

## Early Season Hurricane Risk Assessment: Climate-Conditioned HITS Simulation of North Atlantic Tropical Storm Tracks

JENNIFER NAKAMURA,<sup>a</sup> UPMANU LALL,<sup>b</sup> YOCHANAN KUSHNIR,<sup>a</sup> PATRICK A. HARR,<sup>c</sup> AND KYRA MCCREERY<sup>d</sup>

<sup>a</sup> *Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York*

<sup>b</sup> *Department of Earth and Environmental Engineering, Columbia University, New York, New York*

<sup>c</sup> *Jupiter Intelligence, San Mateo, California*

<sup>d</sup> *North Shore High School, Glen Head, New York*

(Manuscript received 15 October 2020, in final form 11 March 2021)

**ABSTRACT:** We present a hurricane risk assessment model that simulates North Atlantic Ocean tropical cyclone (TC) tracks and intensity, conditioned on the early season large-scale climate state. The model, Cluster-Based Climate-Conditioned Hurricane Intensity and Track Simulator ( $C^3$ -HITS), extends a previous version of HITS. HITS is a nonparametric, spatial semi-Markov, stochastic model that generates TC tracks by conditionally simulating segments of randomly varying lengths from the TC tracks contained in NOAA's Best Track Data, version 2, dataset. The distance to neighboring tracks, track direction, TC wind speed, and age are used as conditioning variables.  $C^3$ -HITS adds conditioning on two early season, large-scale climate covariates to condition the track simulation: the Niño-3.4 index, representing the eastern equatorial Pacific Ocean sea surface temperature (SST) departure from climatology, and main development region, representing tropical North Atlantic SST departure from climatology in the North Atlantic TC main development region. A track clustering procedure is used to identify track families, and a Poisson regression model is used to model the probabilistic number of storms formed in each cluster, conditional on the two climate covariates. The HITS algorithm is then applied to evolve these tracks forward in time. The output of this two-step, climate-conditioned simulator is compared with an unconditional HITS application to illustrate its prognostic efficacy in simulating tracks during the subsequent season. As in the HITS model, each track retains information on velocity and other attributes that can be used for predictive coastal risk modeling for the upcoming TC season.

**KEYWORDS:** Atmosphere; North Atlantic Ocean; Hurricanes/typhoons; Stochastic models; Clustering

### 1. Introduction

Tropical cyclones (TCs) in the North Atlantic Ocean basin have immense impacts on the inhabitants, economies, and environments of Central and North American coastal and inland locations. Human, infrastructure, and societal impacts of TCs far exceed the duration of landfall and can last for years, as seen recently in the case of Hurricanes Sandy (Comes and Van de Walle 2014; Sobel 2014), Harvey (Sebastian et al. 2017; Fan et al. 2018), and Maria (Meléndez and Hinojosa 2017; Zorrilla 2017). The development of statistical and dynamical models that assess the basinwide level of TC activity, landfall potential, and complex relationships among North Atlantic TC activity and various regional and global climatic factors has been an area of concern for researchers over the past several decades (Villarini et al. 2010, 2012; Vimont and Kossin 2007; Tang and Neelin 2004; Mestre and Hallegatte 2009; Kossin et al. 2010; Hall and Yonekura 2013).

A TC track reflects an interaction of the storm with its ambient thermodynamic and dynamic conditions. In turn, these conditions are affected by direct and indirect, short- and long-term factors that modulate atmospheric and oceanic conditions to determine patterns of TC activity (Elsner and Kara 1999). Seasonal large-scale and background climatic

conditions examined in the recent literature in this context include the atmospheric influence of El Niño–Southern Oscillation (ENSO; Klotzbach 2010; Hall and Yonekura 2013) and the North Atlantic Oscillation (Elsner and Jagger 2004). Also important are sea surface temperature anomalies (SSTA) in the tropical North Atlantic Ocean, in the TC main development region (MDR) (Shapiro and Goldenberg 1998; Emanuel 2005; Webster et al. 2005) or by the related, so-called Atlantic meridional mode (AMM; Vimont and Kossin 2007).

Numerical, dynamical climate models (Goerss 2000; Vitart et al. 2007; Knutson et al. 2013) have been used to model North Atlantic TC activity (occurrence, landfall rates, precipitation) and their sensitivity to changes in baseline climate conditions, as part of the evolution of the global climate. However, these models tend to be of low resolution and are expensive to run repeatedly as part of achieving a probabilistic risk assessment (Vitart et al. 1997; Camargo et al. 2005). Climate-conditioned statistical models have thus been proposed to predict the properties of an upcoming hurricane season (Elsner and Jagger 2006; Hall and Jewson 2008; Hall and Yonekura 2013; Klotzbach et al. 2017). Elsner and Jagger (2006) use May–June climate indices to predict July–November basinwide counts of North Atlantic hurricane counts using a Bayesian approach. Hall and Jewson (2008) and Hall and Yonekura (2013) create synthetic tracks under different climate conditions with propagation conditioned on sea surface temperature.

*Corresponding author:* Jennifer Nakamura, jennie@ldeo.columbia.com

DOI: 10.1175/JAMC-D-20-0237.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

The Colorado State University has issued an annual hurricane outlook since 1984 that uses a variety of environmental parameters that have evolved over the years and been shown to influence seasonal TC activity at varying lead times (Klotzbach et al. 2017). Seasonal landfall prediction as a function of background and large-scale climate conditions for the North Atlantic has been explored by Poisson regression (Elsner and Schmertmann 1993; Elsner et al. 2001; Elsner 2003; Sabbatelli and Mann 2007; Kossin et al. 2010).

This paper extends our stochastic model for North Atlantic TC risk assessment—the previously developed Hurricane Intensity and Track Simulator (HITS; see Nakamura et al. 2015). Here, we extend HITS to incorporate the dependence of North Atlantic TC activity on early season large-scale climate conditions, thereby providing dynamic predictive risk estimation through simulated ensembles of hurricane tracks. HITS is a machine-learning model developed using the historical TC tracks maintained by the National Oceanic and Atmospheric Administration (NOAA; Landsea and Franklin 2013). HITS simulates new stochastic tracks using a  $k$ -nearest-neighbor-based nonparametric, spatial semi-Markov model that can preserve the long-range persistence of TC tracks better than the traditional Markov models, as well as models that seek to locally parameterize the stochastic track movement. By considering spatially local information related to track attributes for conditioning and stochastic simulation of random track segments for the track conditionally selected, HITS can exploit information at the current track location as well as the potential track continuity or memory implied in the historical track data. HITS was intended to produce a stochastic catalog of storms for probabilistic risk assessment that uses historical tracks as a basis for extracting spatiotemporal TC evolution properties.

HITS has the advantage of carrying all of the track information such as sea level pressure at the storm center and wind speed information provided in the NOAA archive, and rainfall fields available at landfall from other sources. Here we introduce the Cluster-Based Climate-Conditioned Hurricane Intensity and Track Simulator ( $C^3$ -HITS), which conditions the initial segment of a hurricane track on early season, large-scale climate conditions or covariates, before the simulation of the track by HITS. For now, we consider that ocean temperature conditions inform hurricane formation locations and frequency and that steering of the TC is still dependent on the distance to, the direction of, and other attributes of neighboring storms.  $C^3$ -HITS uses a cluster analysis of historical tracks to sort tracks into groups based on similar track geometry and other spatial properties. A Poisson regression is then used to determine the expected number of TCs formed in each cluster as a function of the large-scale climate covariates. The formation locations are selected randomly from historical storms associated with each cluster. HITS is then used to evolve each randomly generated storm. A catalog of stochastic storms from which the risk profile for the upcoming hurricane season is thus developed.

The TC dataset used is described in the next section along with the climate indices chosen to represent the remote and local underlying seasonal background conditions. A description of

$C^3$ -HITS is provided next in the methods section. A description of the simulations and a comparison with observations and the unconditioned HITS output are given in the results section. The paper concludes with a summary and thoughts about future modifications of the model.

## 2. Data

### a. HURDAT2

We use the recently revised, North Atlantic basin historical hurricane track data (HURDAT2), available from NOAA (<https://oasishub.co/dataset/hurdat-2-atlantic-hurricane-database>). Tracks are provided for each TC season, from 1851 to 2016. Each track constitutes sequential storm locations at fixed time intervals (originally every 6 h), with information on the time and date, the storm position (latitude/longitude), the maximum wind speed, and central pressure. HURDAT2 improves on the original HURDAT by the inclusion of non-synoptic track times for landfalls and maximum intensity locations and the inclusion of nondeveloping tropical depressions. The original 6-h dataset was linearly interpolated to every 2 h for clustering purposes. Because of the relatively small sample size of TC tracks, the full HURDAT2 data were used in the creation of clusters.

### b. Climate indices

To specify the underlying large-scale seasonal climate conditions, we chose indices that specify the major phenomena governing North Atlantic TC activity as determined by earlier studies. The North Atlantic TC MDR (10°–20°N, 80°–20°W) is regarded as a prime hurricane formation location for North Atlantic landfalling TC (Shapiro and Goldenberg 1998; Goldenberg et al. 2001). Increased SSTAs in the MDR typically contribute to an increase in North Atlantic TC activity (Vecchi et al. 2013). The tropical Pacific SSTA in the Niño-3.4 region, (5°S–5°N, 170°–120°W), is typically utilized as an indicator of ENSO events. El Niño events (defined when the anomaly exceeds +0.5°F for at least 6 months and the atmosphere shows a response) typically suppresses hurricane activity in the North Atlantic basin (Klotzbach 2010). The Niño-1+2 index is another ENSO indicator defined in the eastern Pacific (0°–10°S, 90°–80°W). SSTAs in a region of the Gulf of Mexico (GoM: 20°–35°N, 100°–80°W) are associated with the formation of several major North Atlantic landfalling TCs. Increased sea surface temperatures in the Gulf of Mexico are associated with an increase in North Atlantic TC frequency, but most researchers contend that ocean–atmosphere interactions in this region depend on a multitude of factors (Vecchi et al. 2013). The AMM is a large-scale climate pattern that encompasses the southern Atlantic Ocean and is defined by a shift in the intertropical convergence zone caused by fluctuations in SSTs and easterly trade winds (Chiang and Vimont 2004). The AMM has been linked to ENSO and other climate oscillations such as the North Atlantic Oscillation. On inter-annual time scales, the positive phase of AMM is correlated with a higher frequency of North Atlantic named storms (Kossin et al. 2010).

