

# Hurricane Simulation and Nonstationary Extremal Analysis for a Changing Climate

Meagan Carney<sup>a</sup> Holger Kantz<sup>b</sup> Matthew Nicol<sup>c</sup>

<sup>a</sup>The University of Queensland, Brisbane, AUS; <sup>b</sup>Max Planck Institute for the Physics of Complex Systems, Dresden, DE; <sup>c</sup>University of Houston, Houston, USA

## ARTICLE HISTORY

Compiled November 23, 2021

## ABSTRACT

Particularly important to hurricane risk assessment for coastal regions is finding accurate approximations of return probabilities of maximum windspeeds. Since extremes in maximum windspeed have a direct relationship to minimums in the central pressure, accurate windspeed return estimates rely heavily on proper modeling of the central pressure minima. Using the HURDAT2 database, we show that the central pressure minima of hurricane events can be appropriately modeled by a nonstationary extreme value distribution. We also provide and validate a Poisson distribution with a nonstationary rate parameter to model returns of hurricane events. Using our nonstationary models and numerical simulation techniques from established literature, we perform a simulation study to model returns of maximum windspeeds of hurricane events along the North Atlantic Coast. We show that our revised model agrees with current data and results in an expectation of higher maximum windspeeds for all regions along the coast with the highest maximum windspeeds occurring in the northern part of the coast.

## KEYWORDS

Extreme Value Theory; Extreme Weather; Hurricane Simulation; Climate Change

## 1. Introduction

Hurricanes and tropical storms bring massive societal impacts and cause economic instabilities. Known for their high windspeeds and downpours, these storms are often accompanied by flooding, wind damage, and travel hazards that lead to large-scale evacuations and a national emergency response. Talk of climate change in recent years and more frequent observations of extreme weather events has inspired research into techniques that provide more accurate estimates of returns and return times of extremes [3, 4, 13].

Particularly important in hurricane risk assessment for coastal regions is finding accurate approximations of the return probabilities of maximum windspeeds. There have been several studies surrounding maximum windspeed return estimates of hurricanes occurring along the North Atlantic Coast [1, 2, 5, 12, 17]. Many of these studies use the retired HURDAT database which has since been discounted as an unreliable source for future prediction modeling. In 1999, Casson and Coles purposed a hurricane model that allows for approximations of maximum windspeed returns using the tracks

---

<sup>a</sup>m.carney@uq.edu.au; <sup>b</sup>kantz@pks.mpg.de; <sup>c</sup>nicol@math.uh.edu

and central pressure minima [5]. The advantage of a model over raw data analysis is that a large number of hurricanes can be simulated to provide more accurate estimates of the tail probabilities and longer year returns of such rare events. The simulation results of this model are in good agreement with the other models and analyses of that decade. However, our findings suggest that this model does not hold up in accuracy when fitted to the updated HURDAT2 database. These inaccuracies can be almost entirely attributed to systematic trends in the observed central pressure and frequency of hurricane events over recent years.

Although there are many factors in a hurricane event that affect the maximum windspeed, we find that the most influential for risk assessment are the central pressure minima and translational velocity of a hurricane at the time of impact with the coast. Since the central pressure minima have a direct relationship to the windspeed maxima, a better estimate of their probability distribution can provide more accurate returns of extreme highs of windspeed maxima along the coast. Models of an extreme (e.g. minima or maxima) most often take the form of an extreme value distribution [6, 13]. These distributions have been studied extensively; however, revisions for more complex data analysis settings are often required.

Following the work in [5], we show that we can still reliably model the central pressure minima of a hurricane event using the generalized extreme value distribution (GEV); however, a previously unobserved time dependent trend in the central pressure minima requires adaptations in both the model and methodology. We also provide evidence for a Poisson distribution with a time-dependent rate parameter to model the number of yearly hurricane events that continues into the modern era (post 1965) which previous literature has assumed to be stationary.

Our revised model results in two major differences in the simulation of coastal risk analysis of hurricane events: 1.) higher maximum windspeeds are expected for all regions along the North Atlantic Coast, including the Gulf Coast and 2.) the highest maximum windspeeds are expected to occur in the northern part of the coast. Higher maximum windspeeds are likely due to a combination of the central pressure minima time dependence and increase in the number of observed hurricane events incorporated into the model. The second observation is arguably more surprising since the number of hurricane events hitting the coast in the north is much lower than regions near the Gulf of Mexico. An increase in the translational velocity as hurricanes travel northward explains this effect.

## 2. Methodology

### 2.1. *The Wind Field Model, Maximum Windspeeds, and Minimum Pressure*

We describe the Wind Field Model introduced in [11] and the relationship between maximum windspeeds and minimum central pressure.

Given the center location  $(\phi_t, \psi_t)$  in the usual geographic coordinates (degrees) and central pressure  $p_t$  in hPa of a hurricane measured at the eye at time  $t$ , the Wind Field Model [11] allows us to model the stochastic process of maximum windspeeds of a hurricane as a sequence of random variables sampled at any given time  $t$  by,

$$V(R_{\max}, \phi_t, p_t, u_t) = 0.865 \left( K \sqrt{\Delta p_t} - \frac{R_{\max}(\phi_t) f}{2} \right) + 0.5u_t \quad (1)$$

where  $K$  is a constant given in  $m(s \text{ hPa}^{1/2})^{-1}$ ,  $f = \omega \sin \phi_t$  is the Coriolis parameter  $\omega = 7.2982 \times 10^{-4} \text{ s}^{-1}$ ,  $\Delta p_t = 0.75(1013 - p_t)$  is the pressure differential,  $R_{\max}(\phi_t)$  is the radius to maximum windspeeds in  $m$  sampled from the distribution in the Appendix C1, and  $u_t$  is the translational velocity in  $m/s$  (meters per second) at time  $t$ . The translational velocity  $u_t$  at a time  $t$  is estimated as the change in the distance of the center of the hurricane  $\sqrt{(\phi_t - \phi_{t-1})^2 + (\psi_t - \psi_{t-1})^2}$  over the change in time  $t - 1$  to time  $t$ . For more information on how the variables in the Wind Field Model are related see Table 1.

variable:	$V$	$R_{\max}$	$p_t$	$\phi_t$
depends on:	$R_{\max}, \phi_t, p_t, u_t$	$\phi_t$ and sampled	historical data	historical data
variable:	$\psi_t$	$u_t$	-	-
depends on:	historical data	$(\phi_{t,t-1}, \psi_{t,t-1})$	-	-

Table 1.: Description of variables in the Wind Field Model and their dependence.

Throughout this article, we will define a *hurricane event* as a tropical cyclone taking any form (e.g. tropical depression, tropical storm, hurricane) and denote the total lifetime of a hurricane event as a length of indexed time  $T$  representing the total number of 6-hour time intervals passed since formation. For any given hurricane event, if we are given the *track*,  $(\phi_t, \psi_t)$ , and central pressure timeseries,  $p_t$ , we may use (1) to reconstruct the maximum windspeed  $V(R_{\max}, \phi_t, p_t, u_t)$  for all  $t = 1, \dots, T$  where  $t$  is the index number of 6-hour time intervals passed since formation. From (1) we can see that extreme highs of the maximum windspeed occur for extreme lows of central pressure. Hence, it is important to accurately model the central pressure minima of a hurricane event in order to estimate longer year returns and rare threshold exceedances of maximum windspeeds. Furthermore, central pressure minima often occur at or near landfall so they are particularly important for estimating coastal risk.

We can use tools from extreme value theory to model extremes of a timeseries (e.g. minima or maxima). One well-known strategy is to approximate the set of maxima (or negative minima) taken over blocks of a fixed length  $m$  of a set of independent and identically distributed random variables by the generalized extreme value distribution (GEV) given by,

$$G(x) = \exp \left[ - \left\{ 1 - \frac{k(x - \mu)}{\sigma} \right\}^{-1/k} \right] \quad (2)$$

for  $x : 1 + k(\frac{x - \mu}{\sigma}) \geq 0$  where  $\mu$  is the location parameter,  $\sigma$  is the scale parameter, and  $k$  is the shape parameter that defines the tail behavior of  $G$ . Under certain regularity conditions, we may use maximum likelihood estimation of the parameters  $\mu$ ,  $\sigma$ , and  $k$  to fit the GEV to the block maxima (or negative minima) where each parameter estimate is asymptotically normal provided  $k > -0.5$  [6].

By a standard max-stable argument it is not necessary that the block length  $m$  be fixed, as long as it is *long enough* so that the maxima (or negative minima) can be modeled by its asymptotic GEV. By the same argument, the GEV that is fit to blocks of varying length is related to  $G$  from (2), with different  $\mu = \mu^*$  and  $\sigma = \sigma^*$  parameters. A result of this max-stability property is that we may model the central pressure minima of hurricane events coming from historical records with varying lifetimes  $T$ .

That is, we may model the negative central pressure minima

$$p_{\min} = \min_t p_t \quad \text{for } t = 1, \dots, T \quad (3)$$

by the GEV provided the lifetime of each hurricane event is long enough. We will refer to  $t_{p_{\min}}$  as the time in the total lifetime of the hurricane event that the central pressure minimum is reached.

