Identifying Distributional Differences in Convective Evolution Prior to Rapid Intensification in Tropical Cyclones

Trey McNeely [1]   **Galen Vincent** [1]   **Rafael Izbicki** [2]   **Kimberly M. Wood** [3]   **Ann B. Lee** [1]

[1]Department of Statistics & Data Science, Carnegie Mellon University
[2]Department of Statistics, Federal University of São Carlos, São Carlos, Brazil
[3]Department of Geosciences, Mississippi State University
imcneely@stat.cmu.edu

Abstract

Tropical cyclone (TC) intensity forecasts are issued by human forecasters who evaluate spatio-temporal observations (e.g., satellite imagery) and model output (e.g., numerical weather prediction, statistical models) to produce forecasts every 6 hours. Within these time constraints, it can be challenging to draw insight from such data. While high-capacity machine learning methods are well suited for prediction problems with complex sequence data, extracting interpretable scientific information with such methods is difficult. Here we leverage powerful AI prediction algorithms and classical statistical inference to identify patterns in the evolution of TC convective structure leading up to the rapid intensification of a storm, hence providing forecasters and scientists with key insight into TC behavior.

1   Introduction

Background. Tropical cyclones (TCs) are powerful, organized atmospheric systems that draw energy from the upper levels of the ocean. They rank among the deadliest and costliest natural disasters in the United States, and rising ocean heat content promises to amplify their danger over time [12]. TC intensity and motion forecasting alike have improved since the 1990s, but intensity forecasting has lagged behind [3]. Cases of rapid intensification (RI), defined for this work as an intensification of at least 25 knots within 24 hours, are especially difficult to predict [7, 8, 9, 20]. Some of the most impactful TCs in recent years underwent RI prior to landfall (e.g., Hurricanes Irma [2017], Dorian [2019], and Ida [2021]). Understanding TC evolution in the lead-up to RI events is therefore a critical component of forecasting and damage mitigation efforts. In this work, we leverage powerful machine learning (ML) methods to provide *scientific insight* into the evolution of TCs leading up to and during such events.

Multiple internal and external factors drive TC behavior [9]. This work focuses on a key internal factor: the spatio-temporal structure of deep convection—that is, vertical circulation of air—within the storm. Convective patterns within TCs serve as reliable indicators of intensity and drivers of intensity change; for example, identifying convective patterns is critical to the Dvorak Technique for TC intensity estimation [4, 15, 16, 5]. Since temperature generally decreases with height in the troposphere, the temperature of a cloud top is a proxy for its altitude. Thus taller clouds have colder tops, implying deeper (and thus stronger) convection. Stronger convection, particularly near the TC center, tends to be associated with stronger TCs. Infrared (IR) imagery from geostationary satellites such as the Geostationary Operational Environmental Satellites (GOES) provide high spatio-temporal resolution estimates of cloud-top temperature ($10.4\mu m$ wavelength) [19].

Scientific Goals. TC intensity forecasts are ultimately issued by human experts who synthesize domain knowledge with the outputs of many forecasting tools to issue a single forecast. It is thus unproductive to provide another statistical model which only predicts TC intensity directly from IR imagery; such black-box models are difficult for forecasters to utilize. Our goal is instead to provide meteorologists with an enhanced understanding of how convective evolution relates to RI events (Figure 1, blue arrow). Currently operational statistical forecast tools are generally limited to a handful of scalar summary statistics from each GOES image to aid forecasters in quickly digesting convective behavior [2, 1, 9]. Conversely, we work with the rich suite of ORB functions, a dictionary of scientifically meaningful one-dimensional functional features [14]. In this paper, we illustrate our methods