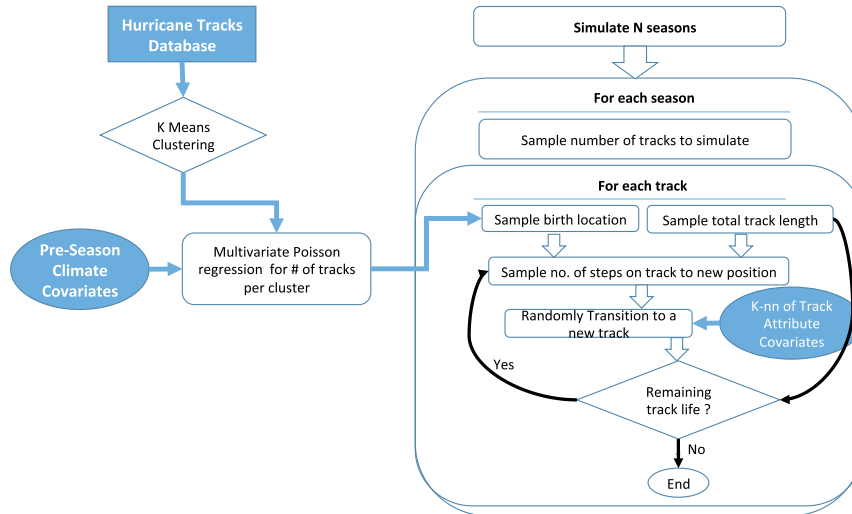


FIG. 1. Flowchart of the climate-conditioned North Atlantic TC track simulator. North Atlantic TC tracks from 1851 to 2016 are  $k$ -means clustered. The multivariate Poisson regression determines the mean number of initial track segments. The number of TC formed in each cluster is sampled from the Poisson distribution. The corresponding TC formation locations are then randomly sampled from the TCs assigned to each cluster for the historical 1950–2016 tracks. The final track length is sampled randomly from the track lengths of the  $k$  nearest neighbors within a  $2.5^\circ$  radius of the formation location. The number of steps taken along a track is randomly sampled. If track length remains, a jump to a  $k$ -nearest-neighbor track within a  $2.5^\circ$  radius is based on transition probabilities across all tracks in the 1851–2016 database.

All of these climate indices were computed from the Kaplan extended, version 2, SST data, and a combination of Kaplan et al. (1998) and Reynolds and Smith (1994) SSTA with respect to 1950–2016 except for the AMM (which was obtained online from <http://www.aos.wisc.edu/~dvimont/MModes/Data.html>). All indices were calculated for April–June, preceding the nominal North Atlantic TC season. Although the official North Atlantic hurricane season starts 1 June, very few tropical cyclones form in June, so for the purpose of the paper, pre-season is defined as being prior to 1 July. Data for 1950–2016 were used for the indices to be concurrent with the computed AMM index and the NCEP reanalysis data in appendix A.

### 3. Methods

The  $C^3$ -HITS algorithm consists of three major steps: (i) clustering of the historical TC tracks; (ii) a nonhomogeneous Poisson process model for the probabilities of the number of TCs formed in each cluster, conditional on pre-season climate covariates; and (iii) simulating the full track evolution as in the original HITS algorithm. The conceptual structure of our model is depicted schematically in Fig. 1, which summarizes the analysis steps detailed below.

#### a. TC cluster identification

We first classify the historical TCs using a cluster analysis procedure based on their full track and movement information. The clustering follows our prior work (Nakamura et al. 2009, hereafter N09), where we classified North Atlantic TC tracks

by their attributes, namely, formation location, duration (or life), distance covered (or track length), and the actual orientation and curvature (or spatial geometry) of tracks. In HURDAT2 (as in its predecessor HURDAT) each track is defined by a collection of TC locations separated by a fixed time interval. We use this sequence of locations to define the first and second spatial moments of the track's trajectory. Thus, each track is described by five parameters: the location (latitude and longitude) of the track centroid, the variance of track HURDAT2 2-hourly locations in longitude and latitude with respect to the centroid, and the cross covariance between them. These summarize the shape of the track. Tracks consisting of a single observation were not considered and track locations eastward of  $0^\circ$  longitude were truncated.

A  $k$ -means algorithm with the number of clusters selected using “silhouette analysis” was used. The details of the resulting clusters are provided in appendix A. Using HURDAT2 we identified five clusters, which are depicted in Fig. 2. In N09, using the older HURDAT data, six clusters had been identified. The five clusters differ from one another by formation location, length, and orientation of the track and by proximity of the track to the U.S. coast.

Cluster 1 (the west-MDR cluster; Fig. 2a) and cluster 2 (the subtropical cluster; Fig. 2b) TCs form in the western MDR and north of  $20^\circ\text{N}$ , with a larger proportion of cluster 2 storms forming in the western flank of the North Atlantic subtropical anticyclone (Fig. A3 of appendix A). Many cluster-2 storms form close to the U.S. coast. Storms from clusters 1 and 2 can make landfall along the U.S. East Coast and GoM coastlines.

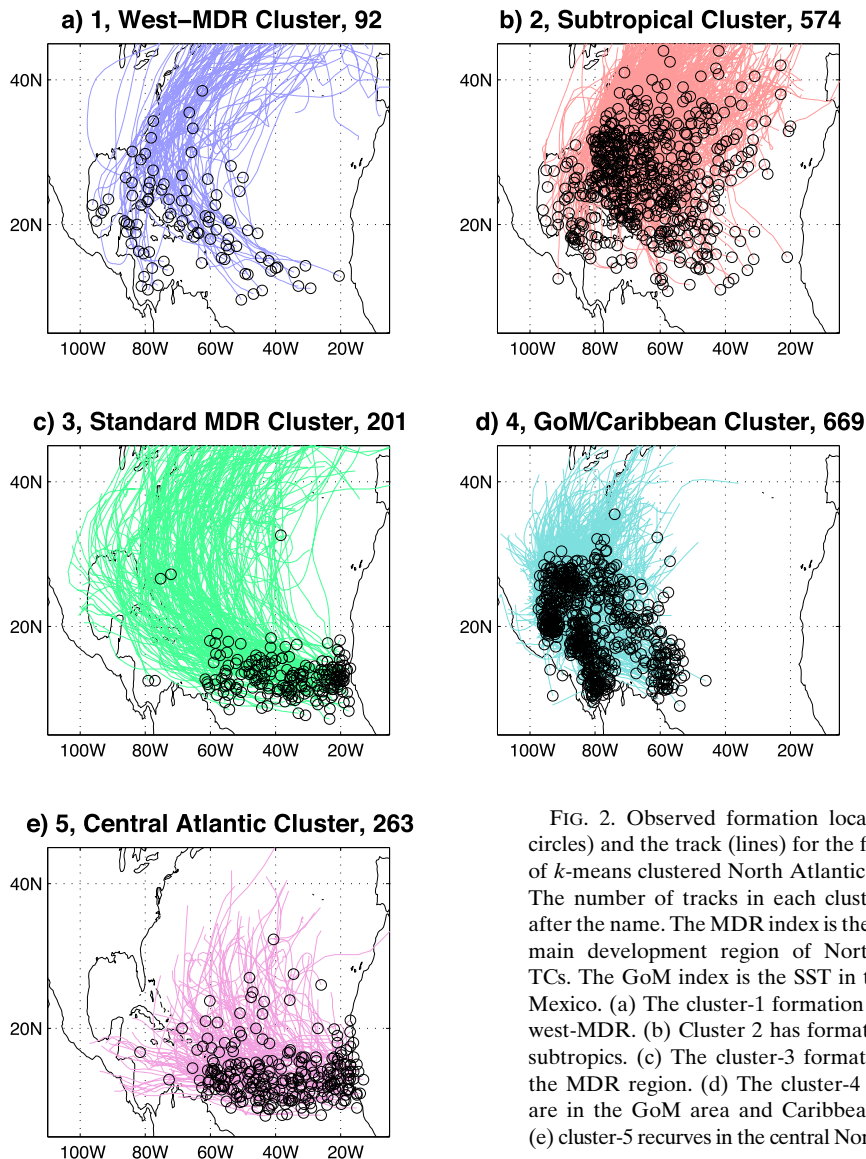


FIG. 2. Observed formation location (open circles) and the track (lines) for the five clusters of  $k$ -means clustered North Atlantic TC tracks. The number of tracks in each cluster is listed after the name. The MDR index is the SST in the main development region of North Atlantic TCs. The GoM index is the SST in the Gulf of Mexico. (a) The cluster-1 formation area is the west-MDR. (b) Cluster 2 has formations in the subtropics. (c) The cluster-3 formations are in the MDR region. (d) The cluster-4 formations are in the GoM area and Caribbean Sea, and (e) cluster-5 recurves in the central North Atlantic.

Here landfall is defined as crossing one of many  $2.5^\circ$  contiguous gates created along the Gulf and East Coasts of the United States up to  $42^\circ\text{N}$ . These gates are meridionally oriented on the East Coast and zonally oriented on the Gulf Coast and are counted separately. Cluster-1 tracks account for 3% of both East and Gulf Coast landfalls; cluster 2 accounts for 21% of East Coast landfalls and 12% of Gulf Coast landfalls.

Cluster 3 (Fig. 2c), which we refer to as the standard MDR cluster, contains storms that form in the MDR. The TCs in this cluster are often referred to as Cape Verde hurricanes because of their formation location, and they move in parabola-shaped tracks. The Standard MDR cluster is associated with an intensification of the subtropical high (Fig. A3). The tracks are the longest of any cluster, and they make up 21% of East Coast landfalls and 9% of Gulf Coast landfalls.

Cluster-4 storms (Fig. 2d; GoM/Caribbean cluster) typically have short tracks but form close to the Caribbean islands or in

the Caribbean Sea and the GoM and move northward around a local anomalous low pressure system (Fig. A3). With their proximity to land, cluster-4 tracks make landfall more than any other cluster, accounting for 55% of the East Coast landfalls and 76% of the Gulf Coast landfalls.

Cluster-5 storms (Fig. 2e, central Atlantic cluster) predominantly drift northward. They can reach the Caribbean islands but rarely make landfall in the United States or in Central America.

#### b. Climate state-based initialization

To initialize the simulation of climate-state dependent tracks, we hypothesize that the TCs that tend to occur in a given storm season originate in a preferred set of clusters determined by the pre-season state of the large-scale climate as represented by the value of key climate covariates. From these preferred clusters, a number of formation locations with their

corresponding initial track segments are drawn according to a probability that is set by a multivariate nonhomogeneous Poisson process, whose rate vector (the mean number of TCs expected from that cluster) is a function of the covariates. We applied linear and nonlinear nonparametric local regression (Loader 2006) Poisson regression models with and without overdispersion (Coxe et al. 2009). Linear dependence models with no overdispersion were selected using the Akaike information criterion (AIC) and the Schwarz Bayesian information criterion (BIC). Both AIC and BIC are often used to estimate the accuracy of the fit of the regression model by providing a quantitative indication of the balance between the goodness of the fit and the number of parameters in the model. The regression model that results in the minimization of AIC or BIC is considered to be superior. In summary, we considered 1) linear or nonlinear regression, 2) sensitivity to AIC or BIC for model selection, 3) whether overdispersion is present, and 4) whether a zero-inflated Poisson process improves the model.

Five SSTA-based indices were considered as predictors (MDR, Niño-1+2, Niño-3.4, GoM, and AMM). For the best subset of regression models, only the MDR and Niño-3.4 indices were retained most often. Nonlinearity, overdispersion, and zero inflation were indicated as significantly better only in a small percentage of the experiments, and hence just a general linear model (glm) with the best subset regression (bestglm in R) and BIC as the selection criteria was finally used.

The model can be represented as

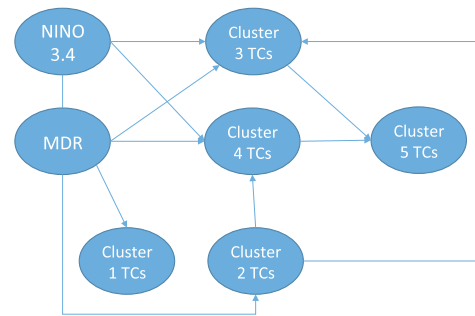
$$f(y_{it}|\mathbf{x}_t) \sim \text{Poisson}(\lambda_{it}), \quad \text{with} \quad (1)$$

$$\lambda_{it} = \mathbf{x}_t \boldsymbol{\beta}_i, \quad (2)$$

where  $y_{it}$  is the number of TCs that form in cluster  $i$  in year  $t$ ,  $\lambda_{it}$  is the corresponding mean of the nonhomogeneous Poisson process,  $\mathbf{x}_t$  is a vector of predictors for year  $t$ ,  $\boldsymbol{\beta}_i$  is a vector of regression coefficients for cluster  $i$ , and  $f(\cdot)$  refers to a conditional probability distribution.