We can relax the requirement of strict independence for the GEV in (2) provided the time-series is weakly dependent and stationary, see for example [10, Chapter 3] or [13]. Using historical recordings from the HURDAT2 database, we find that the central pressure minima have the same dependence as in [5] on the lifetime  $T$  and the latitude  $\phi_{t_{p_{\min}}}$  where the central pressure minima occurs. Figure C3 in the Appendix depicts scatter plots of the central pressure minima against the lifetime  $T$  and latitude  $\phi_{t_{p_{\min}}}$  for landfalling and non-landfalling hurricanes.

We limit our model to hurricanes with lifetimes  $T \geq 25$  to ensure convergence of the negative central pressure minima to a GEV distribution. There are 642 hurricane events over the years 1851-2019 in the HURDAT2 database that satisfy this requirement. We use maximum likelihood estimation on the parameters of the stationary GEV model proposed in [5]. Although the central pressure minima  $p_{\min}$  are sampled from independent hurricane events, they have some underlying dependence on both the lifetime  $T$  and latitude  $\phi_{t_{p_{\min}}}$  of the hurricane event which is accounted for in the location  $\mu$  and scale  $\sigma$  parameters of this stationary model. Figure C4 illustrates still poor fits for quantile plots of this stationary model.

A natural question is whether there exists some time-dependence in the distribution of central pressure minima.

## 2.2. A Nonstationary Model for Central Pressure Minima

We investigate the time-dependence in the location  $\mu$  and scale  $\sigma$  parameters of central pressure minima for landfalling and nonlandfalling hurricanes. We perform an  $F$ -test for equal variance that indicates the variance of the central pressure minima for landfalling hurricanes has significantly changed ( $p = 0.0037 < 0.05$ ) in the last 40 years. We obtain a similar result using a  $T$ -test for equal means of the central pressure minima for nonlandfalling hurricanes ( $p = 0.0121 < 0.05$ ). Preliminary investigations into the time-dependence of the shape  $k$  parameter showed no obvious trend, so it is taken as constant.

Motivated by the observed difference in the statistical parameters of the central pressure minima in the last 40 years, we now investigate the possibility of a time-dependent trend in the location and scale parameters of the stationary model proposed by [5]. This stationary model asserts a dependence of the location  $\mu$  and scale  $\sigma$  parameters on the lifetime  $T$  of the hurricane and latitude  $\phi_{t_{p_{\min}}}$  of the central pressure minima. We use this model as a basis for checking the time-dependence of the  $\mu$  and  $\sigma$  parameters in the GEV described by (2). We begin by performing maximum likelihood estimation of the all the coefficient parameters used in the stationary model where this estimation is performed on subsets of the 642 historical values of central pressure minima taken over moving time windows of 40 years with a timestep of 1 year. Our final result is a set of time-series representing the maximum likelihood values of the coefficient parameters in the stationary model. We then reconstruct the time-series of the location,  $\mu(t_{\text{yr}})$ , and scale,  $\sigma(t_{\text{yr}})$ , using the relationships described in the

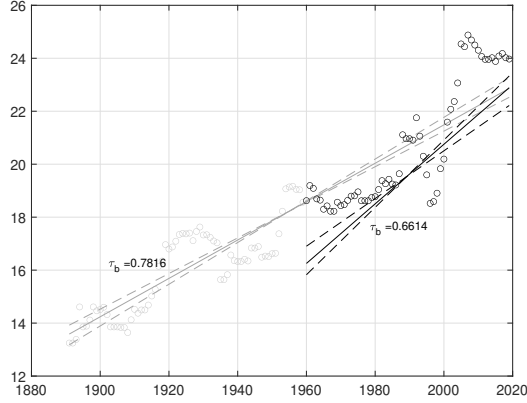


Figure 1.: Timeseries of parameter  $\sigma(t_{\text{yr}}) = \sigma_0(t_{\text{yr}})$  coming from the stationary model for  $-p_{\text{min}}$  in hPa of landfalling hurricanes constructed from likelihood estimates. The value  $\tau_b$  is the Kendall correlation coefficient.

stationary model ( $t_{\text{yr}} = \text{yr} - 1851$ ) and the historical values of  $T$  and  $\phi_{t_{p_{\text{min}}}}$ . From now on, we will refer to the time-series  $\mu(t_{\text{yr}})$  and  $\sigma(t_{\text{yr}})$  as the *location time-series* and *scale time-series*, respectively, to differentiate between the other time-series in this investigation.

Unreliable maximum likelihood estimates of the location and scale parameters in the years 1851-1960 are due to low numbers of recorded hurricane events. Nevertheless, continuous time-dependent trends are noticeable after 1960 for parameters in both the landfalling and nonlandfalling case.

A Mann-Kendall test for trend is performed on each of the location and scale time-series constructed as described above from the 642 historical recordings of central pressure minima. We remark that there are 300 location time-series in the landfalling case and 342 location and scale time-series in the nonlandfalling case. This is because our estimated location (similarly, location and scale) parameter(s) depend on the lifetime  $T$  and the location  $\phi_{t_{p_{\text{min}}}}$  of minimum central pressure of the hurricane where we have 300 (similarly, 342) historical recordings of such lifetimes and locations. This is in contrast to the scale parameter of landfalling hurricanes which does not have a dependence on  $T$  or  $\phi_{t_{p_{\text{min}}}}$  and, as a consequence, results in a single scale time-series.

To determine whether a trend is reliable, we perform the Mann-Kendall test for trend on all the location and scale time-series. We find a positive statistically significant trend for the scale parameter in the landfalling case and a negative statistically significant trend for all time-series of the location parameter in the nonlandfalling case. We do not find clear evidence for a reliable trend in the location time-series for the landfalling case or the scale time-series for the nonlandfalling case. Although, some of the 300 (similarly, 342) location (similarly, scale) time-series do exhibit a significant trend. All trends are determined at the  $\alpha = 0.05$  significance level. The Kendall correlation coefficient is estimated for all years and for years from 1960-2020, for comparison. For an illustration of the trend results and parameter time-series see Figures 1 and 2.

From these results, we propose the following nonstationary model ( $t_{\text{yr}} = \text{yr} - 1851$  is the yearly index),

$$\mu = \mu_0 + \mu_1 \log(T) + \mu_2 \phi_{t_{p_{\text{min}}}}, \quad \sigma = \sigma_0 + \sigma_1 t_{\text{yr}}, \quad k = k_0 \quad (4)$$

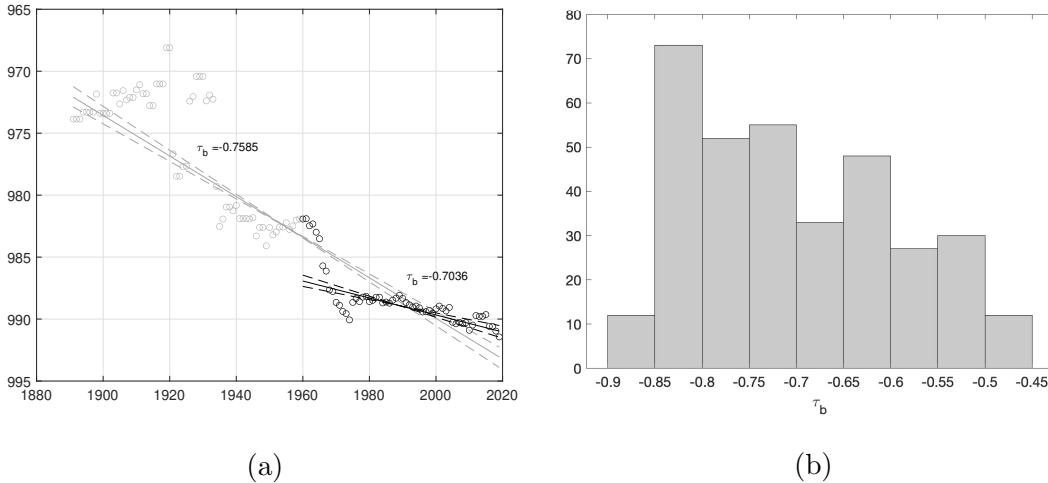


Figure 2.: (a) An example time-series of parameter  $\mu(t_{\text{yr}}) = \mu_0(t_{\text{yr}}) + \mu_1(t_{\text{yr}}) * \log(T)$ , for a single sampled  $T$ . The value  $\tau_b$  is the Kendall correlation coefficient. The  $y$ -axis represents the end year of the 20-year time window chosen for likelihood parameter estimation (b)  $\tau_b$  for all time-series of  $\mu(t_{\text{yr}})$  illustrating that all time-series have a statistically significant negative Kendall correlation coefficient.

for landfalling hurricanes, and

$$\mu = \mu_0 + \mu_1 \log(T) + \mu_2 \log(t_{\text{yr}}), \quad \sigma = \sigma_0 + \sigma_1 \phi_{t_{\text{pmin}}}, \quad k = k_0 \quad (5)$$

for nonlandfalling hurricanes. Maximum likelihood estimates and standard errors estimated from the information matrix are provided in Table 2. We find using the likelihood ratio test that our revised nonstationary model for central pressure minima offers a statistically significant better fit to the data than the stationary model with test statistics well beyond  $\mathcal{L}_{0.05,1} = 3.84$ , the statistic corresponding to the  $\alpha = 0.05$  significance level with 1 degree of freedom (see  $\mathcal{L}$  in Table 2).