Figure 1: From Observation to Machine Learning to Science. Meteorologists seek to connect observations and summaries of TC evolution to processes such as RI (blue path). We pose this scientific question as a hypothesis test (bottom right). By casting the hypothesis test as a prediction problem, we are able to leverage high-capacity AI techniques, here ANNs (top right). Such black-box models often struggle to extract scientifically meaningful insight (red path). Our framework facilitates insight on labeled sequences of functional data via local results in the original space of scientific functional features (black path).
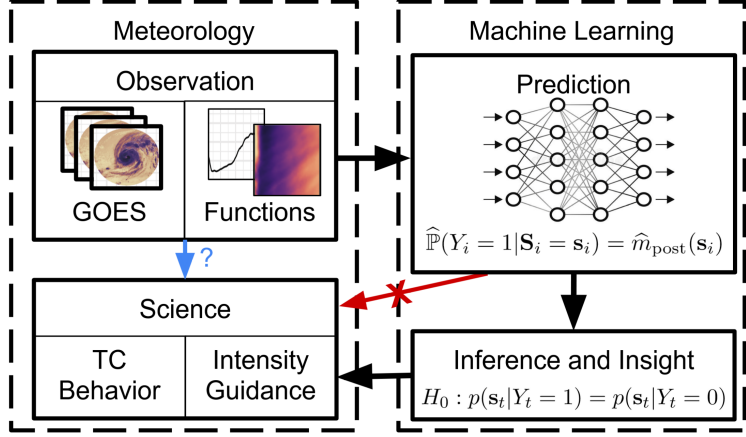
on one such ORB function—the radial profile—but our framework can generalize to jointly analyze multiple functional inputs. The radial profile, defined as $\overline{T}(r) = \frac{1}{2\pi}\int_0^{2\pi} T_b(r,\theta)d\theta$, captures the structure of cloud-top temperatures T_b as a function of radius r from the TC center and serves as an easily interpretable description of the depth and location of convection near the TC core [14, 18]. A single profile describes the instantaneous convective structure ($\mathbf{X}_t$; Figure 2, center) of a TC, whereas a 24-hour (48 observation) sequence of profiles provides a concise, human-readable summary of evolution over time; we call the latter sequence a *structural trajectory* ($\mathbf{S}_t = \{\mathbf{X}_t, \mathbf{X}_{t-1}, \ldots, \mathbf{X}_{t-48}\} \in \mathcal{S}$; Figure 2, right).

Contribution. We seek to answer the meteorological questions: "Are the sequences $\mathbf{S}_t$ leading up to RI events different from those leading up to non-RI events?" and, if so, "How do RI events differ from non-RI events over $\mathcal{S}$?" We pose this question as a hypothesis test comparing the distributions of $\mathbf{S}_t$ on $\mathcal{S}$ between classes (Equation 1). By recasting the hypothesis as a prediction problem, we are able to leverage artificial neural networks (ANNs) to **assess the relationship between labeled sequence data** $\{(\mathbf{S}_t, Y_t)\}_{t\geq 0}$ and TC behavior. Within this framework, shown in Figure 1, we not only detect the differing behavior of convection leading up to RI, but we are able to identify locations in $\mathcal{S}$ corresponding to RI, thus **bridging the gap from ML back to meteorological insight** (Figures 3,4). To our knowledge, this is the only principled framework that can detect distributional differences in sequences of images, and in addition provide local diagnostics which can identify spatio-temporal patterns commonly found in the lead-up to RI events.

2   Data

Inputs. Raw GOES image data are drawn from NOAA's MERGIR database [6], which contains 30-minute×4-km resolution IR imagery dating back to 2000 (Figure 2, left). The NOAA HURDAT2 database provides TC best track data: the best estimate of TC intensity and location produced after each season using all available data, including data not available in real time [13]. HURDAT2 has a 6-hour resolution.

RI Labels. We attach labels $Y_t \in \{0,1\}$ to structural trajectories $\mathbf{S}_t$ at each synoptic time (0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC) based on whether the TC was undergoing RI at time t (where $Y = 1$ indicates RI) according to Algorithm 2. The non-synoptic times are labeled $Y_t = 1$ if they fall between two consecutive synoptic $Y_t = 1$ observations, and $Y_t = 0$ otherwise.