The Poisson process for the cluster-dependent number of TCs each season is also correlated across the clusters. To address this correlation, two approaches were explored: (i) a multivariate regression on the climate covariates; and (ii) a sequential model for cluster Poisson rates that iteratively considers dependence of the Poisson rate for some clusters on the climate covariates, and for other clusters, on the rates for the clusters initially modeled with climate covariates, and possibly a climate covariate. The latter strategy was found to perform better. It recognizes the MDR and Niño-3.4 SSTA climate covariates were not always statistically significant predictors for the Poisson rate for all clusters, but that the rate for some of the clusters was well predicted by the rate for clusters that did have a strong dependence on the climate covariates. Thus, a reduced network representation of the multivariate Poisson process is sought.

An example of the directed acyclic graph (DAG) or causal dependence network associated with the process is provided in Fig. 3. Initial candidate DAGs were selected based on exploratory Poisson regressions for each cluster TC considering the climate covariate and other cluster TCs as potential predictors.



$$f(C1, C2, C3, C4, C5, \text{MDR}, \text{NINO3.4}) = f(C5 | C3, C4) * f(C3 | C2, \text{MDR}, \text{NINO3.4}) * f(C4 | C2, \text{MDR}, \text{NINO3.4}) * f(C2 | \text{MDR}) * f(C1 | \text{MDR}) * f(\text{MDR}, \text{NINO3.4})$$

FIG. 3. The DAG or causal dependence network for the North Atlantic TCs, shown symbolically, with the corresponding factorization of the joint distribution of the variables indicated below. All arrows indicate a directed conditional dependence. The line between MDR and Niño-3.4 represents the joint distribution (no causal direction) of the two variables. The values of both of these pre-season variables are known in the pre-season and are not simulated from a distribution. All other variables are sequentially simulated using Poisson regression models for each conditional distribution as indicated by the DAG, accounting for the parameter uncertainty distribution for each model.

The final causal dependence network selected was based on the best performance from candidate networks based on the aggregate cross-validated likelihood of the network under leave one season out at a time cross validation of the Poisson regressions. A description of the Poisson regression procedure, including a table presenting the regression coefficients corresponding to the links presented in the DAG (Fig. 3) is given in appendix B.

Once the rate parameters  $\lambda_{it}$  and the final predictors  $\mathbf{x}_t$  (climate covariates Niño-3.4 and MDR as well as the TC formation in clusters) are selected consistent with the DAG in Fig. 3, we sequentially simulate the number of TC formed in each cluster from the associated conditional distributions as indicated in the DAG. This allows us to maintain the dependence on the covariates and the mutual dependence in the Poisson process across the clusters.

Given the  $y_{it}$  values simulated, the corresponding number of TC formation locations are selected by randomly drawing the corresponding number of tracks from the subset of tracks associated with that cluster in the historical 1950–2016 dataset. Given these initial TC formation locations and associated track segments, the HITS model as detailed in Nakamura et al. (2015) is used to simulate each track to completion. The key parameters are 1) the total track length, 2) the number of steps to take along a track (analogous to a renewal process or spell length), and 3) track similarity as measured by the distance between tracks, the wind speed, direction-vector of the tracks, and the age of the tracks. The total track length is simulated randomly for each track from the track lengths of the  $k$ -nearest-neighbor tracks. The number of steps to take along a track is sampled randomly between 1 and the remaining track length at each state, with uniform probability.

At each transition point, one of the tracks that is a  $k$  nearest neighbor in terms of similarity is selected to transition to using a probability kernel based on distance as described in Lall and Sharma (1996), and the process is repeated until the track length simulated is reached. Only  $k$  nearest neighbors that are within a  $2.5^\circ$  radius of the current track location are considered across all tracks in the 1851–2016 database. Thus, while the clusters determine the formation locations, subsequent track evolution relies on the full TC track data.

A single simulation of a hurricane season is completed after the number of tracks sampled from the DAG model across all clusters has been simulated. The total number of simulations for a given season may be as large as the user's desired reliability of the estimation of the probability of TC attributes, for example, landfalls in a specific area. In the current version of C<sup>3</sup>-HITS, we do not condition the evolution of the tracks beyond the first segment on the climate covariates.

#### 4. Results and discussion

Using the Poisson model, we simulated the overall number of storms in the entire North Atlantic basin, year by year for 1950–2016, based on each year's pre-season, values of the climate variates. The results of 1000 such simulations are shown in Figs. 4a–f, separated by cluster and for all clusters together. In each panel, the area spanned by the range between the 5th and 95th percentiles of the Poisson model simulations is colored in light blue. The thick blue lines depict the model median, and the red dots represent the observed number of formations in each cluster. The Kendall's tau correlations of the model medians and observed cluster TC formations per season are shown in the upper-left-hand corner of each subplot. Significant correlations at the  $\alpha$  0.05 level are given in boldface font. Most observed TC formed per cluster fall within the 5th–95th-percentile envelope. The correlation between observed and model median for clusters 1, 3, 4, and "all" (Figs. 4a,c,d,f) is significant. The relationship between predictions and observed for clusters 2 and 5 (Figs. 4b,e) is weaker and not significant. The predictive ability of the DAG varies by cluster (appendix B).

Figure 5 contains a number of plots addressing the quality of the climate-conditioned model simulation of the total number of storms in the basin, as a function of the season's Niño-3.4 and MDR index values in comparison with observations. Figure 5a is calculated from the observations and the C<sup>3</sup>-HITS model simulation is shown in Fig. 5b. Also shown is a comparison between the observed number of storms and those generated by HITS and C<sup>3</sup>-HITS models, using the metric of RMS difference. For C<sup>3</sup>-HITS (in Fig. 5d), the median of the simulated 1000 values each season is used in the comparison. For HITS (Fig. 5c), where the simulation is not dependent of the climate variates, the median of a single set of 1000 simulations is used as a reference. The difference between HITS simulation errors and the simulation errors of C<sup>3</sup>-HITS is presented in Fig. 5e, the black dot indicates the MDR and Niño-3.4 index values in 2018. In Fig. 5f, the difference between the simulation error variances of HITS and C<sup>3</sup>-HITS is tested using a parametric  $F$  test.

In observations (Fig. 5a), North Atlantic basin totals are low when the Niño-3.4 index is positive and high when the Niño-3.4 index is negative or the MDR index is positive. This statement is true broadly, but the distribution shows pockets of lower or higher TC activity per year, which may be an actual variation or the outcome of sampling issues due to a small number of data points in these cases. Open black circles in Fig. 5a indicate where the data points are. Consistent with its design, the C<sup>3</sup>-HITS simulations (Fig. 5b) yield a smooth and robust relationship as to the sign and magnitude of the response to Niño-3.4 and MDR indices.

The RMSE differences with observations in both the HITS (Fig. 5c) and the C<sup>3</sup>-HITS (Fig. 5d) runs are relatively low overall, but they are visibly lower in the latter case. This is seen in the RMSE control—RMSE climate (Fig. 5e), where a smaller value indicates a smaller error in the climate case. Thus, the warm (yellow-red) portion of the difference plot (Fig. 5e) indicates the C<sup>3</sup>-HITS simulation was closer to observations than the nonconditioned HITS simulation. Overall, adding climate information reduces the RMSE error of the modeled basin total formations per season and adds more value to the model in the warm Niño-3.4 index and the warm MDR index cases. The region between the two warm extremes at the plot edges is a region of transition between increased TC activity in the MDR index and less TC activity in the Niño-3.4 index. This area has a higher error in both control and climate runs as shown in Figs. 5c and 5d, which may be due to a low number of observations in these cases.

Saunders et al. (2020) discussed the 25% above-normal 2018 North Atlantic TC total basin count and found only a 5% chance of occurrence with knowledge of the pre-season environmental fields. They state that the above-normal season was not forecasted because of the lack of skill in statistical seasonal forecast models of North Atlantic basin total. The year 2018 was run as an independent realization as the Poisson regression models used in C<sup>3</sup>-HITS were computed based on data for the years 1950–2016. We used the MDR and Niño-3.4 2018 pre-season values in the Poisson regression models to determine basin cluster totals (illustrated in Fig. 1). The 2018 black dot in Fig. 5e is in the warm (red) portion of the MDR and Niño-3.4 index space, which indicates skill for including climate information. The 2018 observed total of named TC activity in the North Atlantic basin was 15. The run for 2018 gave a mean overall TC count of 14.18, a median of 13, and a 75th percentile of 19. This is a slightly lower total for the mean (0.82 named TC), but the observed total falls between the 50th and 75th percentiles. The Colorado State 2018 statistical seasonal forecast of named storms in the North Atlantic basin using several predictors issued 31 May 2018 was 13. Including the pre-season Subtropical Storm Alberto, this gave 14, near the mean of the Poisson regression model ([https://tropical.colostate.edu/Forecast/Archived\\_Forecasts/2010s/2018-06.pdf](https://tropical.colostate.edu/Forecast/Archived_Forecasts/2010s/2018-06.pdf)). With knowledge of pre-season MDR and Niño-3.4 indices, the Poisson regression models used in C<sup>3</sup>-HITS thus exhibits visible skill in the 2018 total basin named TC count.

An  $F$  test is used to test the equality of two variances. The variances compared in this case are both error variances with respect to observed total basin formations per season. At a 5%

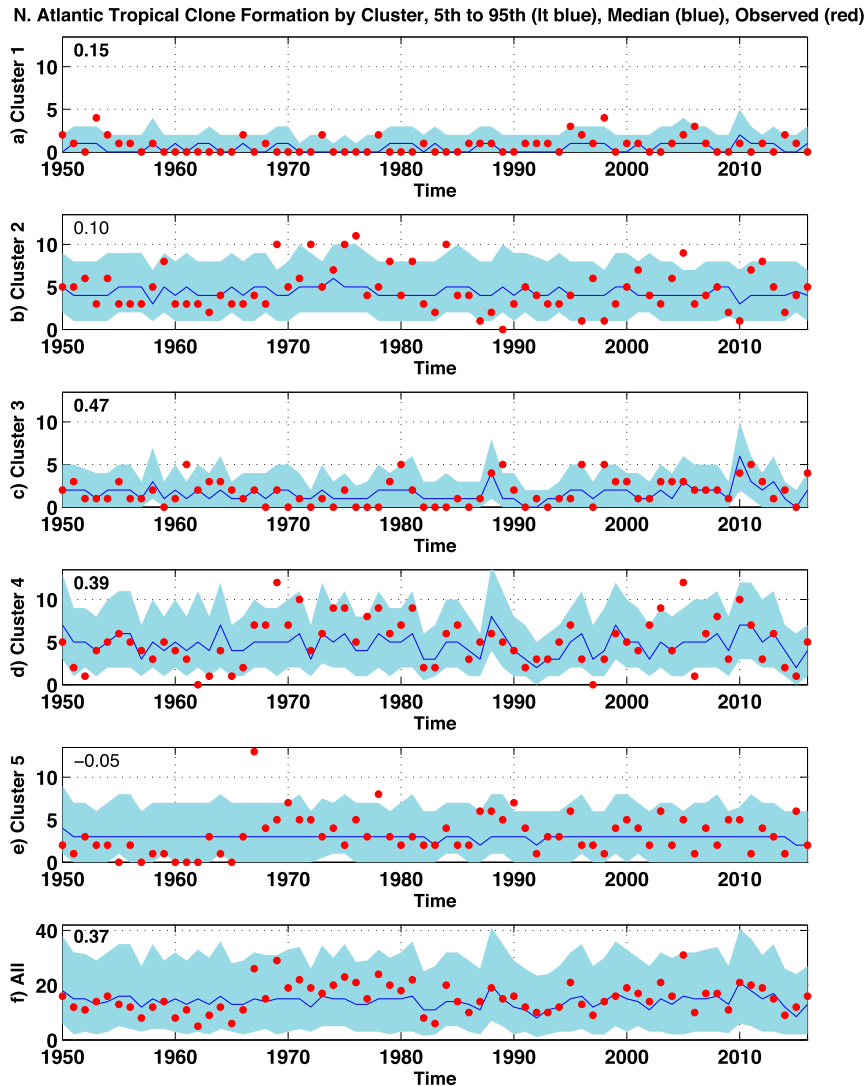


FIG. 4. A comparison of the historical number of TC formations (red dots) in each  $k$ -means cluster of North Atlantic TC tracks and the 5th and 95th percentile of a Poisson regression model of basin formations per season. The  $r$  value in the upper-left corners is the correlation coefficient between historical and model median basin TC formations per season by cluster membership.