Type	$\mu_0(\text{se})$	$\mu_1(\text{se})$	$\mu_2(\text{se})$	
Landfalling	-1078.97(14.48)	32.87(3.60)	-0.52(0.16)	
Nonlandfalling	-1027.13(13.65)	27.40(2.76)	-11.47(1.84)*	
Type	$\sigma_0(\text{se})$	$\sigma_1(\text{se})$	$k(\text{se})$	$\mathcal{L}$
Landfalling	12.47(1.86)	0.07(0.02)*	-0.13(0.04)	15.62
Nonlandfalling	20.48(2.04)	-0.16(0.06)	-0.13(0.04)	30.77

Table 2.: Maximum likelihood estimates of the parameters in the nonstationary generalized extreme value distribution model for  $-p_{\text{min}}$ . Time-dependent parameters are marked with \*. Likelihood ratio test statistics for our revised nonstationary model of  $-p_{\text{min}}$  against the stationary model are indicated by  $\mathcal{L}$ .

### 2.3. Poisson Returns of Hurricane Events and a Nonstationary Rate Parameter

We investigate a Poisson model for yearly returns of hurricane events where the number of expected yearly hurricane events is increasing over time.

Returns of extreme hurricane events, such as low central pressure minima or high maximum windspeeds, are often reported in terms of an  $n$ -year return. In order to interpret returns in this way, our model must consider how often a hurricane event occurs in a given year. Classically, it is expected that a rare event, such as a hurricane, is modeled by a Poisson distribution given by,

$$P(X = \mathcal{K}) = \frac{\lambda^{\mathcal{K}} e^{-\lambda t}}{\mathcal{K}!} \quad (6)$$

where  $\lambda = r/t$  is the rate parameter estimated as the number of events  $r$  in a given time  $t$ .

Under the assumption of stationarity, the authors in [5] estimate a fixed rate parameter,  $\lambda = 5.45$  hurricane events per year, as the average number of returns of a hurricane in a given year over the years 1965-1994. With more data available in the HURDAT2 database, we are able to estimate the time-dependent yearly rate parameter  $\lambda_{t_{\text{yr}}}$  over 20 year sliding windows from 1851-2019. We refer the reader to Figure 3 for an illustration of the estimated yearly rate parameter.

**Remark 1.** We use only hurricane events that have lifetimes  $T \geq 25$  in this analysis to estimate the time-dependent rate parameter  $\lambda_{t_{\text{yr}}}$ ; however, all hurricane events are observed at a further increased rate of  $\lambda_{2019} \approx 16$  which corresponds to current NOAA estimates.

We use likelihood estimation to fit an exponential to the time-dependent Poisson rate parameter  $\lambda_{t_{\text{yr}}}$ . In particular, our model is given as,

$$\lambda(t_{\text{yr}}) = ae^{bt_{\text{yr}}}. \quad (7)$$

Likelihood estimates and confidence intervals of  $a$  and  $b$  can be found in Table 3. Our model for hurricane returns does not differentiate between landfalling and nonland-falling hurricane events due to the nature of the simulation in the final section. This is because tracks of a simulated hurricane are generated by randomly sampling a historical track and adding noise. To compare our results against current literature, we separate the discussion of returns of hurricane events for landfalling and nonlandfalling hurricanes in the paragraphs below.

There is some debate on whether the average number of hurricane events is increasing generally; some literature suggests that low ship density is the underlying cause for the low number of recorded hurricanes for years up to 1965 [8, 15], while others report significant increases in frequency after the late 1980s [14]. When averaging yearly frequency over moving time windows, the authors in [14] report a small nominally positive upward trend post 1878. The work of [8] finds an increase in the occurrence of short lifetime hurricanes only, leading the authors to conclude ship density as a plausible cause for the observed trend. It is important to note that the literature described here uses the retired HURDAT database for their analyses rather than the HURDAT2 database used in this investigation; however, this certainly does not rule out the possibility of historically unrecorded storms in the updated database. There

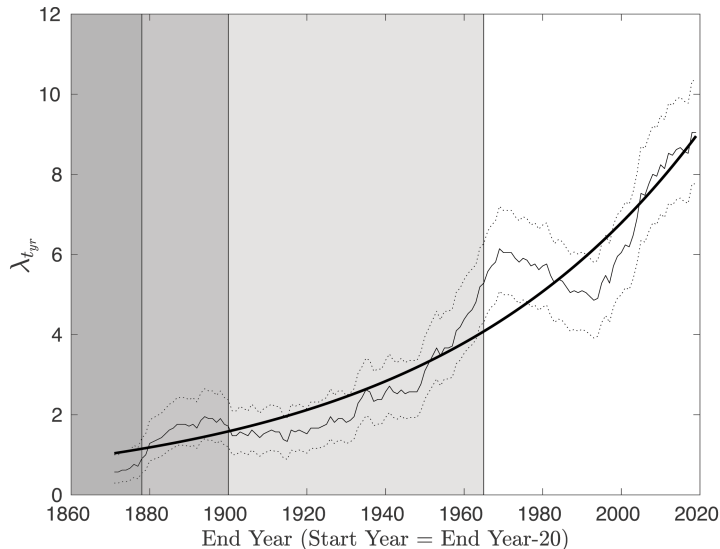


Figure 3.: Likelihood estimate of the Poisson parameter for yearly hurricane event rates with lifetimes greater than 6.25 days. Estimates are taken over 20 year moving time windows. Standard errors are marked with dotted lines. Fitted exponential model is represented by a thick line. Grayed areas correspond to those in [16]: (1) 1878 - year when the U.S. Signal Corps began cataloging all Atlantic hurricanes (2) 1900 - year when the U.S. Coast was sufficiently well-populated for monitoring (3) modern-era with appropriate ship density.

is active research on the frequency of hurricane events recorded in the HURDAT2 database where an observed late-20th century trend is attributed to a possible unusually low minima in the 1980s. Currently, this research is only available in preprint form in [16].

We find an increasing trend in frequency of hurricane events longer than 6.25 days using the Poisson rate parameter, which differs from the results in [8]. This trend holds even into the modern era (post 1965) where ship density is expected to remain steady. One explanation for this difference could be our use of a Poisson rate estimate over a moving average. Rate estimates expect that an increase in the mean results in an increase in the variance. This phenomenon is observed in the raw data. In the case of a moving average estimate this increase in variance can cause statistical tests of the mean difference to be near zero due to large standard errors. We also do not separate hurricane events by windspeed where differences in trend have been reported [14]. Since we limit our investigation to hurricanes with lifespans longer than 6.25 days, our findings may also be a result of some underlying increase in the lifespan of hurricane events as a whole. Finally, using the yearly estimates from our model for the rate over 1965-94 we find that the average is identical to past literature [5] which provides some reasonable benchmark.

An argument could be made that this increase in the total number of observed hurricane events post-1965 comes from our ability to more readily observe nonlandfalling hurricanes. However, an increase in the Poisson rate parameter is also observed for strictly landfalling hurricane events of lifetimes longer than 6.25 days; however, this rate parameter follows a similar pattern (with a low minima in the 1980s) to that of [16] with a slight increase in the current peak compared to that of 1965. We refer to



the Appendix figure C6 for an illustration of the estimated Poisson rate parameter for landfalling hurricane events and figure C7 for the raw data of total yearly observed hurricane events, both with lifetimes longer than 6.25 days.

$a$ (ci)	$b$ (ci)
1.024 (0.925, 1.123)	0.015 (0.014, 0.015)

Table 3.: Maximum likelihood estimates of the exponential model for the time-dependent Poisson parameter  $\lambda_{t_{\text{yr}}}$ .

#### 2.4. Verification of the Nonstationary Model for Central Pressure Minima

We use a combination of established statistical methods to illustrate the reliability of our nonstationary model at predicting returns of central pressure minima.

To test the reliability of our model to accurately predict the distribution of central pressure minima without updating, we break the HURDAT2 database up into a *training* set which we will use to simulate hurricanes from the model and *test* set which we will use to compare risk probability outcomes estimated from the training set against the 'true' probabilities. Our training set will be defined as the set of all years in our dataset minus the number of years  $n$  used to obtain the  $n$ -year returns and our test set will be the  $n$  last years in our dataset. For example, if we are interested in finding the 50-year returns, our training set would be defined as the set of all hurricanes occurring between 1851-1970 and our test set would be the set of all hurricanes occurring between 1971-2019.

Under the assumption that our negative central pressure minima follow some generalized extreme value distribution, [6, Section 6.2.3] suggests the use of a sequence of standardized variables  $z_{t_{\text{yr}}}$  defined for our purposes by,

$$z_{t_{\text{yr}}} = \frac{1}{k} \log \left\{ 1 + k \left( \frac{-p_{\min}(t_{\text{yr}}) - \mu(t_{\text{yr}})}{\sigma(t_{\text{yr}})} \right) \right\} \quad (8)$$

each having a standard Gumbel distribution,

$$P(z_{t_{\text{yr}}} \leq z) = \exp\{-e^{-z}\}, \quad z \in \mathbb{R}. \quad (9)$$

The advantage of using this sequence is that the 'true' quantile plots of the observed and standardized  $-p_{\min}(t_{\text{yr}})$  in the test set can be made with reference to the distribution for the simulated and standardized  $-p_{\min}(t_{\text{yr}})$  from the training set.

