the bias of the important parameters for TC rainfall simulated by the synthetic storm models. The main findings of this study are summarized as follows:

1. Maximum intensity, minimal distance, and duration are the three most important parameters that control the ETR of a TC rainfall event. Higher intensity, shorter minimal distance, and longer duration are favorable for both historical storms and TCR to produce ETR.

2. Examination of the probability distribution of important parameters and ETR from Stage-IV observations and TCR simulation shows that low intensity and short duration favor ordinary rainfall events while high intensity and long duration favor extreme rainfall events. TCR and observations share similar conditions that produce extreme rainfall events, indicating that the simplified TC rainfall theory employed by TCR can explain the occurrence of extreme rainfall. However, in TCR simulation a weak storm is unlikely to produce heavy ETR, while in observations such events are non-neglectable.
3. The CHIPS-TCR and PepC-TCR have some discrepancies in terms of the estimated rainfall hazard. On the East Coast, PepC-TCR estimates higher TC rainfall hazard than CHIPS-TCR, as a result of longer TC duration predicted by PepC than CHIPS. On the Texas Coast, PepC-TCR predicts higher rainfall hazard than CHIPS-TCR due to the shorter minimal distance and longer TC duration predicted by PepC than by CHIPS. This analysis shows that the differences in the simulated important TC parameters may largely explain the differences in the simulated TC rainfall hazards.
4. Bias of the TCR coupled with synthetic storm simulation can be largely explained by the bias of the three important TC-related parameters estimated by the synthetic storm models. Correcting the distribution of the most biased parameter may significantly improve rainfall hazard estimation. Bias correction based on the joint distribution of the important parameters may render more accurate rainfall hazard estimations in most but not all cases, and the matching of the probability distributions of all important parameters is not guaranteed. Bias correction based on the joint probability distribution suffers from the general technical difficulties in resampling from high-dimensional joint probability distributions, which may be further explored in future research.

This study identifies the important parameters for TC rainfall hazards as a way to understand and potentially reduce the bias in TC rainfall hazard estimation. The results are obtained based on observations and a physics-based TC rainfall model that has satisfactory performance in comparison with observations (Figure 2, see also Feldmann et al. 2019, Xi et al. 2020) and full-physics numerical models (Lu et al. 2018). Thus, the identified TC parameter-rainfall hazard relationships may also be used to test other TC rainfall models. The findings of the current study suggest future research also in the following ways. First, the impact of climate change on TC rainfall hazard may be assessed by investigating the changes of the identified important TC parameters in various climate simulations and statistical downscaling datasets. For example, TC intensity (Knuston et al. 2020, Emanuel et al. 2021), track (Studholme et al. 2021, Wang and Toumi 2021), and impact duration (Xi and Lin 2021) and the correlations between these parameters may change, likely leading to changes in TC rainfall hazard. Since the TC parameter-rainfall relationships are identified based on both observations and physics-based modeling, they may not change significantly under climate change. Second, the discussion of the important parameters for ETR, wind hazard (Section 4.1), and flash flooding (Section 4.2), together with previous research on the important parameters for storm surge (Resio et al. 2009, Irish et al. 2009) and compound flooding (Gori et al. 2020), provide powerful tools for understanding various hazards associated with TCs and their modeling uncertainties. Moreover, the Joint Probability Method for TC rainfall hazard and multi-hazards could be developed based on the results of this study for engineering applications. Finally, in this study, we focused on the bias in rainfall hazard estimation caused by storm simulations. However, to accurately estimate TC rainfall hazard, the intrinsic bias in TC rainfall modeling (e.g., TCR) should also be investigated and corrected in future research.

Acknowledgments. This material is based upon work supported by the National Science Foundation (NSF award 1854993) and by the U.S. Army Corps of Engineers (USACE) Flood and Coastal Systems R&D Program through the Oak Ridge Institute for Science and Education (ORISE) Research Participation Program as part of an interagency agreement between the U.S. Department of Defense (DOD) and the U.S. Department of Energy (DOE). ORISE is managed by ORAU under DOE contract DE-SC0014664. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of NSF, USACE, DOD, DOE, or ORAU/ORISE.

Data Availability Statement. The observed TC dataset IBTrACS can be found in https://www.ncdc.noaa.gov/ibtracs/ . Radar observation of TCs (Stage IV) can be found in https://data.eol.ucar.edu/dataset/21.093 . Rain gauge observations are obtained from National Centers For Environmental Information (https://www.ncdc.noaa.gov/). The PepC synthetic data followings the data availability of Jing and Lin (2020) https://doi.org/10.1029/2019MS001975 , and the CHIPS synthetic data following the data availability of Emanuel (2021) https://doi.org/10.1175/JCLI-D-20-0367.1

References:

Aryal, Y.N., G. Villarini, W. Zhang, and G.A. Vecchi, Long term changes in flooding and heavy rainfall associated with North Atlantic tropical cyclones: Roles of the North Atlantic Oscillation and El Niño-Southern Oscillation, Journal of Hydrology, 559, 698-710, 2018.