significance level, the  $F$  test is 1 when it rejects the null hypothesis of the same variance and 0 when it cannot reject. Of 67 years (1950–2016), 65 are 1 or 0.97 and 2 of 67 years are 0 or 0.03. In an overwhelming number (97%) of years, HITS and  $C^3$ -HITS have significantly different variances (Fig. 5f). These results show that the new model performs significantly better in predicting the seasonal number of storms than the original HITS with no climate conditioning.

Figure 6 is the same as Fig. 5, but for U.S. mainland landfall of North Atlantic TCs (on the East Coast and Gulf coastlines). Local minimum and maximum pockets are even more prevalent in the observed U.S. landfall (Fig. 6a) than basin total (Fig. 5a), which due to a lower amount of data could be a result of sampling errors. The median of 1000 runs of the  $C^3$ -HITS is a

function of both the linear Poisson regressions and the transition probabilities generated by the model and that determine the landfall probability. The RMSE for the HITS and  $C^3$ -HITS are low overall relative to the basin total formations per season (Figs. 5c,d) and higher in the landfall case (Figs. 6c,d). HITS transition probabilities were originally selected for success in representing mainland U.S. landfall correctly.

The 2018 MDR and Niño-3.4 index value dot in Fig. 6e is in a skill area (red) for  $C^3$ -HITS. The most populous cluster when running the Poisson regression models for  $C^3$ -HITS in 2018 was cluster 2, the subtropical cluster (5.02 of 14.18 total named TC). Cluster 2 has landfalls on the East Coast and is the cluster to which Hurricane Florence, which made landfall in 2018, would have belonged. The next most popular cluster in the 2018 run

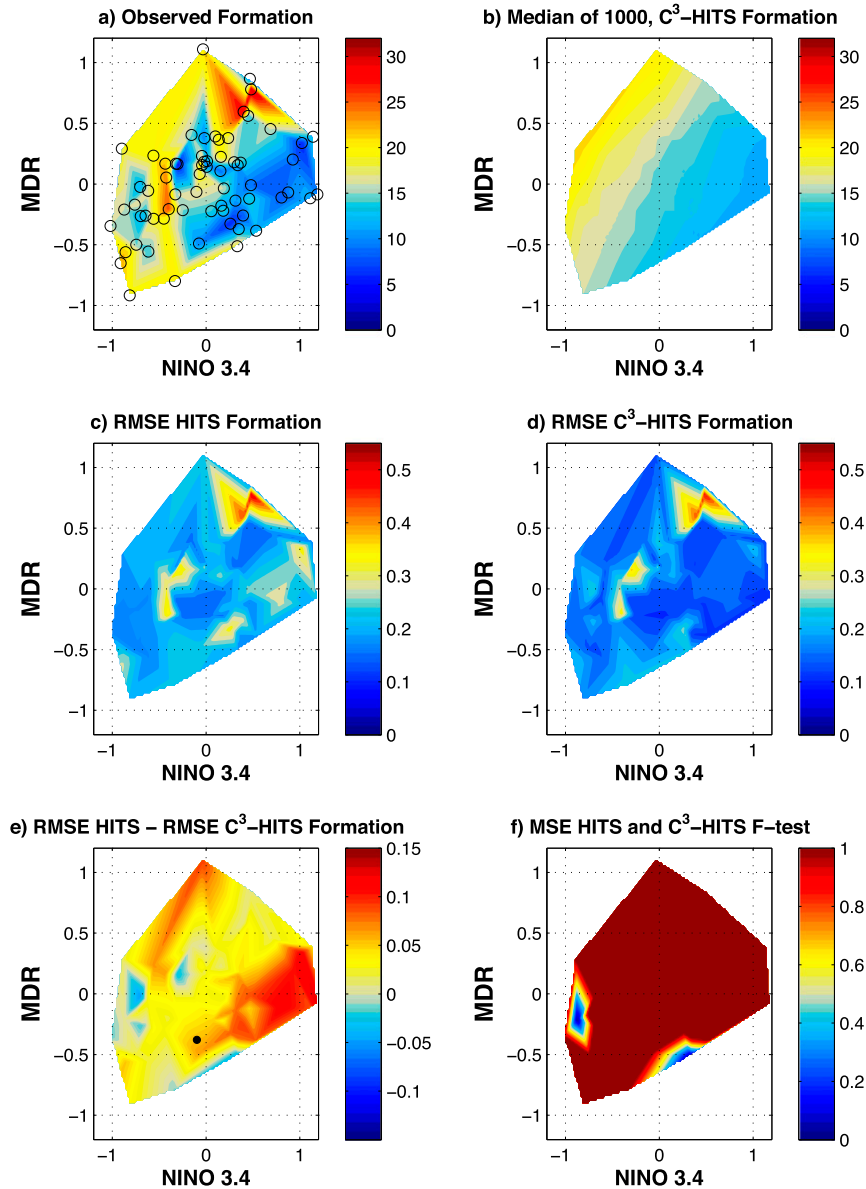


FIG. 5. The  $x$  axis is the Niño-3.4 index, the SST in a region of the tropical Pacific Ocean. The  $y$  axis is the MDR index, the SST in the main development region of TCs in the North Atlantic Ocean. (a) The observed North Atlantic basin formation total in each year is contoured. Open black circles indicate data points. (b) Contours of the  $C^3$ -HITS model North Atlantic Ocean basin formation total per year; the median of 1000 model runs. (c) The RMSE of the HITS model and observed TC basin formation total in each year. (d) The RMSE of  $C^3$ -HITS with the observed basin total. (e) The difference of (c) minus (d). The black dot is the Niño-3.4 and MDR value in 2018. (f) A significance test ( $F$  test) of the mean-square error of HITS and  $C^3$ -HITS compared with observed TC basin totals per year. The  $F$ -test results indicate whether the variances come from the same normal distributions. The  $F$  test is 1 if it rejects the same variance at a 5% significance level and is 0 if it cannot reject at a 5% significance level.

was cluster 4, the GoM/Caribbean cluster (4.38 of 14.18 total named TC). Cluster 4 has landfalls on the Gulf Coast and is the cluster to which another 2018 storm, Michael, would have belonged. The  $F$  tests (Fig. 6f) show less certainty in

the significance of the difference between HITS and  $C^3$ -HITS than the basin total formations per season with 75% of the years for landfall (50 of 67) showing a significant difference.



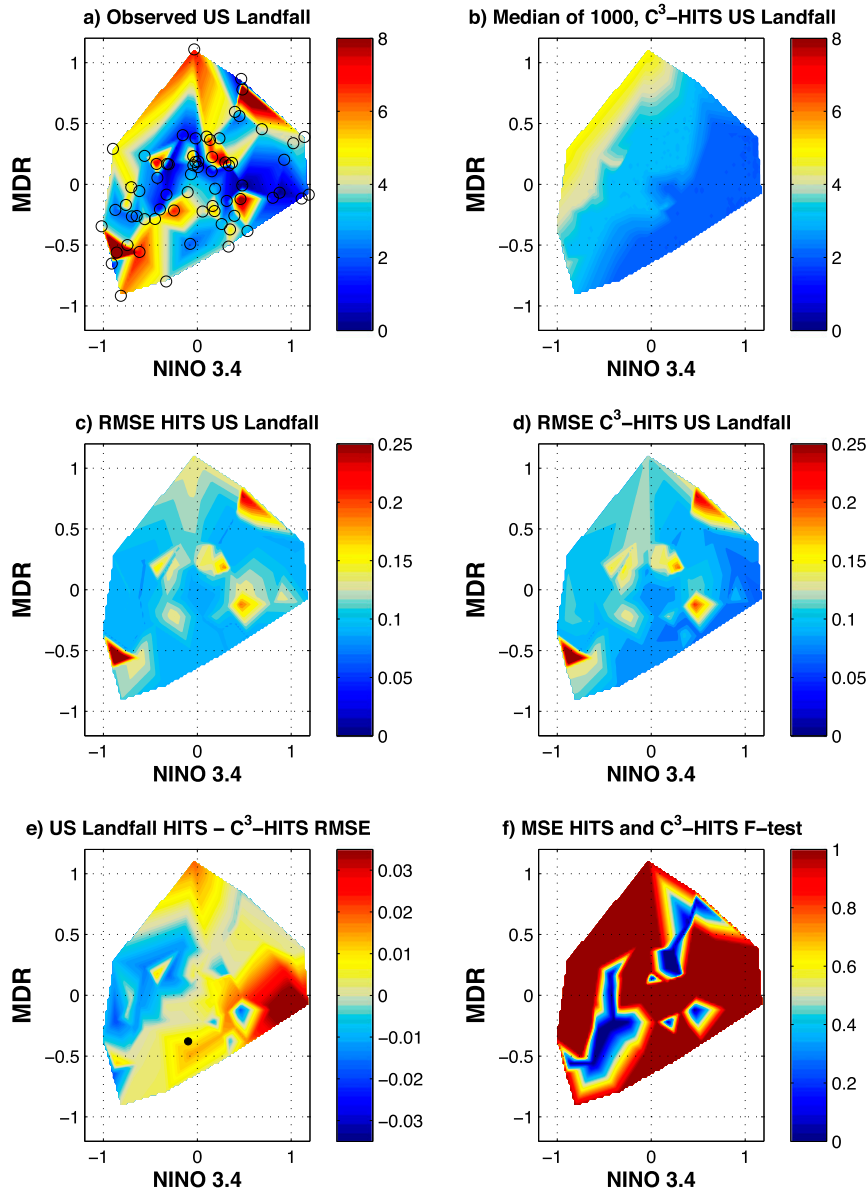


FIG. 6. As in Fig. 5, but for the count of TCs landfalling on the East and Gulf Coasts of the United States.

Post-air reconnaissance landfalls (1950–2016) as compared with the earlier period (1851–1949) were used for Fig. 6 for improved accuracy. However, large uncertainty is introduced into the root-mean-square error calculations because of the very small sample size with the order of magnitude of 1. The sample size of Fig. 5, basin total formations per season, is an order of magnitude 10, larger by an order of magnitude than landfalls (Fig. 6). This makes Fig. 5a more reliable estimate of the relative error of the two models.