We generate data to model negative central pressure minima  $n$ -year returns for the years in the test set using the 1.) parameter likelihoods of  $\mu(t_{\text{yr}})$ ,  $\sigma(t_{\text{yr}})$  and  $k$  defined by the model in (4) and (5) estimated from the training set and 2.) the appropriate rate parameters defined by (7) to compute returns of hurricane events using (6) where  $t_{\text{yr}}$  indices are chosen to correspond to those of the test set. We use this data to compute the model standardized quantile plots for 20-, 30-, and 50-year return periods for both landfalling and nonlandfalling hurricanes. Figures 4, 5, and 6 show model results against the actual data in the test set. Not surprisingly, better approximations for both the landfalling and nonlandfalling case are made for shorter  $n$ -year returns; however,

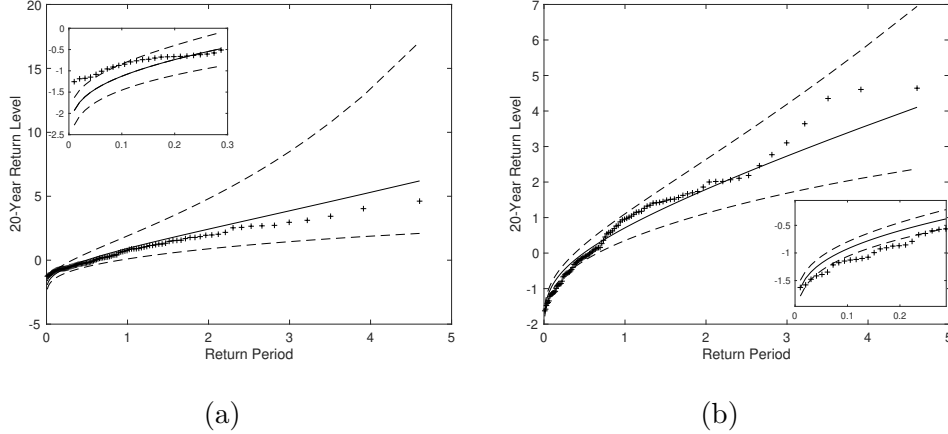


Figure 4.: 20-year return-levels of the standardized  $-p_{\min}$  coming from (8) for (a) nonlandfalling and (b) landfalling hurricanes. Solid lines and dashed lines represent the model and 95% confidence intervals approximated from the training set over the years 1851-2000. Symbol (+) indicates the true return-levels calculated from the test set over the years 2001-2019. Return periods and return levels here are based on (8) are non-dimensional and expected to follow the Gumbel distribution (9).

estimates for 50-year returns still fall reasonably within the 95% confidence interval of the model estimated from the information matrix.

Using the standardized negative central pressure minima allows us to estimate the accuracy of the nonstationary model against true data; however, it does not provide us with a complete way of interpreting the  $n$ -year returns. At best, we are able to fix a year index  $t_{\text{yr}}$  and state the probability of the negative central pressure minima being above a certain threshold in that given year. Most risk analysis involves directly computing  $n$ -year return-levels where a new definition needs to be introduced in the nonstationary setting. We discuss this in detail in the next section.

### 3. Application of Methodology for Coastal Windspeed Risk

#### 3.1. Time-dependent Returns of High Maximum Windspeeds

We discuss a definition for time-dependent  $n$ -year return-levels of maximum windspeeds.

Return-level is often used in risk analysis to communicate the threshold that we are expected to exceed in a give amount of time. For example, we may ask what is the maximum value of the windspeed that we are expected to exceed in  $n$  years. When accounting for nonstationary effects, such as those brought on by climate change, the probability of observing values above or below a threshold varies over time so that terms like return-level no longer make physical sense.

[18, Section 4.2] introduces the idea of extending the definition of the  $n$ -year return-level to the nonstationary case by taking the threshold where the expected number of exceedances in  $n$  years is 1 to the non-stationary case. In the context of nonstationary

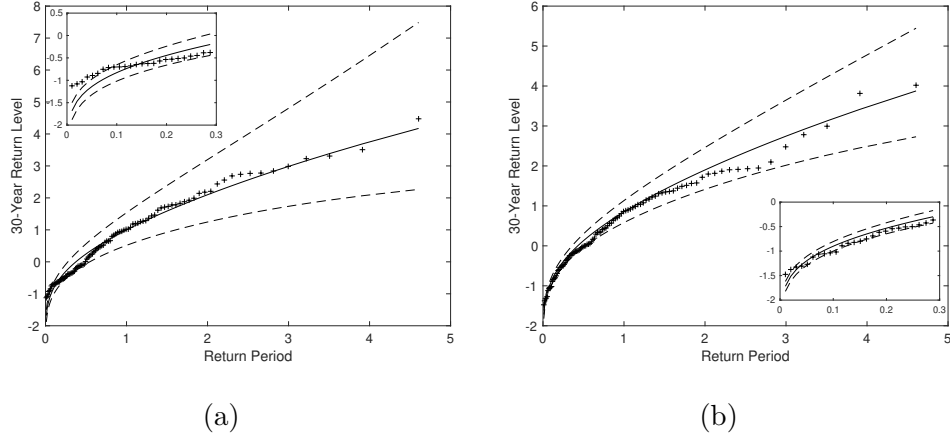


Figure 5.: 30-year return-levels of the standardized  $-p_{\min}$  coming from (8) for (a) nonlandfalling and (b) landfalling hurricanes. Solid lines and dashed lines represent the model and 95% confidence intervals approximated from the training set over the years 1851-1990. Symbol (+) indicates the true return-levels calculated from the test set over the years 1991-2019. Return periods and return levels here are based on (8) are non-dimensional and expected to follow the Gumbel distribution (9).

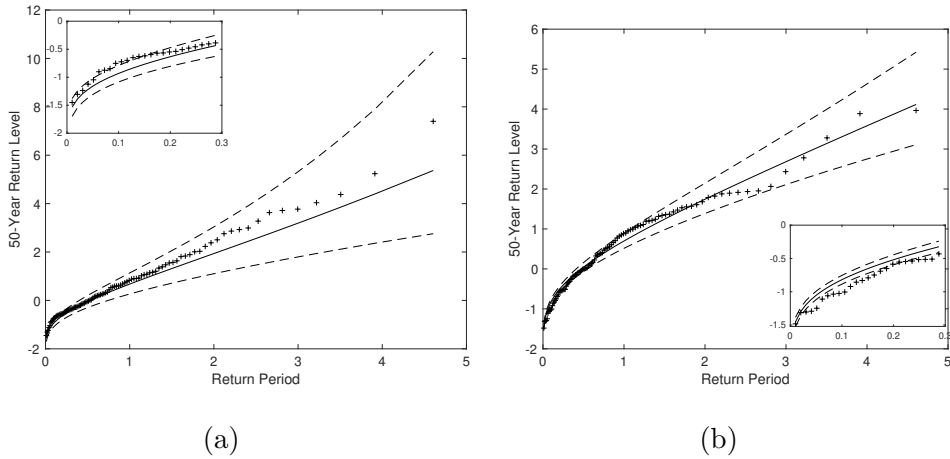


Figure 6.: 50-year return-levels of the standardized  $-p_{\min}$  coming from (8) for (a) nonlandfalling and (b) landfalling hurricanes. Solid lines and dashed lines represent the model and 95% confidence intervals approximated from the training set over the years 1851-1970. Symbol (+) indicates the true return-levels calculated from the test set over the years 1971-2019. Return periods and return levels here are based on (8) are non-dimensional and expected to follow the Gumbel distribution (9).

windspeed prediction this would be equivalent to solving for  $r_n$  in (10),

$$1 = \sum_{t_{\text{yr}}=1}^n (1 - F_{t_{\text{yr}}}(r_n)) \quad (10)$$

where  $r_n$  is the  $n$ -year return-level beginning with year  $t_{\text{yr}} = 1$  and ending with year  $t_{\text{yr}} = n$  and  $F_{t_{\text{yr}}}$  is the unknown indexed yearly cumulative distribution function of maximum windspeed. For example, if we are interested in finding the 50-year return-level  $r_{50}$  of windspeed, (10) would become,

$$1 = \sum_{t_{\text{yr}}=1}^{50} (1 - F_{t_{\text{yr}}}(r_{50})) = P_1(\text{ws} > r_{50}) + \dots + P_{50}(\text{ws} > r_{50}) \quad (11)$$

The corresponding  $n$ -year return-level can be numerically estimated for future years by extrapolating the trend in the model and approximating  $r_n$  by calculating the  $1 - \frac{1}{n}$  quantile of the equal weight mixed probability density function of windspeed occurring over  $t_{\text{yr}} = 1, \dots, n$  years given by,

$$f(x; t_1, \dots, t_n) = \sum_{t_{\text{yr}}=1}^n f_{t_{\text{yr}}}(x). \quad (12)$$

where  $f_{t_{\text{yr}}}$  is the unknown and numerically approximated probability density function of the windspeed corresponding to the yearly time index  $t_{\text{yr}}$ . In fact, the definition in (12) has also been used to model regional returns of extremes where  $f(x; \ell_1, \dots, \ell_n)$  varies by location  $\ell_i$  instead of time [4].