Sample Size. We train our ANN on 33,873 labeled structural trajectories from 2000-2012 (further divided into 60% train/40% validation). We hold out 14,447 labeled trajectories from 2013-2019 as test data. When only labels

Figure 2: Structural Trajectories as Image Data. The raw data for each trajectory $\mathbf{S}_t$ is a sequence of cloud-top temperature images from GOES (T_b). We convert each of these into a radial profile ($\mathbf{X}_t$). The 24-hour sequence of radial profiles are combined into a *structural trajectory* or image ($\mathbf{S}_t$). These images are the high-dimensional inputs to $\widehat{m}_{\text{post}}(\mathbf{s}_t)$.
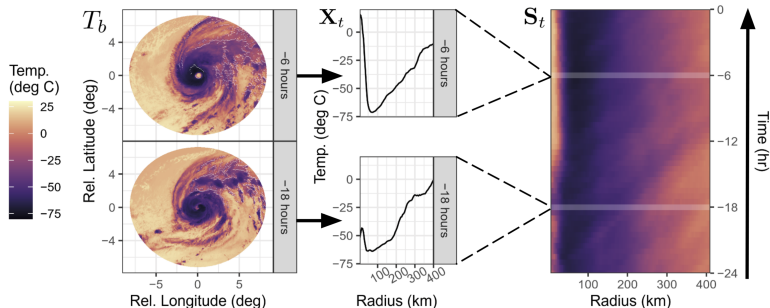
Y_t — and not trajectories $\mathbf{S}_t$ — are required ($\widehat{m}_{\text{seq}}$; Alg. 1(3)), we train on 31,203 synoptic best track entries from 1979-2012; we did not use data prior to 1979 due to reduced availability of satellite observations to support post-season analysis and thus the HURDAT2 best track database.

3 Methods

Testing Hypotheses. Our goal is to detect distributional differences in labeled sequence data $\{(\mathbf{S}_t, Y_t)\}_{t \geq 0}$, where the "labels" $Y_t \in \{0, 1\}$ are binary, and each covariate $\mathbf{S}_t \in \mathcal{S}$ is a sequence of functional quantities (our structural trajectory, Figure 2). Formally, we test the hypothesis

$$H_0 : p(\mathbf{s}_t | Y_t = 1) = p(\mathbf{s}_t | Y_t = 0) \text{ for all } \mathbf{s}_t \in \mathcal{S} \text{ versus} \tag{1}$$
$$H_1 : p(\mathbf{s}_t | Y_t = 1) \neq p(\mathbf{s}_t | Y_t = 0) \text{ for some } \mathbf{s}_t \in \mathcal{S}.$$

In our setting, there are two methodological challenges: (i) Equation 1 is a high-dimensional two-sample test for dependent sequence data, and (ii) to glean scientific insight, an answer of whether we reject/fail-to-reject H_0 is not enough: if we reject H_0, then we need to identify the *regions* in $\mathcal{S}$ where the two distributions differ.

From Tests to Prediction. For independent and identically distributed (IID) data $\{\mathbf{S}_i\}_{i=1}^n$, Equation 1 can be re-cast as a classification problem via Bayes Theorem, resulting in an equivalent test of $H_0 : \mathbb{P}(Y_i = 1 | \mathbf{s}_i) = \mathbb{P}(Y_i = 1)$, $\forall \, \mathbf{s}_i \in \mathcal{S}$. This idea is at the core of the recent classification accuracy [11] and regression-based [10] two-sample tests. The latter regression test goes beyond classification to provide *local* results, indicating where in $\mathcal{S}$ the two distributions differ. The key quantity to estimate in this procedure is the *class posterior* $m_{\text{post}}(\mathbf{s}_i) := \mathbb{P}(Y_i = 1 | \mathbf{s}_i)$. For this, one can leverage any probabilistic classification technique; hence converting a challenging inference and testing problem to a prediction problem. We reject H_0 if the test statistic $\lambda := \sum_{t \in \mathcal{V}} (\widehat{m}_{\text{post}}(\mathbf{s}_i) - \widehat{m}_{\text{prior}})^2$ is large on a held-out evaluation set $\mathcal{V}$ (not used in training). Here $\widehat{m}_{\text{prior}}$ is the proportion of $Y_i = 1$ cases in the train sample, and $\widehat{m}_{\text{post}}$ is the estimated class posterior.