The C<sup>3</sup>-HITS model gives the ability to examine multihazard probability, such as the situation of one or more TC landfalls per season by climate state. Figure 7 shows the probability in the climate index domain of Niño-3.4 versus MDR index for one

(Fig. 7a) and as many as six and up (Fig. 7f) TCs landfall per season in C<sup>3</sup>-HITS. The progression of increasing TC per season shows a decrease that starts in the Niño-3.4 warm and MDR cold to the neutral corner and works its way slowly over to Niño-3.4 cold and MDR warm until finally, in six and up (Fig. 7f) TC per season, there is just an edge of higher probability in the Niño-3.4 cold and MDR warm space. These panels are very similar to the diagonal pattern of the median of the C<sup>3</sup>-HITS model in Figs. 5b and 6b implying that when there are more storms in the basin the chances of multiple landfalls in a season increases.

To compare this model-based landfall assessment to actual observations of TC activity per season in the years 1950–2016, 1000 estimates were drawn from the observation-based

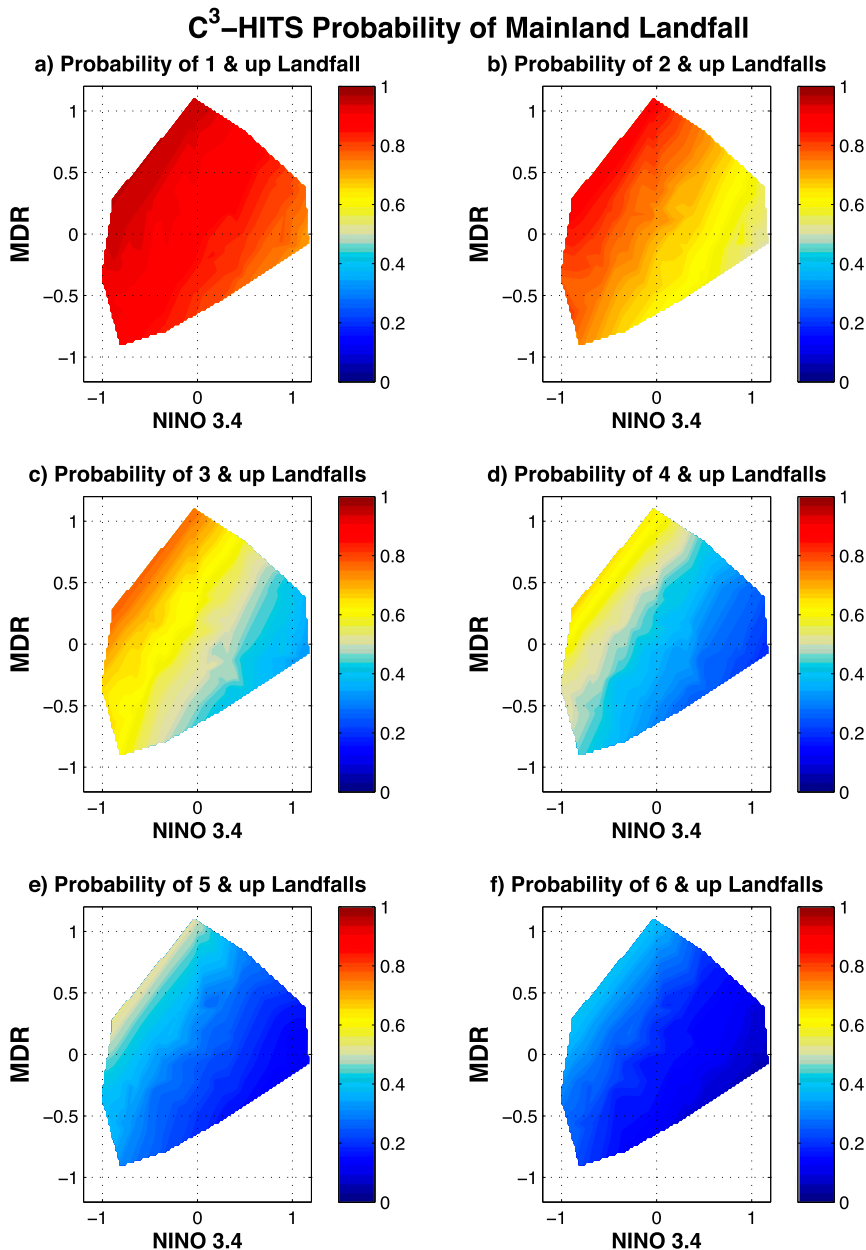


FIG. 7. The  $x$  axis is the Niño-3.4 index, the SST in a region of the tropical Pacific Ocean. The  $y$  axis is the MDR index, the SST in the main development region of TCs in the North Atlantic Ocean. Shown is the C<sup>3</sup>-HITS model with climate information: counts of TCs landfalling on the East and Gulf Coasts of the United States, over the 1000 runs for each year divided by 1000, giving the probability of landfall in each category as indicated by the sub-panel titles.

estimated Poisson distribution with overdispersion of these events. Overdispersion allows the fitted distribution to stray slightly from the exact theoretical values of the Poisson distribution. The results in Fig. 8 illustrate a similar diagonal pattern. This pattern is a match to C<sup>3</sup>-HITS probabilities in Fig. 7. It is not expected that observed and C<sup>3</sup>-HITS would be similar in this regard. A Poisson distribution of landfall was not imposed in the

model and the tracks disperse based on the HITS parameters. In all, C<sup>3</sup>-HITS provides TC tracks and their accompanying parameters along with landfalls that have a similar distribution to the observed TC tracks.

Quantitative comparisons of the probabilities of landfall or multiple landfalls in the model (Fig. 7) and observed (Fig. 8) at MDR and Niño-3.4 values per year are computed by Hellinger distances as defined in Cha [2007, Eq. (26)]:

**Observed Poisson Estimated Probability Mainland Landfall**

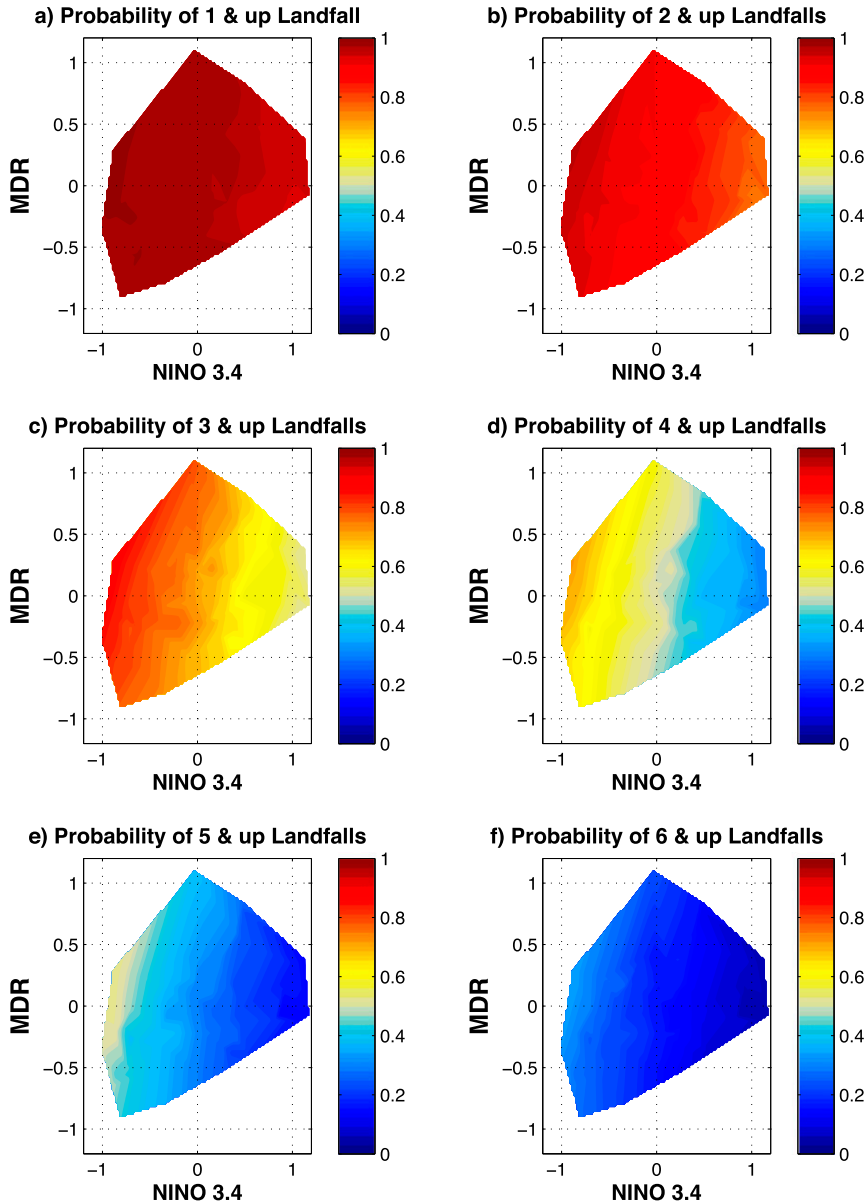


FIG. 8. As Fig. 7, but for observations rather than model results.

$$d_h = 2\sqrt{1 - \sum_{i=1}^n (P_i Q_i)^2}, \tag{3}$$

where  $d_h$  is the Hellinger distance,  $n$  is the number of years,  $P$  is the observed probability of landfall at year  $i$ , and  $Q$  is the C<sup>3</sup>-HITS average model probability over 1000 simulations at year  $i$ .

Values of the Hellinger distance indicate 0 for complete overlap and 1 for maximum separation. The distances between the model and observations for one and up TCs per year and six and up TCs per year are 0.15 (Figs. 7a,f and 8a,f), two and up are 0.33 (Figs. 7b and 8b), three and up are 0.43 (Figs. 7c and

8c), four and up are 0.41 (Figs. 7d and 8d), and five and up are 0.30 (Figs. 7e and 8e). Both one and up and six and up per year are the closest at 0.15—a very close match. Three and up per year is the farthest at 0.43 but still overlaps in more than half of the years.