### ***3.2. A Simulation to Estimate Maximum Windspeed Risk Along the US North Atlantic Coast***

We run a simulation using the adaptations described in earlier sections to estimate high maximum windspeed risk for specified regions along the US North Atlantic coast.

From the Wind Field Model described in (1), we observe that returns of low central pressure minima have a large and direct effect on returns of high maximum windspeeds. This relationship makes appropriately modeling central pressure minima vital when considering returns of extreme windspeeds along the coast. However, it is not enough to know the central pressure minima to estimate coastal windspeed risk. This is because maximum windspeeds for a coastal region depend, among other things, on the translational velocity of the hurricane, the location at which landfall occurs, and whether the central pressure minima is achieved at landfall.

We now consider a more complex hurricane simulation to estimate the unknown distribution described in (12) of maximum windspeeds for a particular coastal location with the adaptations described in this investigation. The simulation is outlined in the Appendix A; however, we refer the reader to the original literature [5] for a detailed description. In essence, the process described in the Appendix A simulates a series of hurricane events for a given year by sampling the number of events to occur and the random variables used in the Wind Field Model represented by (1) at each time  $t$  along a simulated hurricane track. Once all hurricane events for a set of years have been simulated, we sample the windspeed for each simulated hurricane landing along

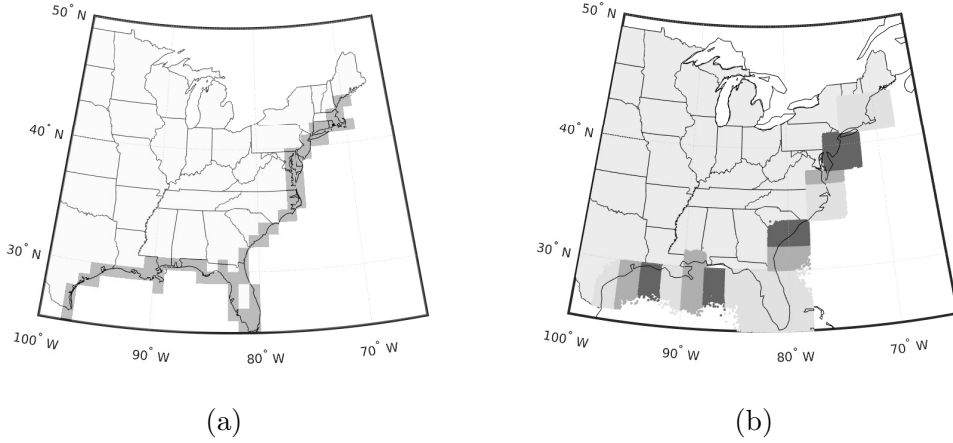


Figure 7.: (a) Coarse grid representing the coastal line. (b) Simulated hurricane locations along the coastal line. Different regions are indicated in gray-scale.

a specified coastline to form the unknown distribution described in (12). The North Atlantic coastline is first approximated by a coarse grid, illustrated in figure 7, then divided into coastal regions: N-Texas, S-Texas, W-Louisiana, E-Louisiana, Mississippi, Alabama-Florida, Florida, Florida-Georgia, South Carolina, North Carolina, Virginia, Maryland-New Jersey, and Connecticut-Massachusetts-New Hampshire. A simulated hurricane is said to be "on the coast" if the eye of the hurricane is within 2 degrees of the coastal line.

To estimate the  $n$ -year return-levels for regions along the coast, we must numerically approximate the probability distribution function of windspeeds described in (12). Then the  $n$ -year return-level is simply the  $1 - \frac{1}{n}$  quantile of the combined frequency distribution of maximum windspeed data for each coastal region. We do this for 20-, 30- and 50-year return-levels for each region taken along the coast by generating 20, 30 and 50 years of data (that is, 2020-2040, 2020-2050, and 2020-2070) for  $N = 1,000$  trials and estimating the 0.95, 0.97, and 0.98 quantiles, respectively.

It is reasonable to assume that each likelihood parameter in our simulation of maximum windspeeds (there are several),  $\theta$ , has reached its asymptotic normal distribution  $\mathcal{N}(\hat{\theta}, s_\theta)$  with mean  $\hat{\theta}$  equal to the maximum likelihood estimate of the parameter  $\theta$  and standard deviation given by the standard error  $s_\theta$  approximated from the Hessian. We can be confident that the true population distribution of maximum windspeeds, which the model is meant to represent, falls within some combination of these parameters; each coming from their corresponding distribution  $\mathcal{N}(\hat{\theta}, s_\theta)$ .

To estimate the confidence intervals of maximum windspeed return-levels, we independently sample from each of the parameter distributions to obtain 100 different combinations of parameters. We then run the simulation with each set of parameters for 2,000 (e.g. 20 years and  $N = 100$  trials), 3,000 and 5,000 years of hurricane simulations and estimate the 20-, 30- and 50- year return-levels. Given that each of these simulations is independent, we are left with a sequence of quantile estimates (return-levels) coming from an i.i.d. sequence of maximum windspeeds for each coastal region. It show in [7] that quantiles coming from i.i.d. sequences can be well-approximated by a normal distribution. Confidence intervals of each return-level are then estimated by

assuming an underlying normal distribution so that,

$$CI_{0.95} = \left[ F_{\text{norm}}^{-1}(0.025) \frac{\sigma_q}{\sqrt{100}}, F_{\text{norm}}^{-1}(0.975) \frac{\sigma_q}{\sqrt{100}} \right] \quad (13)$$

where  $F_{\text{norm}}^{-1}$  is the inverse standard normal distribution and  $\sigma_q$  is the estimated standard deviation of the 100 quantiles obtained from the 100 different parameter combinations. Quantiles to estimate 20-, 30-, and 50-year return-levels of maximum windspeed for each coastal region and their 95% estimated confidence intervals can be found in figure 8.

#### 4. Discussion

The Wind Field Model [11] has provided a convenient way of calculating the maximum windspeed of a hurricane event at any given time along a track, provided the central pressure is known. According to this model, high maximum windspeeds are obtained for low central pressure measurements. It is shown in [5] that this relationship can be used as a guide for estimating returns of extremely high maximum windspeeds along the coast by appropriately modeling the pressure minima. They found using the HURDAT database that central pressure can be modeled by the generalized extreme value distribution with stationary location and scale parameters depending on the lifetime and latitude of the central pressure minima. The simulation results of [5] using a stationary model of central pressure minima are in good agreement with the other models and analyses of the decade [1, 2]. However, our investigation shows that this stationary model does not appropriately fit the central pressure minima of the updated HURDAT2 database. These poor fits can be almost entirely blamed on a time-dependent component of the scale and location parameters in the model.

We have proposed a new, nonstationary model that accounts for this observed time-dependence in the location and scale parameter of the central pressure minima. Our model shows very reasonable fits to the true central pressure minima. Following a standard approach, we assume a Poisson distribution for yearly returns of hurricane events; however, we show that this model is also time-dependent with an exponentially increasing rate parameter for hurricane events with lifetimes greater than 6.25 days. We discuss this against current literature where stationarity of hurricane returns is assumed. We show that our models can reliably predict up to at least 50-year returns for the central pressure minima without the need for updating by comparing the generated training set model against a test set and standardizing the central pressure minima using extreme value methods.

We have used our nonstationary model of central pressure minima and Poisson returns for yearly hurricanes in a more complex simulation to estimate 20-, 30-, and 50-year return-levels of maximum windspeeds for sections along the US North Atlantic coastline. Compared to other analyses of maximum windspeed returns for coastal regions based on the HURDAT database like those in [1, 5, 6], our model has two significant results: 1.) higher maximum windspeeds are expected for all regions along the US North Atlantic Coast and 2.) the highest maximum windspeeds occur in the northern part of the coast.

Specifically for landfalling hurricanes, we find a scale parameter for negative central pressure minima that is linearly increasing with time which suggests an expectation for higher-highs and lower-lows of central pressure minima. This phenomenon coupled

