A Flexible Pipeline for Prediction of Tropical Cyclone Paths

Niccolò Dalmasso [* 1] **Robin Dunn** [* 1] **Benjamin LeRoy** [* 1] **Chad Schafer** [1]

1. Introduction

Hurricanes and, more generally, tropical cyclones (TCs) are rare, complex natural phenomena of both scientific and public interest. The importance of understanding TCs in a changing climate (Knutson et al., 2010) has increased as recent TCs have had devastating impacts on human lives and communities (ABC; American Red Cross). Moreover, good prediction and understanding about the complex nature of TCs can mitigate some of these human and property losses. Though TCs have been studied from many different angles, more work is needed from a statistical approach of providing prediction regions (Camargo & Wing, 2016). The current state-of-the-art in TC prediction bands comes from the National Hurricane Center at National Oceanographic and Atmospheric Administration (NOOA), whose proprietary model provides "cones of uncertainty" for TCs through an analysis of historical forecast errors (NOAA/NHC, 2018).

The contribution of this paper is twofold. We introduce a new pipeline that encourages transparent and adaptable prediction band development by streamlining cyclone track simulation and prediction band generation. We also provide updates to existing models and novel statistical methodologies in both areas of the pipeline respectively.

Our pipeline has many desirable properties. It provides an *easy-to-use* framework to estimate prediction bands for new TCs. The pipeline consists of two separable components for TC track simulation and prediction band creation. Each individual part is designed to be *interchangeable*; provided the input and output are the same, each part can easily be replaced by the user's favorite method. We provide *purely data driven* TC simulation approaches, but one can easily pair these with other meteorological methods. This is especially important in a climate change scenario, where TC behavior might change drastically in a relatively short period of time. Our current pipeline is implemented as an `R` package and is available on `Github`[1]. Section 2 details its structure and components.

We extend the current statistical methods for TC simulations, combining autoregressive path prediction models (Vickery et al., 2000) and models of TC lifespan (Hall & Jewson, 2007). We propose novel statistically valid prediction bands by leveraging geometric properties and statistical notions on the depth and centrality of simulated curves. We benchmark against the prediction bands creation technique found in NOOA's state-of-the-art model (NOOA/NHC, 2018). We construct prediction band methods that provide coverage in either a pointwise or uniform manner, i.e., aiming to cover either the future measurements of the TC path or the full TC track respectively. Section 3 discusses these contributions in detail, highlights experimental results on held-out data, and provides empirical assessments of the prediction band methods' statistical properties.

2. Pipeline Structure and Flow

Figure 1 provides a schematic representation of our proposed pipeline. The leftmost stage corresponds to the input, which is a series of spatiotemporal measurements of a TC track. In the current iteration we rely on the Atlantic Oceanographic and Meteorological Laboratory of NOAA, which publicly shares data on TC paths in the North Atlantic (Landsea & Franklin, 2013; NOAA/AOML, 2017); each path has information on time, latitude and longitude, recorded every six hours until the storm dies. For each new TC, the pipeline uses only a few initial points to simulate a series of tracks the TC could traverse. The curve simulation models are tuned according to the computational complexity of the models and the available processing power. When training the track simulation model, the training TC paths should be representative of the world region for which the pipeline will make predictions. For instance, TC track data over the Atlantic Ocean should not be used to train models over the Pacific Ocean. Once these simulations are generated, the next stage uses this ensemble to produce a prediction band at a chosen statistical confidence level, usually between 90% and 99%. The confidence refers to the statistical coverage of the prediction band, i.e., the rate at which the prediction band covers the true TC points or the full TC track. When considering different prediction band techniques, one could use a held-out set of TC tracks to assess the performance,

---
[*]Equal contribution [1]Department of Statistics & Data Science, Carnegie Mellon University, Pittsburgh, PA, USA. Correspondence to: Benjamin LeRoy <bpleroy@stat.cmu.edu>.

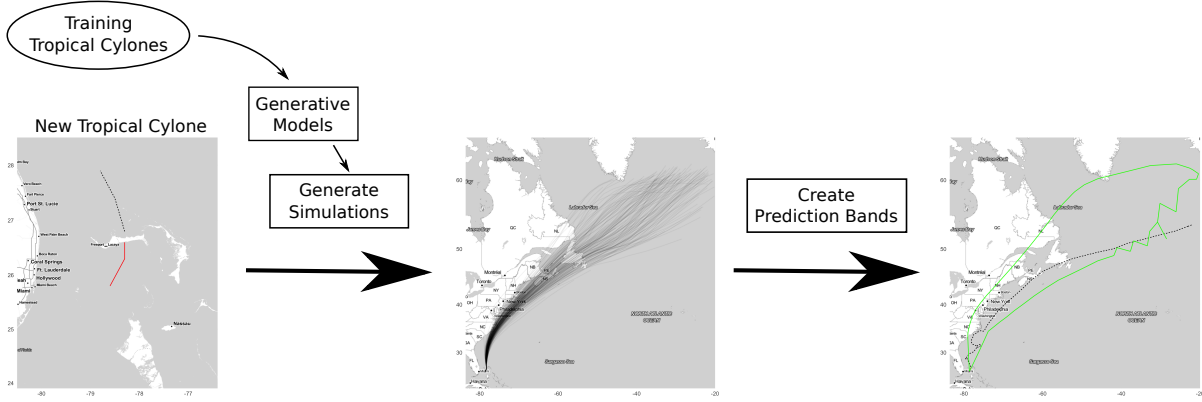[1]github.com/Mr8ND/TC-prediction-bands/

Figure 1. Pipeline structure schema. From left to right, when the start (12-18 hours) of new tropical cyclone is observed, a pre-trained model generates track simulations, and the prediction band methods use the simulated ensemble to construct the prediction band.

taking into consideration the band coverage and size.