The cluster-based climate-conditioned track simulator uses clusters to determine the formation location and the first segment of the simulated track, but what impact does this have on the resulting tracks? The resulting C<sup>3</sup>-HITS tracks were subjected to moments-based clustering, as were the observed tracks, and contrasted in Fig. 9 as a function of the cluster on which they began. If the answer to the question of the impact of

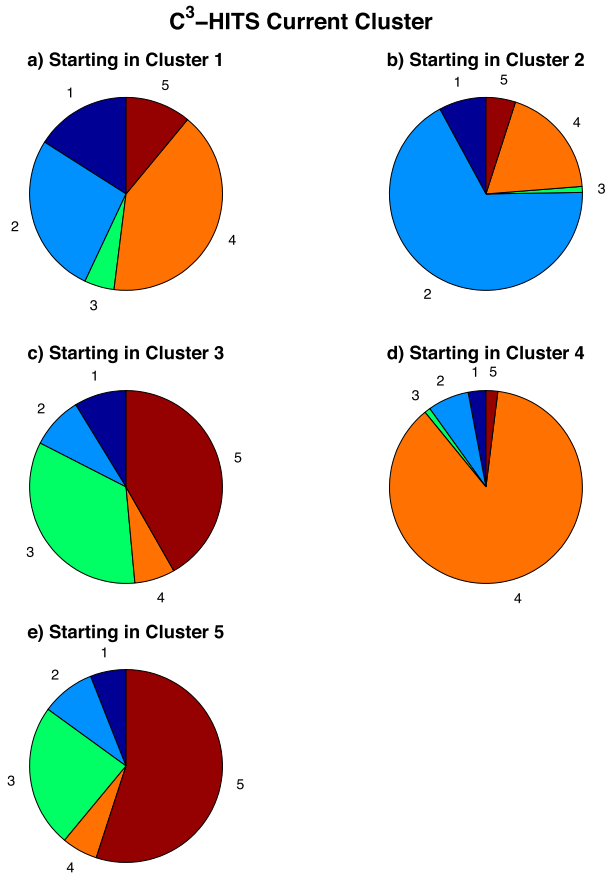


FIG. 9. C<sup>3</sup>-HITS model final cluster membership by the starting  $k$ -means cluster based on observed TC tracks.

resulting tracks were given by cluster 1, the west-MDR cluster, it would be that the starting cluster has little to do with the final one. However, cluster 1 is the least populated cluster in the observed data. The observed clusters 1851–2016, from most to least populated one, are cluster 4 GoM/Caribbean (669), cluster 2 subtropical cluster (574), cluster 5 central Atlantic (263), cluster 3 standard MDR (201), and cluster 1 west-MDR (92) (Fig. 3). In a small number of tracks (the population of cluster 1), 92/1799, or 5%, the starting cluster is not important. In the case of clusters 2 (subtropical cluster), 4 (GoM/Caribbean cluster), or 5 (central Atlantic cluster), the track will likely stay in that cluster. For cluster 4, the GoM/Caribbean cluster, tracks predominately stay in cluster 4, and this is also the cluster with the highest percentage of landfalls (55% of East Coast and 76% of Gulf Coast) preserving the observed mainland landfall distribution. Cluster 3, the standard MDR cluster, is an interesting case, with cluster 5, the central Atlantic cluster, edging out cluster 3 for the most popular resulting cluster. Both clusters 3 and 5 share the MDR formation region (as seen in Figs. 3c,e), and there is a larger portion of cluster-5 tracks than cluster-3 tracks from which to choose (roughly  $\frac{1}{3}$  more, 263 as compared with 201). According to our analysis, a random draw of the region will result in more cluster 5, central Atlantic

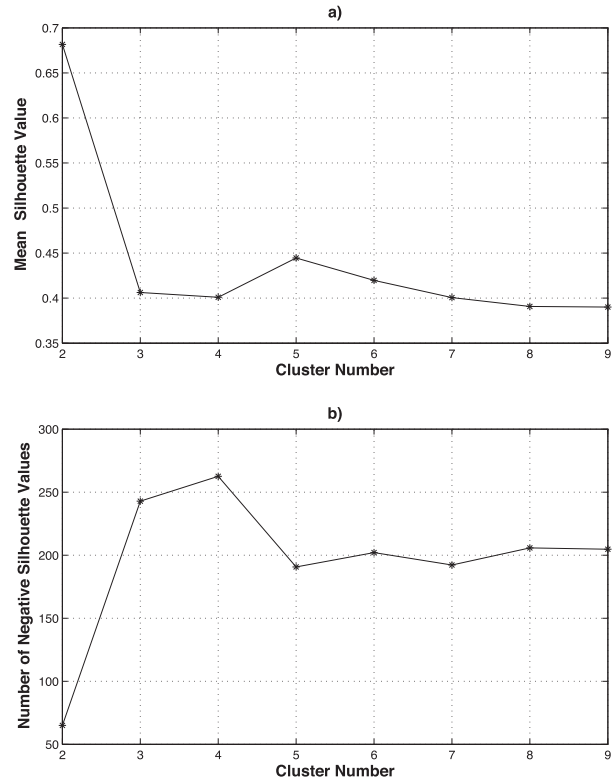


FIG. A1. (a) A measure of the  $k$ -means cluster cohesiveness (mean silhouette values) and (b) a measure of the number of misclassified members (number of negative silhouette values), of clustered observed North Atlantic TC track data.

cluster tracks. This will result in fewer mainland landfalls, because cluster-5 tracks tend not to make landfall.

## 5. Summary and future work

C<sup>3</sup>-HITS is a data-based model for the stochastic simulation of North Atlantic hurricane tracks given preseason climate covariates. It employs machine learning and statistical estimation tools to preserve the spatial continuity and orientation of the tracks simulated and to account for spatial dependence in the hurricane formation process. One aim is to aid in the creation of accurate landfall statistics. The underlying track simulation model, HITS, is especially effective near land and in the main development region of North Atlantic tropical storms, where there is the densest best track data (Nakamura et al. 2015). We have consistently found it to perform better than Markovian models that do one-step conditioning (Nakamura et al. 2015). The trajectories generated are more realistic and less diffusive than those generated by Markovian models, and as a result, the simulations translate into better occurrence statistics. Since all tracks created are drawn from segments of historical tracks, all the associated track information is automatically inherited. This provides a consistent multivariate representation of all historical track information including latitude, longitude, time, maximum 1-min sustained surface wind, landfall, central pressure from

TABLE A1. Observed mean cluster mass moments of  $k$ -means clustered North Atlantic TC tracks.

Cluster	Centroid $X$ (°E)	Centroid $Y$ (°N)	Variance $X$	Variance $Y$	Variance $XY$
1	-56.93	35.73	322.81	143.00	162.14
2	-61.92	33.42	43.58	32.91	23.25
3	-57.64	25.33	148.08	101.80	-39.65
4	-85.05	24.20	23.31	13.52	-3.09
5	-47.94	16.51	45.81	9.98	-10.36

1979, and wind radii maximum extent in quadrants at 34, 50, and 64 kt ( $1 \text{ kt} \approx 0.5 \text{ m s}^{-1}$ ) from 2004.

The  $C^3$ -HITS model adds the impact of the state of the climate at the beginning of the season on the simulated tracks at the time of storm formation. Here, we conditioned the stochastic storm-track generation on SST, in the tropical Pacific and tropical North Atlantic that is known to affect the interannual variability of hurricane tracks. Some statistical models create index-based runs by selecting years under those conditions (dividing the data in half or thirds).  $C^3$ -HITS is unique in that it can include all data instead of subselecting. Index-based runs can be set to favor an individual index value so that gradual changes in the formation and track shape can be observed. Also, all the indices can be run together for covarying traits.

Our simulation results show that adding climate information ( $C^3$ -HITS) reduces the RMSE error of the modeled basin total formations and mainland landfalls per season and adds more value to the model in the warm Niño-3.4 index and the warm MDR index cases (Figs. 5 and 6). Model skill was shown using an out-of-sample hindcast of the 2018 formation total and landfalls (Figs. 5e and 6e). When examining multihazard probability for  $C^3$ -HITS modeled (Fig. 7) and observed (Fig. 8), increasing TC per season shows a decrease in the probability that starts in the Niño-3.4 warm and MDR cold to the neutral corner and works its way slowly over to Niño-3.4 cold and MDR warm. It is notable that the  $C^3$ -HITS probability calculated from the simulated tracks and the observed probabilities calculated from the Poisson fit to observations mirror each other even though a Poisson distribution of landfall was not imposed in the model.

Further development of  $C^3$ -HITS will bring in surface winds and precipitation fields associated with the historical data and use them to generate an analog simulation of the wind and precipitation fields associated with the HITS tracks. This would include directional exposure of the track over land, to provide the key information needed to determine moderate to extreme flooding in a subsequent application of a hydrologic model. The storm-related fields can be pulled from historical reanalyses and/or gridded station observations. Uncertainty can be addressed by producing a large number of track simulations that will generate a distribution of jointly selected, two-dimensional, storm-related wind, and precipitation fields.

*Acknowledgments.* This work was funded by Jupiter Intelligence under Grant JI CU18-0639. This paper is Lamont-Doherty Earth Observatory Publication Number 8483.

*Data availability statement.* Data analyzed in this study were a reanalysis of existing data, which are openly available at locations cited in section 2. Computed mean cluster moments are located in appendix Table A1. Computed coefficients of the Poisson model are located in appendix Table B2.

## APPENDIX A

### Cluster Analysis for Initializing Climate-States Dependent Tracks

As indicated in section 2, the new, HURDAT 2 (1851–2016) was used to identify track families using a  $k$ -means clustering

TABLE B1. Preseason (April–June) Kendall correlation matrix of climate indices and the five clusters (C1–C5) of  $k$ -means clustered North Atlantic TC tracks. The MDR index is the SST in the main development region of North Atlantic tropical cyclones. The Niño-3.4 and Niño-1+2 indices are regions of tropical Pacific SST. The GoM index is Gulf of Mexico SST, and the AMM is the Atlantic meridional mode, a climate pattern in the South Atlantic Ocean defined by a shift in the intertropical convergence zone caused by fluctuations in SSTs and easterly trade winds.

	Year	MDR	Niño-3.4	Niño-1+2	GoM	AMM	C1	C2	C3	C4	C5
Year	1.00	0.23	0.22	0.11	-0.12	-0.26	0.06	-0.08	0.19	0.08	0.20
MDR	0.23	1.00	0.38	0.14	-0.26	0.79	0.30	-0.19	0.37	0.05	-0.01
Niño-3.4	0.22	0.38	1.00	0.64	-0.10	0.13	0.12	-0.16	-0.22	-0.40	-0.09
Niño-1+2	0.11	0.14	0.64	1.00	-0.02	-0.01	0.04	0.04	-0.13	-0.28	-0.12
GoM	-0.12	-0.26	-0.10	-0.02	1.00	-0.18	-0.13	0.20	0.03	0.08	-0.01
AMM	-0.26	0.79	0.13	-0.01	-0.18	1.00	0.16	-0.02	0.32	0.08	-0.10
C1	0.06	0.30	0.12	0.04	-0.13	0.16	1.00	-0.24	0.05	-0.16	-0.07
C2	-0.08	-0.19	-0.16	0.04	0.20	-0.02	-0.24	1.0	-0.25	0.35	0.10
C3	0.19	0.37	-0.22	-0.13	0.03	0.32	0.05	-0.25	1.0	0.09	-0.15
C4	0.08	0.05	-0.40	-0.28	0.08	0.08	-0.16	0.35	0.09	1.0	0.40
C5	0.20	-0.01	-0.09	-0.12	-0.01	-0.10	-0.07	0.10	-0.15	0.40	1.0

TABLE B2. Coefficients of Poisson model North Atlantic TC formations per season by  $k$ -means cluster of the moments of the TC tracks. Plain text (step 1), boldface type (step 2), and italics (step 3) indicate successive steps in creating a model of all clusters.