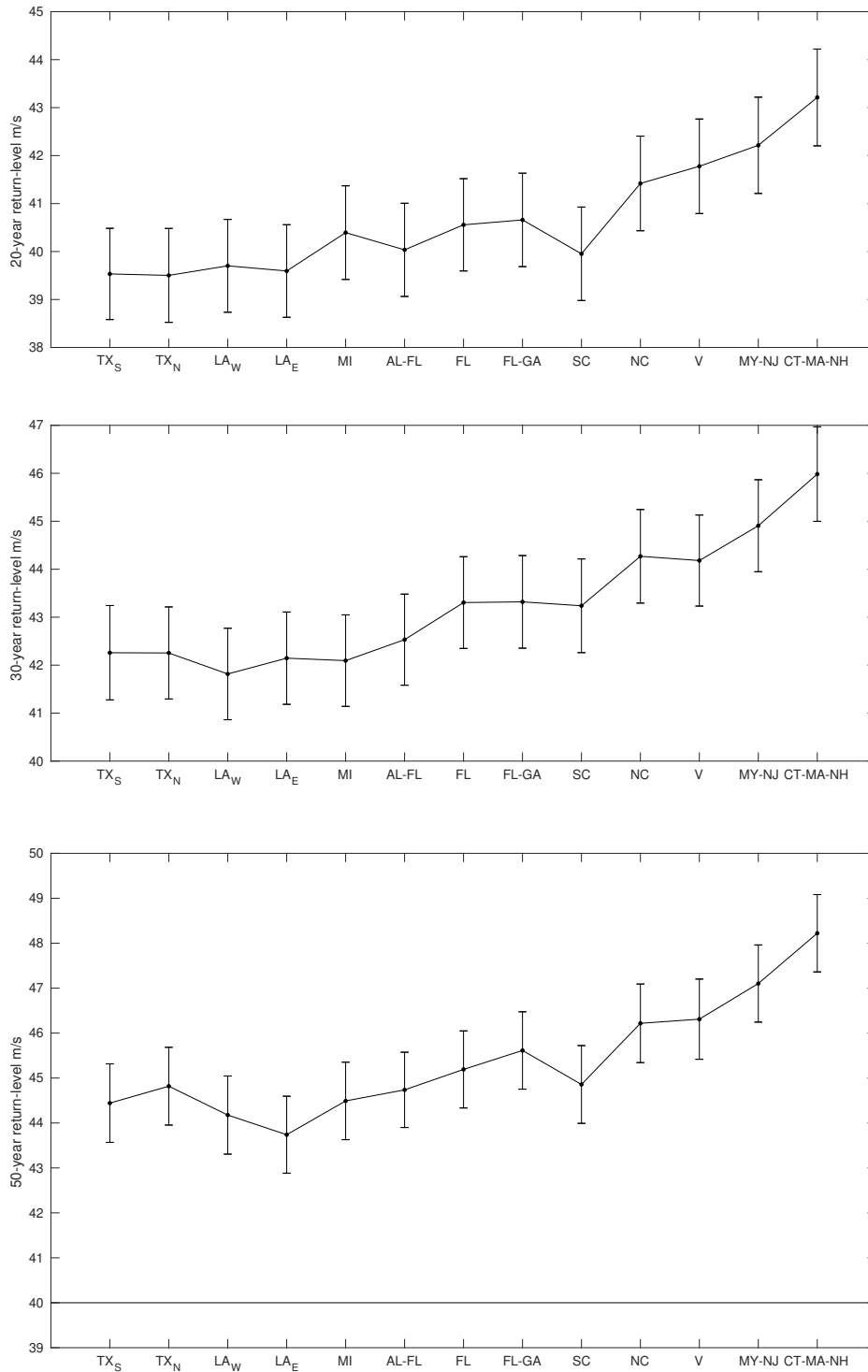


Figure 8.: Plots of the 20-, 30-, and 50- year return-levels of maximum windspeed along the coast **estimated for year 2021**. Central estimates are the quantiles of the distribution of 6-hourly windspeeds for  $N = 1,000$  trials of 20, 30, and 50 years of simulated hurricanes, respectively. Errorbars represent the 95% confidence interval estimates from eq. 13. Solid horizontal line indicates the maximum windspeed return estimated from stationary models of previous literature.

with a general increase in the observed number of hurricane events can certainly lead to higher maximum windspeeds everywhere along the coast.

An increase in the maximum windspeed for higher latitudes is actually a nontrivial observation because return-levels are affected by the number of hurricanes observed in a coastal region. In general, the number of observed hurricanes in the north tend to be lower. For example, it is well-known that the coastal region around Florida has many more hurricane events than those regions along the north-eastern coast. In fact, we find this to be true in our simulations as well. So, one would expect to have a higher 20-year maximum wind-speed return-level for the Florida region than the northeast. On the other hand, translational velocity plays a critical role in the maximum windspeed of a hurricane hitting the coastal region where translational velocity is always greater for higher latitudes.

**Remark 2.** The uneven latitudinal distribution of hurricane tracks is accounted for in the simulation by the use of historical tracks with some small added noise.

We have tested our model to determine the cause of this northern increase in maximum windspeed and have found that translational velocity has the greatest influence over the observed trend. Furthermore, our simulated values of average translational velocity, estimated from the simulated hurricane tracks, almost identically follow those found in the literature [19]. We refer the reader to figure 9 for a plot of translational velocity over latitude. This result provides us with reasonable confidence in our model for coastal risk analysis.

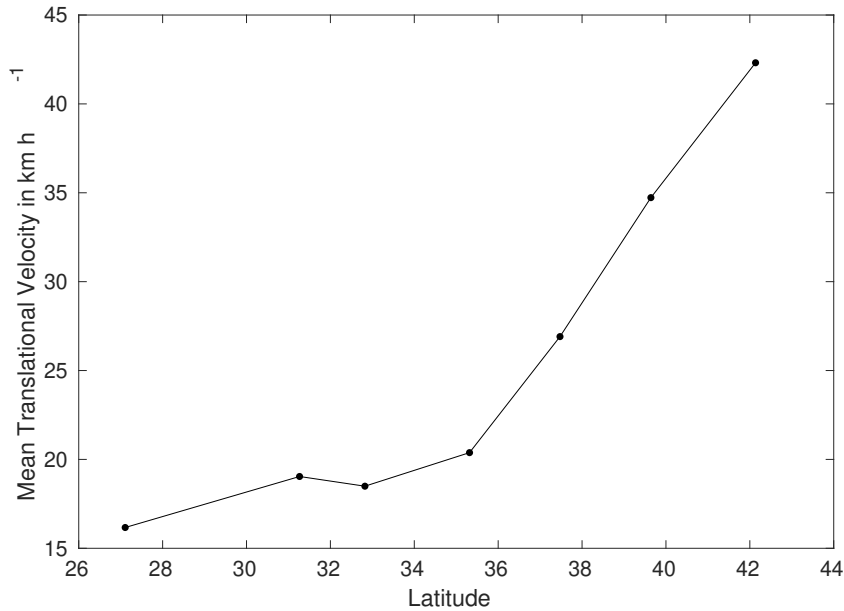


Figure 9.: Average translational speed for simulated hurricanes from our model along the coast plotted against latitude.



## **Code and Data Availability**

The historical central pressure and track data are freely available on the NOAA website: <https://www.nhc.noaa.gov/data/hurdat/hurdat2-1851-2020-052921.txt> [9]. The code used to perform the nonstationary investigation, model fits, and wind-speed simulation was written in MATLAB. For access to this code, please contact the corresponding author.

## **Acknowledgements**

MC would like to thank MPIPES for their hospitality where part of this work was completed. MN would like to thank the NSF for support on NSF-DMS Grants 1600780 and 2009923.

## References

- [1] M. E. Batts, M. R. Cordes, L. R. Russell, J. R. Shaver, E. Simiu. Hurricane wind speeds in the United States. *Report BSS 124*. National Bureau of Standards, US Department of Commerce, Washington DC. (1980).
- [2] F. P. Ho, J. C. Su, K. L. Smith, F. P. Richards. Hurricane climatology for the Atlantic and Gulf Coasts of the United States. *Technical Report NWS 38*. National Weather Service, National Oceanic and Atmospheric Administration, US Department of Commerce, Silver Spring. (1987).
- [3] M. Carney, R. Azencott and M. Nicol. Nonstationarity of Summer Temperature Extremes in Texas. *International Journal of Climatology*. (2019) **40** 1, 620-640.
- [4] M. Carney, H. Kantz. Robust Regional Clustering and Modeling of Nonstationary Summer Temperature Extremes across Germany. *Advances in Statistical Climatology, Meteorology, and Oceanography*. (2020) **6** 1, 61-77.
- [5] E. Casson and S. Coles. Simulation and Extremal Analysis of Hurricane Events. *Appl. Statist.* (2000) **49**, 227-245.
- [6] S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer (2001).
- [7] K. Knight. What are the Limiting Distributions of Quantile Estimators?. In: Dodge Y. (eds) *Statistical Data Analysis Based on the L1-Norm and Related Methods*. Statistics for Industry and Technology. Birkhäuser, Basel. (2002) 47-65.
- [8] C. W. Landsea, G. A. Vecchi, L. Bengtsson, and T. R. Knutson. Impact of Duration Thresholds on Atlantic Tropical Cyclone Counts. *Journal of Climate*. (2010) 2508-2519.
- [9] C. W. Landsea and J. L. Franklin. Atlantic Hurricane Database Uncertainty and Presentation of a New Database Format. *Mon. Wea. Rev.* (2013) 141, 3576-3592.
- [10] M. R. Leadbetter, G. Lindgren and H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. Springer Verlag New York Inc. (1983).
- [11] National Oceanic and Atmospheric Administration. *HUR 7-120* revised Standard Project Hurricane criteria for the Atlantic and Gulf Coasts of the United States. Technical Report. National Weather Service, National Oceanic and Atmospheric Administration, US Department of Commerce, Silver Spring. (1972).
- [12] E. Simiu, N. A. Heckert, and T. Whalen. Estimates of hurricane wind speeds by the 'Peaks Over Threshold' method. *Technical Report 1416*. National Institute of Standards and Technology, Gaithersburg. (1995).
- [13] V. Lucarini, D. Faranda, A.C. Freitas, J.M. Freitas, M. P. Holland, T. Kuna, M. Nicol, M. Todd, S. Vaienti, *Extremes and Recurrence in Dynamical Systems*, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs, and Tracts. (2016).
- [14] G. A. Vecchi and T.R. Knutson. On Estimates of Historical North Atlantic Tropical Cyclone Activity. *Journal of Climate*. (2007) 3580-3600.
- [15] G. A. Vecchi and T. R. Knutson. Estimating Annual Numbers of Atlantic Hurricanes Missing from the HURDAT Database (1878-1965) Using Ship Track Density. *Journal of Climate*. (2010) 1736-1745.
- [16] G. A. Vecchi, C. Landsea, W. Zhang, G. Villarini and T. R. Knutson. Changes in Atlantic Major Hurricane Frequency Since the Late-19th Century. Preprint: Research Square. (2021).
- [17] P.J. Vickery and L. A. Twisdale. Prediction of Hurricane Windspeeds in the US. *Journal of Structural Engineering*. (1995). **121** 11, 1691-1699.
- [18] D. Cooley. Return Periods and Return Levels Under Climate Change. In: AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., Sorooshian, S. *Extremes in a Changing Climate. Detection, Analysis and Uncertainty*. (2013) 97-113.
- [19] M. Yamaguchi, J. Chan, I-J Moon, K. Yoshida, R. Mizuta. Global warming changes tropical cyclone translation speed. *Nature Communications*. **11** 47 (2020).