It is important to remark that, as said in the introduction, the pipeline is model-agnostic. The choices of statistical framework or machine learning models do not affect the structure or flow of the pipeline, provided the input and output of each stage is not changed. Similarly, a more granular or richer input will not modify the pipeline structure but will only affect the models in the simulation and band prediction stages.

3. Current Methodology

Our methodology for predicting TC paths involves two major modeling components: the simulation of potential paths and the construction of prediction bands from those paths.

To simulate TC paths, first we have developed models that make sequential predictions for the location of a TC. Our current approaches build upon foundational work by Vickery et al., who predicted change in TC bearing and speed from linear models with basin-specific coefficients (Vickery et al., 2000). By sequentially predicting changes in TC bearing and speed at each time step, these models propagate complete TC paths from the initial points. Whereas Vickery et al. include two lag terms in their bearing model and one lag term in their speed model, we have developed two sets of models. A non-autoregressive (non-AR) model that predicts change in bearing and change in speed at time step i from latitude (at time step i), longitude, bearing (calculated between time steps $i-1$ and i), and speed (between time steps $i-1$ and i). An autoregressive (AR) model that also includes a lag term for change in bearing or change in speed, respectively. Conceptually, the AR models allow us to model climatological phenomena such as TCs continuing to gain/lose speed or continuing to turn in the same direction. Empirically, table 1 shows that the predic-

tion bands constructed from AR simulations capture greater proportions of the test TC points and full paths. To complete our generative models, we also develop models for TC lysis (death). Our first lysis model is a block-specific logistic regression. Earlier work on Gaussian weighting of historical TC lysis locations showed the promise of location-dependant lysis models (Hall & Jewson, 2007). The logistic regression model predicts TC death at each point based on latitude, longitude, bearing, and speed. We also include a kernel density approach to lysis modeling, which draws TC lengths from a smoothed distribution of TC lifespans. In table 1, we see that kernel density lysis generally yields slightly higher coverage rates on held-out TCs.

We train and evaluate our models on 1008 TCs from the NOAA database between June 1851 and November 2016 (NOAA/AOML, 2017). We use 702 TCs to train the bearing, speed, and lysis models on 10-degree latitude-longitude squares. Then we generate potential paths from the initial observations of the 306 test TCs. To predict trajectories from a non-AR model, we initialize the path with the first two TC observations (6 hours), to observe speed and bearing. To predict trajectories from an AR model, we start with the first three TC observations (12 hours), to observe change in speed and change in bearing. For each new/test TC we generate a set of curves (350 for current analysis) using a combination of AR or non-AR model and lysis method. Upon simulating 350 potential TC paths for each test TC under each method, we construct prediction bands for the TC paths. This is similar to the approach of the National Hurricane Center of NOAA, which predicts circular regions in which a TC center will fall with high probability, up to five days out from the point of prediction (NOAA/NHC, 2018). In contrast, our approach aims to predict the location of the TC across its entire lifespan.

To construct prediction bands, we have developed two

Simulation Curve Type	δ-Ball	Convex Hull	Kernel Density Estimate	Spherical
AR & Logistic	0.88 / 0.43	0.75 / 0.33	0.50 / 0.14	0.31 / 0.03
AR & Kernel	0.88 / 0.44	0.86 / 0.37	0.55 / 0.19	0.31 / 0.04
Non-AR & Logistic	0.85 / 0.40	0.72 / 0.33	0.45 / 0.13	0.24 / 0.03
Non-AR & Kernel	0.88 / 0.43	0.79 / 0.39	0.50 / 0.16	0.24 / 0.03

Table 1. Median proportion of true test TC points captured by their prediction bands / proportion of test TCs with 100% of points captured. Based on prediction bands attempting to obtain 90% coverage.

approaches based on pointwise confidence intervals and two approaches based on uniform confidence structure. One pointwise approach constructs a kernel density from the set of all simulated points; the prediction band is the $100(1-\alpha)\%$ level set at a pre-specified coverage level. The second pointwise approach uses a depth metric proposed by Geenens and Nieto-Reyes to determine the most central simulated path (Geenens & Nieto-Reyes, 2017). At each time step t, this approach finds the smallest sphere, centered at the tth point of the most central path, that contains $100(1-\alpha)\%$ of the other simulated TCs' points at time t. We connect these spheres to form the prediction band. Out of the four approaches, this spherical ball method is closest to the approach of the National Hurricane Center of NOAA (NOAA/NHC, 2018). The uniform approaches treat the entire simulated tracks as the objects of interest. One uniform approach creates a prediction band from the convex hull of the top $100(1-\alpha)\%$ of deepest curves. The second uniform prediction band contains a set of balls centered at all points in the top $100(1-\alpha)\%$ of deepest curves. The balls have a common radius (δ), which is the smallest radius such that all balls are connected to at least one other ball. We call this prediction band method the δ-ball method.