Chavas, D. R., N. Lin, and K. Emanuel, 2015: A model for the complete radial structure of the tropical cyclone wind field. Part I: Comparison with observed structure. *J. Atmos. Sci.*, **72**, 3647–3662, https://doi.org/10.1175/JAS-D-15-0014.1.

Cheung, K. K.W., L.-R. Huang, and C.-S. Lee, 2008: Characteristics of rainfall during tropical cyclone periods in Taiwan. Nat. Hazards Earth Syst. Sci., 8, 1463–1474, https://doi.org/10.5194/nhess-8-1463-2008.

Braun, S. A., and L. Wu, 2007: A numerical study of Hurricane Erin (2001). Part II: Shear and the organization of eyewall vertical motion. Mon. Wea. Rev., 135, 1179–1194, https://doi.org/10.1175/MWR3336.1.

Emanuel, K., 2017: Assessing the present and future probability of Hurricane Harvey's rainfall. *Proc. Natl. Acad. Sci.*, **114**, 12681–12684, https://doi.org/10.1073/pnas.1716222114.

Emanuel, K., R. Sundararajan, and J. Williams, 2008: Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bull. Am. Meteorol. Soc.*, **89**, 347–367, https://doi.org/10.1175/BAMS-89-3-347.

Emanuel, K., 2021: Response of global tropical cyclone activity to increasing CO2 : Results from downscaling CMIP6 models. *J. Climate,* **34**, 57-70, doi:10.1175/jcli-d-20-0367.1.

Feldmann, M., K. Emanuel, L. Zhu, and U. Lohmann, 2019: Estimation of atlantic tropical cyclone rainfall frequency in the United States. *J. Appl. Meteorol. Climatol.*, **58**, 1853–1866, https://doi.org/10.1175/JAMC-D-19-0011.1.

Gori, A., Lin, N., and Xi, D. (2020). Tropical Cyclone Compound Flood Hazard Assessment: From Investigating Drivers to Quantifying Extreme Water Levels. Earth's Future 8 (12). doi:10.1029/2020ef001660

Hall, T. M., & Kossin, J. P. (2019). Hurricane stalling along the North American coast and implications for rainfall. Npj Climate and Atmospheric Science, 2(1), 1– 9. https://doi.org/10.1038/s41612-019-0074-8

Houze, R. A., 2010: Clouds in tropical cyclones. Mon. Wea. Rev., 138, 293–344, https://doi.org/10.1175/2009MWR2989.1

Irish JL, Resio DT, Cialone MA (2009) A surge response function approach to coastal hazard assessment. Part 2: quantification of spatial attributes of response functions. Nat Hazards 51:183–205

Jing, R., and N. Lin, 2020: An Environment-Dependent Probabilistic Tropical Cyclone Model. *J. Adv. Model. Earth Syst.*, **12**, 1–18, https://doi.org/10.1029/2019MS001975.

Knutson, T. R., and Coauthors, 2010: Tropical cyclones and climate change. *Nat. Geosci.*, **3**, 157–163.

Kossin, J. P. (2018). A global slowdown of tropical-cyclone translation speed. Nature, 558(7708), 104–107. https://doi.org/10.1038/s41586-018-0158-3

Kroese, Dirk P.., Rubinstein, Reuven Y (2016) Simulation and the Monte Carlo Method.Lin, Y., and K. E. Mitchell, 2005: The NCEP Stage II/IV hourly precipitation analyses: Development and applications. 19th Conf. on Hydrology, San Diego, CA, Amer. Meteor. Soc.,1.2, https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.

Lin, N., Emanuel, K., Oppenheimer, M., & Vanmarcke, E. (2012). Physically based assessment of hurricane surge threat under climate change. Nature Climate Change, 2(6), 462–467. https://doi.org/10.1038/nclimate1389

Liu, M., G. A. Vecchi, J. A. Smith, and H. Murakami, 2018: Projection of landfalling–tropical cyclone rainfall in the eastern United States under anthropogenic warming. *J. Climate*, 31, 7269–7286, https://doi.org/10.1175/JCLI-D-17-0747.1.