## Appendix A. Numerical Simulation of Maximum Windspeeds for Hurricane Events

A summary of the scheme for simulating the timeseries of maximum windspeeds of a yearly sample of hurricane events is as follows. Adaptations from this investigation are marked with a \*. Sampling tracks or timeseries refers to sampling from the 642 historical hurricane records from the HURDAT2 database. Densities and probabilities are estimated from the 642 historical hurricane records.

- Sample the number of hurricane events to occur for a specified year from (6) with rate parameter (7) \*.
- For each hurricane event:
  - Simulate a hurricane track by uniformly sampling a historical track and adding a small amount of spatial noise at each step sampled from  $\mathcal{N}(0, \sigma^2)$  where  $\sigma \simeq 100$  n.mi.
  - Uniformly sample a central pressure timeseries of a nonlandfalling hurricane.
  - Time scale the central pressure timeseries to match the lifetime of the simulated track.
  - For simulated landfalling hurricanes, run a Bernoulli trial with  $p$  probability that the time of occurrence of the central pressure minimum occurring is at landfall,  $t_{p_{\min}} = t_{\text{lf}}$ , and  $1 - p$  otherwise.  $p$  is estimated from the database.
    - If occurring before landfall, randomly sample the time of occurrence for central pressure minimum using the density of the ratio  $t_{p_{\min}}/t_{\text{lf}}$  in the Appendix figure C2.
  - Randomly sample the central pressure minimum from the nonstationary model described by (4) and (5) \*.
  - Randomly sample the central pressure range using the density model in [5, Section 2.5] with fits in Table B1.
  - Add landfall effects described in [5, Section 2.6] separately for inland and coastal hurricanes.
  - Simulate the radius to maximum windspeeds  $R_{\max}(\phi_t)$  by finding the distribution with coastal latitude  $\phi_t$  illustrated in figure C1.
  - Use the simulated central pressure timeseries  $p_t$  and  $R_{\max}(\phi_t)$  as inputs into the Wind Field Model described in (1).

## Appendix B. Additional Tables

Type	$a(\text{se})$	$b(\text{se})$	$c(\text{se})$
Landfalling	872.33(25.48)	-0.87(0.03)	11.77(0.32)
Nonlandfalling	829.50(19.77)	-0.82(0.02)	7.47(0.19)

Table B1.: Maximum likelihood estimates of the parameters in the normal distributional model from [5, Section 2.5] for  $p_{\text{range}}$ .

## Appendix C. Additional Figures

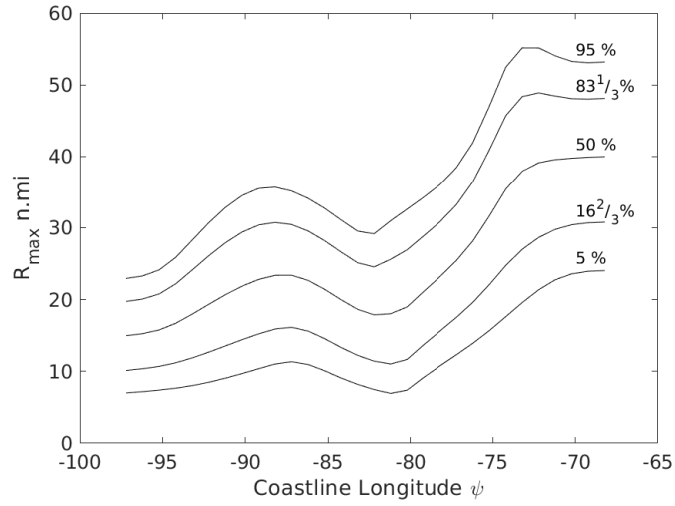


Figure C1.: Observed quantiles of  $R_{\max}$  as a function of the U.S. coastline longitude  $\psi$  reproduced from fig. 37 and 38 of [2] and compared to fig. 7 of [5].  $R_{\max}$  increases as a function of the latitude as seen in the figure (increases and decreases along the U.S. coastline defined longitude). The model for  $R_{\max}$  is created using the longitude  $\psi$  while sampling is performed using the latitude  $\phi$  because  $R_{\max}$  is unique along  $\phi$ .

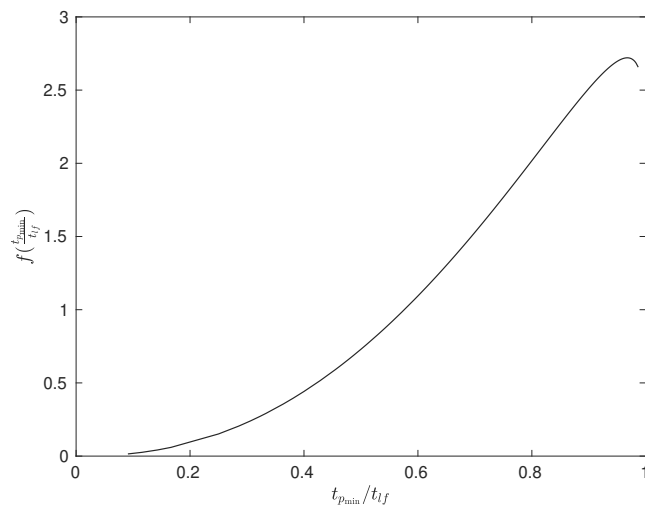
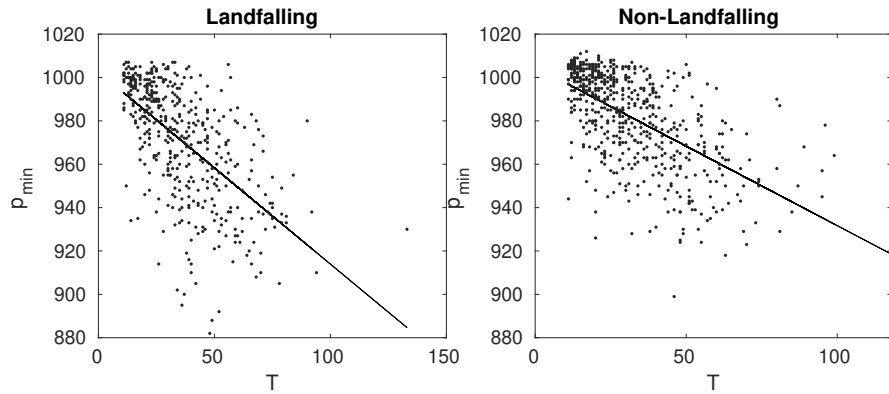
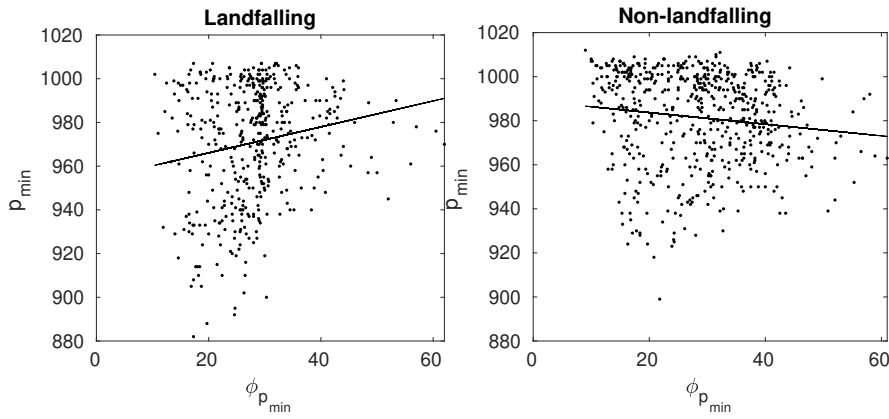


Figure C2.: Empirical density function of the ratio  $t_{p_{min}}/t_{lf}$  for landfalling hurricanes that make landfall after their central pressure minima occurs.



(a)



(b)

Figure C3.: (a) Scatter plot of  $p_{\min}$  against  $T$  and (b)  $\phi_{t_{p_{\min}}}$  for landfalling and non-landfalling hurricanes with a line of best fit. Results indicate the the stationary model including dependence on  $T$  and  $\phi_{t_{p_{\min}}}$  is a reasonable starting point for forming the nonstationary model.