In table 1, we see that the δ-ball prediction bands have both the highest pointwise and uniform TC coverage, across all simulation models at $\alpha = 0.10$. The difference between logistic and kernel lysis coverage is negligible for the δ-ball TCs. Since the logistic approach incorporates location information and tends to yield PBs with smaller area, we recommend the δ-ball approach with logistic lysis and AR simulation models. The δ-ball prediction bands are also smaller than the convex hull prediction bands (the other proposed uniform prediction band method). One can notice that the δ-ball prediction bands still achieve less than 90% pointwise and uniform coverage rates. As a prediction band method, figure 2 shows that the δ-ball method achieves proper uniform coverage under experimental simulations. Specifically, figure 2 shows the coverage results of experiments in which we construct the prediction band-generating TCs and the "true" TC from the same methods. This suggests that the less than 90% coverage rates may be due to a failure of the generative models to capture the true distribution of TC paths. Nevertheless, the δ-ball prediction bands
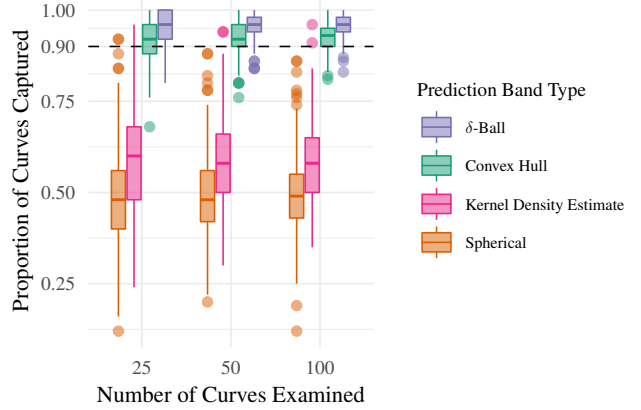

Average Uniform Containment per Curve

Figure 2. Uniform coverage of prediction bands, i.e. proportion of times the prediction band captures the whole TC track, generated from autoregressive and logistic-based lysis models. Pointwise prediction bands are intrinsically not well suited for this task.

with AR simulation models and kernel lysis are able to capture a median of 88% of test TC points, and they capture the full path for 44% of the test TCs.

4. Future Work and Extensions

We anticipate the modeling components to be expanded to include more extensive and potentially nonparametric time series models. From a statistical perspective, the current methodologies fall within a parametric bootstrap framework, but Bayesian models such as multi-dimensional Gaussian processes could be easily implemented, given the pipeline's flexibility. Incorporating inputs from different scientific probes (e.g., satellite images, oceanic surface temperatures) and different climatological models could also be feasible, by establishing parallel threads within the same pipeline.

References

ABC. Hurricane Florence death tool now at 39 in North Carolina. *ABC11*. URL https://abc11.com/weather/hurricane-florence-death-toll-now-at-39-in-north-carolina/4244283/.

American Red Cross, A. Mozambique cyclone: "almost everything is destroyed". *American Red Cross*. URL https://www.redcross.org/about-us/news-and-events/news/2019/mozambique-cyclone.html.

Camargo, S. J. and Wing, A. A. Tropical cyclones in climate models. *Wiley Interdisciplinary Reviews: Climate Change*, 7(2):211–237, 2016.

Geenens, G. and Nieto-Reyes, A. On the functional distance-based depth. *Manuscript in preparation.*, 2017.

Hall, T. M. and Jewson, S. Statistical modelling of North Atlantic tropical cyclone tracks. *Tellus A*, 59(4):486–498, 2007.

Knutson, T. R., McBride, J. L., Chan, J., Emanuel, K., Holland, G., Landsea, C., Held, I., Kossin, J. P., Srivastava, A. K., and Sugi, M. Tropical cyclones and climate change. *Nature Geoscience*, 3:157, feb 2010. URL https://doi.org/10.1038/ngeo779http://10.0.4.14/ngeo779https://www.nature.com/articles/ngeo779{#}supplementary-information.

Landsea, C. W. and Franklin, J. L. Atlantic hurricane database uncertainty and presentation of a new database format. *Monthly Weather Review*, 141(10):3576–3592, 2013.

NOAA/AOML. HURDAT 2 database. http://www.aoml.noaa.gov/hrd/hurdat/hurdat2-1851-2016-apr2017.txt, 2017.

NOAA/NHC. Definition of the NHC track forecast cone. https://www.nhc.noaa.gov/aboutcone.shtml, 2018.

Vickery, P., Skerlj, P., and Twisdale, L. Simulation of hurricane risk in the US using empirical track model. *Journal of Structural Engineering*, 126(10):1222–1237, 2000.