Cluster	MDR	Niño-3.4	Intercept	Cluster 2	Cluster 3	Cluster 4
1	0.9903	0	-0.3651	0	0	0
2	-0.2747	0	1.5168	0	0	0
<b>3</b>	<b>1.0599</b>	<b>-0.6943</b>	<b>0.8140</b>	<b>-0.0782</b>	<b>0</b>	<b>0</b>
<b>4</b>	<b>0.4538</b>	<b>-0.5150</b>	<b>1.2228</b>	<b>0.0729</b>	<b>0</b>	<b>0</b>
5	0	0	0.7624	0	-0.0936	0.1025

procedure. The selection of the appropriate number of clusters is a subjective process. Here we are aided in the selection by the silhouette (sil) evaluation method (see N09). The method provides two numbers based on repeated application

of the clustering process. Figure A1a shows the mean “sil” number resulting from 1000 simulations as a function of the number of clusters. High values indicate a high cohesiveness of the cluster members and good separation from other

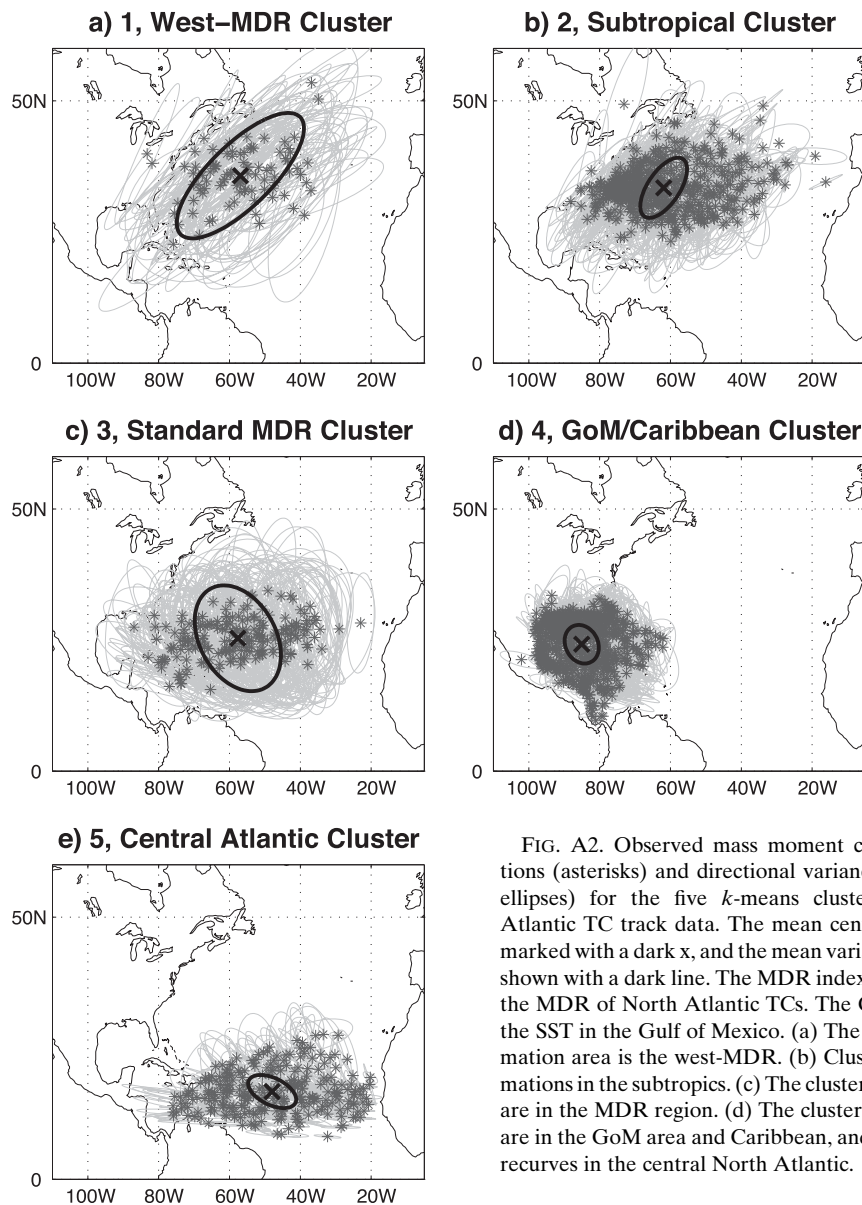


FIG. A2. Observed mass moment centroid locations (asterisks) and directional variance (light-gray ellipses) for the five  $k$ -means clusters of North Atlantic TC track data. The mean centroid value is marked with a dark x, and the mean variance ellipse is shown with a dark line. The MDR index is the SST in the MDR of North Atlantic TCs. The GoM index is the SST in the Gulf of Mexico. (a) The cluster-1 formation area is the west-MDR. (b) Cluster 2 has formations in the subtropics. (c) The cluster 3 formations are in the MDR region. (d) The cluster-4 formations are in the GoM area and Caribbean, and (e) cluster 5 recurs in the central North Atlantic.

### North Atlantic Tropical Storm Formation Day 850 hPa Heights

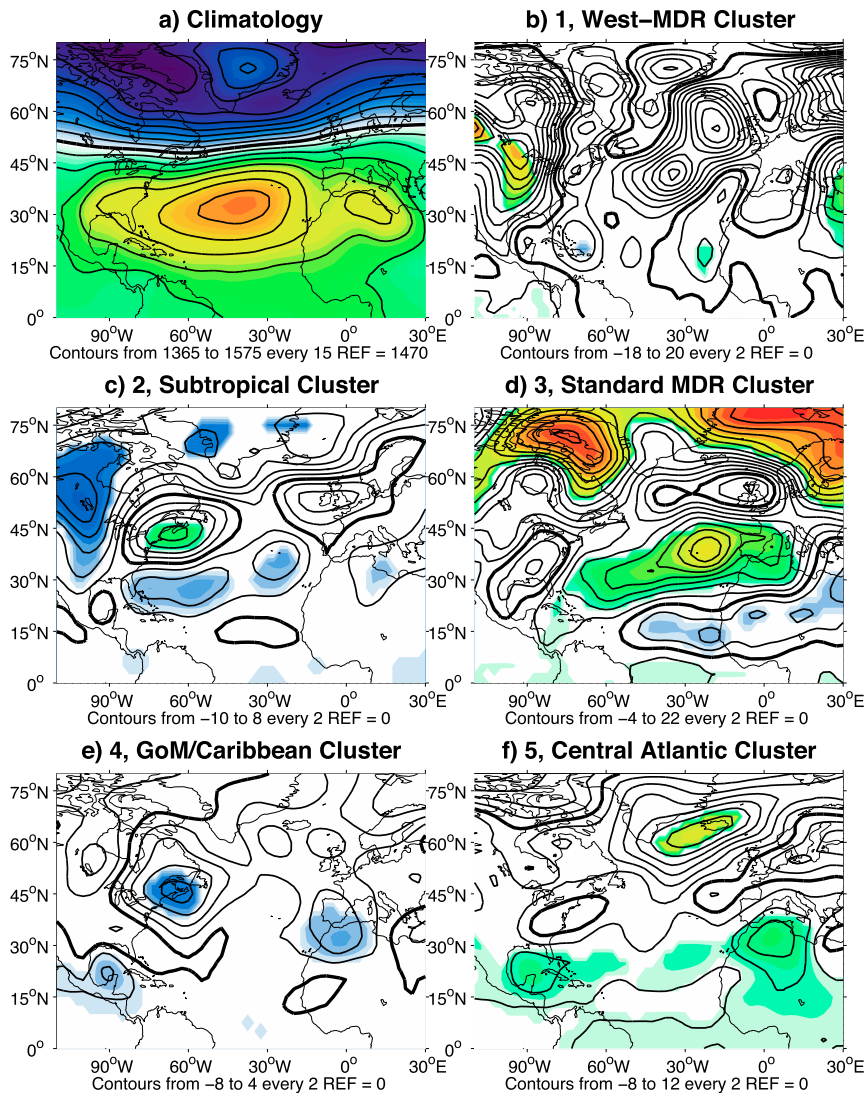


FIG. A3. (a) The average of North Atlantic tropical storm formation day 850-hPa heights for all clusters showing the location of the subtropical high and the predominant low-level flow. (b)–(f) Anomalies of 850-hPa heights for each cluster (contours), with significant areas (to the 5% level using a Student's  $t$  test) indicated by color shades.

clusters (Kaufman and Rousseeuw 1990). Although two clusters represent the highest mean sil value, based on an examination of the degree of physical separation of the tracks and based on the formation, location, and track length distributions, a sil of five, which is the next maximum is selected. In Fig. A1b, the occurrence of negative sil number in 1000 clustering attempts is shown. Low values indicate a low number of misclassified tracks (Kaufman and Rousseeuw 1990). Again, a two-cluster solution is an optimal choice, but five and seven clusters are also local minima. The combination of a local maximum in mean sil value and local minima of negative sil numbers at five show an organization in the data at five clusters that is cohesive and well classified. Based on

these results, five clusters were selected for both a relative maximum in mean sil and a relative minimum for negative sil. Using a similar selection criterion, six clusters were found in N09, where the earlier, HURDAT (1948–2006) historical track record was used. Differences in the clustering include a new expanded dataset, different time periods, 6-h versus 2-h interpolated data, and exclusion of tracks of single point length.

The resulting  $X$  and  $Y$  locations of centroid values, depicted as stars and the three directions of variance values ( $X$ ,  $Y$ , and covariance  $XY$ ), depicted as ellipses, are shown in Figs. A2a–e for each of the five clusters as are the cluster-mean values (dark  $x$  and thick ellipses). Table A1 presents the mean centroid location and directional variance for each cluster (1–5).

Figure A3a is the average of North Atlantic tropical cyclone formation day, 850-hPa heights for all clusters showing the location of the subtropical high and the predominant low-level flow. In Figs. A3b–f anomalies for each cluster are contoured. Colors show anomalies significant to the 5% level using a Student's  $t$  test. Cluster 2, the subtropical cluster (Fig. A3c), has close contours near the East Coast, and the color indicates a small intensification of the subtropical high in the west and steering toward the north along the coast. This agrees with the formation locations and the tracks for cluster 2. Cluster 3, the standard MDR cluster (Fig. A3d), is related to an intensification of the subtropical high, centered on the east and center of the climatological feature (Fig. A3a). With that is increased steering across the basin toward the North American east coast. The pattern for cluster 4, the GoM/Caribbean cluster (Fig. A3e), is consistent between the flow around the low pressure center in the GoM and the low pressure center over the northeastern United States, and the fact that the cyclones stay in the GoM and travel northward toward the coast. Cluster 5, the central Atlantic cluster (Fig. A3f), consists of tracks that predominately drift northward.

## APPENDIX B

### The Poisson Regression Models

Kendall's tau correlations for all cluster formations per season on the select indices are generally weak (Table B1), with MDR and Niño variables emerging as the highest correlations to the clusters with an average absolute value over all clusters of 0.18 and 0.20, respectively. The AMM and MDR indices are highly correlated with each other at 0.79. However, the MDR index correlates a bit higher with cluster 3 than AMM (0.37 vs 0.32) and overall for all clusters. MDR and Niño variables were found to be sufficient, accounting for much or all of the nonrandom component of the total North Atlantic basin TC tracks, in Sabbatelli and Mann (2007).

The Poisson regression model chooses identical predictors in some cases whether AIC or BIC is used as the criterion, and in some cases AIC leads to additional predictors being selected. The coefficients of the best models identified (fitting) for each cluster are shown in Table B2. Then, to estimate (simulate) the Poisson rate,

$$\lambda_{it} = \mathbf{X}_i \boldsymbol{\beta}, \quad (\text{B1})$$

where formations  $\lambda_{it}$  are identified for cluster  $i$  and year  $t$ ,  $\mathbf{X}_i$  is the matrix of predictors, and  $\boldsymbol{\beta}$  is the coefficient vector.

Since our application considers simulation from the model that is fit, we simulate the conditional distribution of the number of TCs for the season for each cluster, considering also the uncertainty distribution of the regression parameters  $\boldsymbol{\beta}$ . Recognizing that the AIC and the BIC penalize bias and variance in model fitting to different degrees, we opted to use the larger model selected by AIC but then consider the additional uncertainty in fitting that model (which potentially has a lower bias than the BIC selected model) and use that in the sampling process for the simulation.