Lu, P., N. Lin, K. Emanuel, D. Chavas, and J. Smith, 2018: Assessing hurricane rainfall mechanisms using a physics-based model: Hurricanes Isabel (2003) and Irene (2011). J. Atmos. Sci., 75, 2337–2358, https://doi.org/10.1175/JAS-D-17-0264.1.

Luitel, B., G. Villarini, and G.A. Vecchi, Verification of the skill of numerical weather prediction models in forecasting rainfall from U.S. landfalling tropical cyclones, Journal of Hydrology, 556, 1026-1037, 2018.

Marsooli, R., N. Lin, K. Emanuel, and K. Feng, 2019: Climate change exacerbates hurricane flood hazards along US Atlantic and Gulf Coasts in spatially varying patterns. *Nat. Commun.*, **10**, 1–9, https://doi.org/10.1038/s41467-019-11755-z.

Resio, D., J. Irish, and M. Cialone (2009), A surge response function approach to coastal hazard assessment – part 1: basic concepts, Nat. Hazards, 51(1), 163– 182, doi:10.1007/s11069-009-9379-y.

Rodgers, E. B., , and R. F. Adler, 1981: Tropical cyclone rainfall characteristics as determined from a satellite passive microwave radiometer. *Mon. Wea. Rev*, **109 ,** 506–521.

Shapiro, L. J., 1983: The asymmetric boundary layer under a translating hurricane. J. Atmos. Sci., 40, 1984–1998, https://doi.org/10.1175/1520-0469(1983)040,1984:TABLFU.2.0.CO;2

Studholme, J., Fedorov, A.V., Gulev, S.K., Emanuel, K. and Hodges, K., 2021. Poleward expansion of tropical cyclone latitudes in warming climates. Nature Geoscience, pp.1-15.

Tamarin-Brodsky, T., and Y. Kaspi, 2017: Enhanced poleward propagation of storms under climate change. *Nat. Geosci.*, **10**, 908–913, https://doi.org/10.1038/s41561-017-0001-8.

Tishbirani, R., 1996: Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Ser. B*, **58**,

Toro, G. R., D. T. Resio, D. Divoky, A. W. Niedoroda, and C. Reed (2010), Efficient joint-probability methods for hurricane surge frequency analysis, Ocean Eng., 37(1), 125– 134.

Tuleya, R. E., M. DeMaria, and R. J. Kuligowski, 2007: Evaluation of GFDL and simple statistical model rainfall forecasts for US landfalling tropical storms. Wea. Forecasting, 22, 56–70, https://doi.org/10.1175/WAF972.1.

Villarini, G., J.A. Smith, M.L. Baeck, T. Marchok, and G.A. Vecchi, Characterization of rainfall distribution and flooding associated with U.S. landfalling tropical cyclones: analyses of Hurricanes Frances, Ivan, and Jeanne (2004), Journal of Geophysical Research, 116, D23116, doi:10.1029/2011JD016175, 2011.

Villarini, G., W. Zhang, P. Miller, D. Johnson, L. Grimley, and H. Roberts, Probabilistic rainfall generator for tropical cyclones affecting Louisiana, Inter.national Journal of Climatology, https://doi.org/10.1002/joc.7335, 2021

Wang, S. and Toumi, R., 2021. Recent migration of tropical cyclones toward coasts. Science, 371(6528), pp.514-517.

Willoughby, H. E., F. D. Marks, and R. J. Feinberg, 1984: Stationary and moving convective bands in hurricanes. J. Atmos. Sci., 41, 3189–3211, https://doi.org/10.1175/1520 0469(1984)041,3189:SAMCBI.2.0.CO;2.

Xi, D., N. Lin, and J. Smith, 2020: Evaluation of a physics-based tropical cyclone rainfall model for risk assessment. Journal of Hydrometeorology, 21 (9), 2197–2218.

Xi, D. and Lin, N., 2021. Sequential landfall of tropical cyclones in the United States: From historical records to climate projections. Geophysical Research Letters, 48(21), p.e2021GL094826.

Yang, M. J., S. A. Braun, and D. S. Chen, 2011: Water budget of Typhoon Nari (2001). Mon. Wea. Rev., 139, 3809–3828, https://doi.org/10.1175/MWR-D-10-05090.1.

Yin, H., 2005: A consistent poleward shift of the storm tracks in simulations of 21st century climate. *Geophys. Res. Lett.*, **32 .**L18701, doi:10.1029/2005GL023684.

Zhu, L., S. M. Quiring, and K. A. Emanuel, 2013: Estimating tropical cyclone precipitation risk in Texas. *Geophys. Res. Lett.*, **40**, 6225–6230, https://doi.org/10.1002/2013GL058284.