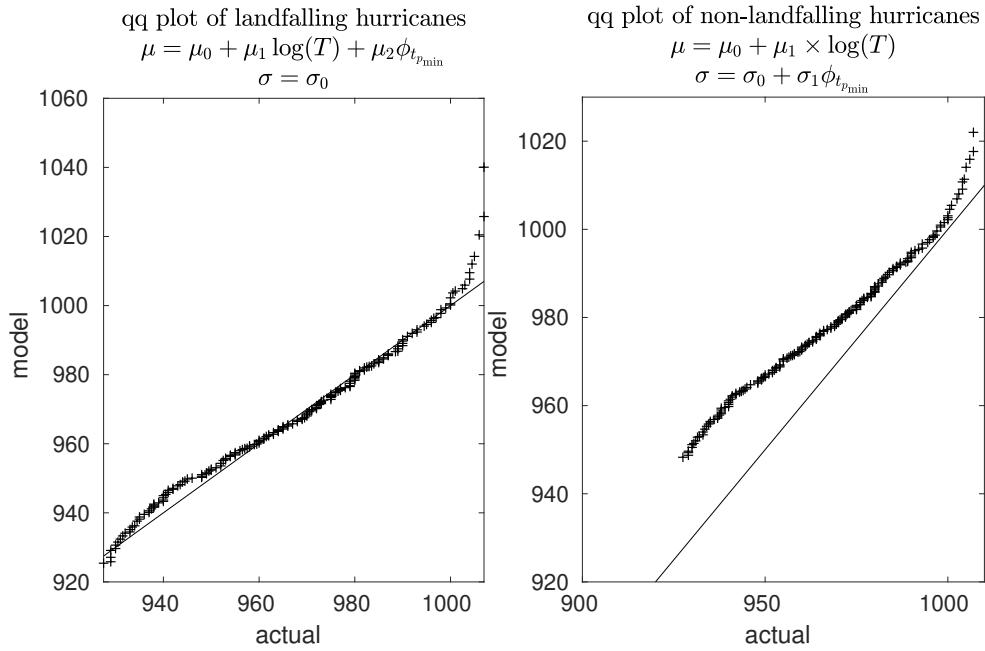


Figure C4.: Quantile plots for landfalling and nonlandfalling hurricanes comparing the stationary model from [5] on the HURDAT2 database.

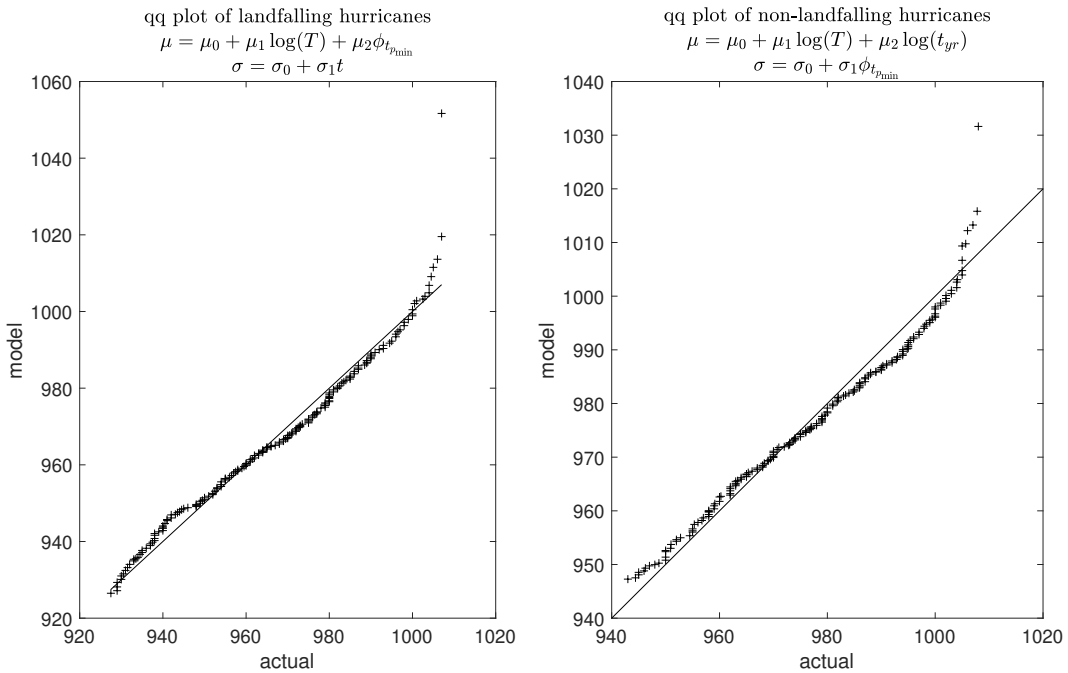


Figure C5.: Quantile plots for landfalling and nonlandfalling hurricanes comparing the proposed nonstationary GEV model to data from the HURDAT2 database.



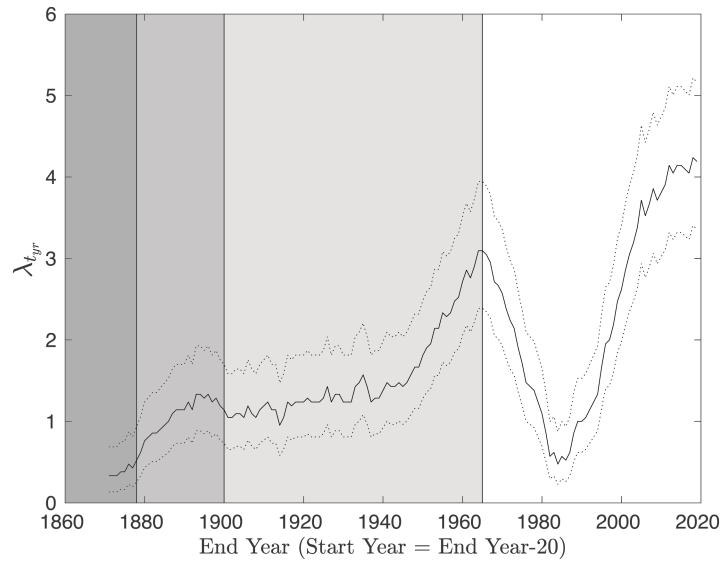


Figure C6.: Likelihood estimate of the Poisson parameter for yearly landfalling hurricane event rates with lifetimes greater than 6.25 days. Estimates are taken over 20 year moving time windows. Standard errors are marked with dotted lines. Fitted exponential model is represented by a thick line. Grayed areas correspond to those in [16]: (1) 1878 - year when the U.S. Signal Corps began cataloging all Atlantic hurricanes (2) 1900 - year when the U.S. Coast was sufficiently well-populated for monitoring (3) modern-era with appropriate ship density.

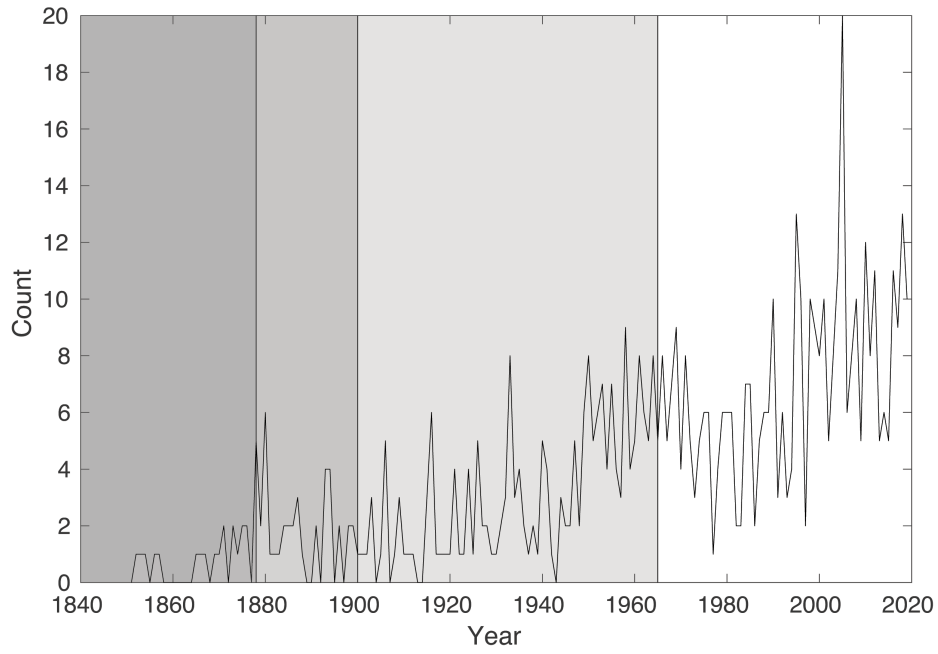


Figure C7.: Number of total hurricane events of lifetimes greater than 6.25 days by year. Grayed areas correspond to those in [16]: (1) 1878 - year when the U.S. Signal Corps began cataloging all Atlantic hurricanes (2) 1900 - year when the U.S. Coast was sufficiently well-populated for monitoring (3) modern-era with appropriate ship density.