For the multivariate DAG considering both climate indicators and TCs, the following structure that respects the spatial dependence across the clusters and the dependence on the climate covariates was identified. Clusters 1 and 2 are functions of the MDR index (Table B1; step 1); clusters 3 and 4 are functions of the MDR index, the Niño-3.4 index, and cluster 2 (step 2); and cluster 5 is a function of clusters 3 and 4 (step 3). Table B2 shows these successive steps in the computation of predictors with plain (step 1), boldface (step 2), and italic (step 3) text, and Fig. 3 shows the corresponding DAG.

To develop the stochastic catalog for the upcoming season, we consider  $N$  simulations from the DAG identified. For each simulation, we use the DAG to simulate the number of TCs,  $h_i$  for each cluster  $i$ . Next, formation locations for  $h_i$  TCs are randomly sampled from the historical TCs identified with that cluster. HITS is then used to simulate tracks starting from each of those locations.

## REFERENCES

- Camargo, S. J., A. G. Barnston, and S. E. Zebiak, 2005: A statistical assessment of tropical cyclone activity in atmospheric general circulation models. *Tellus*, **57A**, 589–604, <https://doi.org/10.3402/tellusa.v57i4.14705>.
- Cha, S.-H., 2007: Comprehensive survey on distance/similarity measures between probability density functions. *Int. J. Math. Models Methods Appl. Sci.*, **4**, 300–307.
- Chiang, J. C. H., and D. J. Vimont, 2004: Analogous Pacific and Atlantic meridional modes of tropical atmosphere–ocean variability. *J. Climate*, **17**, 4143–4158, <https://doi.org/10.1175/JCLI4953.1>.
- Comes, T., and B. Van de Walle, 2014: Measuring disaster resilience: The impact of Hurricane Sandy on critical infrastructure systems. *Proc. 11th Int. Conf. on Information Systems for Crisis Response and Management*, University Park, PA, The Pennsylvania State University, 195–204.
- Coxe, S., S. G. West, and L. S. Aiken, 2009: The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *J. Pers. Assess.*, **91**, 121–136, <https://doi.org/10.1080/00223890802634175>.
- Elsner, J. B., 2003: Tracking hurricanes. *Bull. Amer. Meteor. Soc.*, **84**, 353–356, <https://doi.org/10.1175/BAMS-84-3-353>.
- , and C. P. Schertmann, 1993: Improving extended-range seasonal predictions of intense Atlantic hurricane activity. *Wea. Forecasting*, **8**, 345–351, [https://doi.org/10.1175/1520-0434\(1993\)008<0345:IERSP0>2.0.CO;2](https://doi.org/10.1175/1520-0434(1993)008<0345:IERSP0>2.0.CO;2).
- , and A. B. Kara, 1999: *Hurricanes of the North Atlantic: Climate and Society*. Oxford University Press, 512 pp.
- , and T. H. Jagger, 2004: A hierarchical Bayesian approach to seasonal hurricane modeling. *J. Climate*, **17**, 2813–2827, [https://doi.org/10.1175/1520-0442\(2004\)017<2813:AHBATS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<2813:AHBATS>2.0.CO;2).
- , and —, 2006: Prediction models for annual U.S. hurricane counts. *J. Climate*, **19**, 2935–2952, <https://doi.org/10.1175/JCLI3729.1>.
- , B. H. Bossak, and X. F. Niu, 2001: Secular changes to the ENSO-US hurricane relationship. *Geophys. Res. Lett.*, **28**, 4123–4126, <https://doi.org/10.1029/2001GL013669>.
- Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436**, 686–688, <https://doi.org/10.1038/nature03906>.
- Fan, C., A. Mostafavi, A. Gupta, and C. Zhang, 2018: A system analytics framework for detecting infrastructure-related topics



- in disasters using social sensing. *Workshop of the European Group for Intelligent Computing in Engineering*, Springer, 74–91.
- Goerss, J. S., 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.*, **128**, 1187–1193, [https://doi.org/10.1175/1520-0493\(2000\)128<1187:TCTFUA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<1187:TCTFUA>2.0.CO;2).
- Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nunez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**, 474–479, <https://doi.org/10.1126/science.1060040>.
- Hall, T., and S. Jewson, 2008: SST and North American tropical cyclone landfall: A statistical modeling study. arXiv:0801.1013, 30 pp., <https://arxiv.org/pdf/0801.1013.pdf>.
- , and E. Yonekura, 2013: North American tropical cyclone landfall and SST: A statistical model study. *J. Climate*, **26**, 8422–8439, <https://doi.org/10.1175/JCLI-D-12-00756.1>.
- Kaplan, A., M. Cane, Y. Kushnir, A. Clement, M. Blumenthal, and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res.*, **103**, 18 567–18 589, <https://doi.org/10.1029/97JC01736>.
- Kaufman, L., and P. J. Rousseeuw, 1990: *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley and Sons, 342 pp.
- Klotzbach, P. J., 2010: El Niño–Southern Oscillation’s impact on Atlantic basin hurricanes and U.S. landfalls. *J. Climate*, **24**, 1252–1263, <https://doi.org/10.1175/2010JCLI3799.1>.
- , M. A. Saunders, G. D. Bell, and E. S. Blake, 2017: North Atlantic seasonal hurricane prediction: Underlying science and an evaluation of statistical models. *Climate Extremes: Patterns and Mechanisms*, *Geophys. Monogr.*, Vol. 226, Amer. Geophys. Union, 315–328, <https://doi.org/10.1002/9781119068020.CH19>.
- Knutson, T. R., and Coauthors, 2013: Dynamical downscaling projections of twenty-first-century Atlantic hurricane activity: CMIP3 and CMIP5 model-based scenarios. *J. Climate*, **26**, 6591–6617, <https://doi.org/10.1175/JCLI-D-12-00539.1>.
- Kossin, J. P., S. J. Camargo, and M. Sitkowski, 2010: Climate modulation of North Atlantic hurricane tracks. *J. Climate*, **23**, 3057–3076, <https://doi.org/10.1175/2010JCLI3497.1>.
- Lall, U., and A. Sharma, 1996: A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resour. Res.*, **32**, 679–693, <https://doi.org/10.1029/95WR02966>.
- Landsea, C. W., and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, <https://doi.org/10.1175/MWR-D-12-00254.1>.
- Loader, C., 2006: *Local Regression and Likelihood*. Springer Science & Business Media, 290 pp.
- Meléndez, E., and J. Hinojosa, 2017: Estimates of post-Hurricane Maria exodus from Puerto Rico. City University of New York Hunter College Center for Puerto Rican Studies Research Brief RB2017-01, 7 pp., <https://centropr.hunter.cuny.edu/research/data-center/research-briefs/estimates-post-hurricane-maria-exodus-puerto-rico>.
- Mestre, O., and S. Hallegatte, 2009: Predictors of tropical cyclone numbers and extreme hurricane intensities over the North Atlantic using generalized additive and linear models. *J. Climate*, **22**, 633–648, <https://doi.org/10.1175/2008JCLI2318.1>.
- Nakamura, J., U. Lall, Y. Kushnir, and S. J. Camargo, 2009: Classifying North Atlantic tropical cyclone tracks by mass moments. *J. Climate*, **22**, 5481–5494, <https://doi.org/10.1175/2009JCLI2828.1>.
- , —, —, and B. Rajagopalan, 2015: HITS: Hurricane intensity and track simulator with North Atlantic Ocean applications for risk assessment. *J. Appl. Meteor. Climatol.*, **54**, 1620–1636, <https://doi.org/10.1175/JAMC-D-14-0141.1>.
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948, [https://doi.org/10.1175/1520-0442\(1994\)007<0929:IGSSTA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<0929:IGSSTA>2.0.CO;2).
- Sabbatelli, T. A., and M. E. Mann, 2007: The influence of climate state variables on Atlantic tropical cyclone occurrence rates. *J. Geophys. Res.*, **112**, D17114, <https://doi.org/10.1029/2007JD008385>.
- Saunders, M. A., P. J. Klotzbach, A. S. R. Lea, C. J. Schreck, and M. M. Bell, 2020: Quantifying the probability and causes of the surprisingly active 2018 North Atlantic hurricane season. *Earth Space Sci.*, **7**, e2019EA000852, <https://doi.org/10.1029/2019EA000852>.
- Sebastian, A., and Coauthors, 2017: Hurricane Harvey Report: A fact-finding effort in the direct aftermath of Hurricane Harvey in the Greater Houston Region. Delft University of Technology Rep., 102 pp., <http://resolver.tudelft.nl/uuid:54c24519-c366-4f2f-a3b9-0807db26ff69c>.
- Shapiro, L. J., and S. B. Goldenberg, 1998: Atlantic sea surface temperatures and tropical cyclone formation. *J. Climate*, **11**, 578–590, [https://doi.org/10.1175/1520-0442\(1998\)011<0578:ASSTAT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<0578:ASSTAT>2.0.CO;2).
- Sobel, A. H., 2014: *Storm Surge: Hurricane Sandy, Our Changing Climate, and Extreme Weather of the Past and Future*. 1st ed. Harper-Collins, 336 pp.
- Tang, B. H., and J. D. Neelin, 2004: ENSO influence on Atlantic hurricanes via tropospheric warming. *Geophys. Res. Lett.*, **31**, L24204, <https://doi.org/10.1029/2004GL021072>.
- Vecchi, G. A., and Coauthors, 2013: Multiyear predictions of North Atlantic hurricane frequency: Promise and limitations. *J. Climate*, **26**, 5337–5357, <https://doi.org/10.1175/JCLI-D-12-00464.1>.
- Villarini, G., G. A. Vecchi, and J. A. Smith, 2010: Modeling the dependence of tropical storm counts in the North Atlantic basin on climate indices. *Mon. Wea. Rev.*, **138**, 2681–2705, <https://doi.org/10.1175/2010MWR3315.1>.
- , —, and —, 2012: U.S. landfalling and North Atlantic hurricanes: Statistical modeling of their frequencies and ratios. *Mon. Wea. Rev.*, **140**, 44–65, <https://doi.org/10.1175/MWR-D-11-00063.1>.
- Vimont, D. J., and J. P. Kossin, 2007: The Atlantic meridional mode and hurricane activity. *Geophys. Res. Lett.*, **34**, L07709, <https://doi.org/10.1029/2007GL029683>.
- Vitart, F., J. L. Anderson, and W. F. Stern, 1997: Simulation of interannual variability of tropical storm frequency in an ensemble of GCM integrations. *J. Climate*, **10**, 745–760, [https://doi.org/10.1175/1520-0442\(1997\)010<0745:SOIVOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<0745:SOIVOT>2.0.CO;2).
- , and Coauthors, 2007: Dynamically-based seasonal forecasts of Atlantic tropical storm activity issued in June by EUROSIP. *Geophys. Res. Lett.*, **34**, L16815, <https://doi.org/10.1029/2007GL030740>.
- Webster, P. J., G. J. Holland, J. A. Curry, and H.-R. Chang, 2005: Changes in tropical cyclone number and intensity in a warming environment. *Science*, **309**, 1844–1846, <https://doi.org/10.1126/science.1116448>.
- Zorrilla, C. D., 2017: The view from Puerto Rico—Hurricane Maria and its aftermath. *N. Engl. J. Med.*, **377**, 1801–1803, <https://doi.org/10.1056/NEJMp1713196>.