Accounting for Dependence. However, existing high-dimensional tests do not extend to dependent sequence data. In particular, the regression test [10] relies on a permutation test and the assumption of exchangeability — an assumption that is generally violated for spatio-temporal processes. Here we propose a generalization of the regression test from IID to dependent data by replacing the permutation test with a *Monte Carlo (MC) test*. Our MC test estimates the label distribution $m_{\text{seq}}(y_{t-1}, \ldots, y_1) := \mathbb{P}(Y_t = 1 | Y_{t-1} = y_{t-1}, \ldots, Y_1 = y_1)$ under H_0 in place of permutation. The MC test is asymptotically valid (that is, it controls Type I errors), as long as the estimator $\widehat{m}_{\text{seq}}$ converges to m_{seq}. See Algorithm 1 for a full description of the testing procedure.

Local Importance. If we reject H_0, then the magnitude of the terms $\lambda_t := \widehat{m}_{\text{post}}(\mathbf{s}_t) - \widehat{m}_{\text{prior}}$ in the test statistic λ provide information as to which types of structural trajectories contributed the most to the rejection. Hence, we assign a local *importance score* $I(\mathbf{s}_t) := \text{sign}(\lambda_t) / \widehat{p}_t$ to each sequence $\mathbf{s}_t$, where $\widehat{p}_t$ is an approximate local p-value associated with λ_t, computed according to Algorithm 1. Positive scores $I(\mathbf{s}_t) \gg 0$ indicate $p(\mathbf{s}_t | Y_t = 1) > p(\mathbf{s}_t | Y_t = 0)$, and negative scores $I(\mathbf{s}_t) \ll 0$ the reverse. In other words, our test and diagnostics can identify 24-hour trajectories which are more common to the lead-up to RI events.

4 Results

Testing Hypotheses. We execute Algorithm 1, estimating m_{seq} with a Markov chain model with 48-hour memory and m_{post} with a convolutional neural network (CNN). The CNN classifier achieved $\sim 75\%$ balanced accuracy on test data; note that classifier performance affects the power of the test, but not validity. We find that the distributions of RI vs non-RI trajectories are different at significance level $\alpha = 0.01$.

Local Importance. Figure 3 (left) projects $\mathbf{S}_t \in \mathcal{S} \subset \mathbb{R}^{48 \times 120}$ onto their first two principal components. We use principal component analysis (PCA) for visualization only; $\widehat{m}_{\text{post}}$ is trained on the full structural trajectories. In Figure 3 (left), a single point represents an *entire* 24-hour trajectory. Temporally adjacent trajectories are highly correlated in PCA space, as evidenced by the "strands" in Figure 3. We divide the importance scores $I(\mathbf{s}_t)$ for each trajectory into 3 categories: negative ($I(\mathbf{s}) < -40$; blue, RI less likely), null ($-40 < I(\mathbf{s}) < 40$; gray), and positive ($I(\mathbf{s}) > 40$; red, RI more likely). Positive importance scores tend to have low values of the first principal component score, corresponding to cold trajectories whose temperatures are decreasing.

Scientific Insight. Six sample trajectories from two example TCs (Hurricanes Nicole [2016] and Jose [2017]) are labeled in Figure 3 (left) and displayed in Figure 3 (right). The PCA map clearly show two patterns: (i) core convection is stronger for low values of PCA1 and (ii) eye-eye wall structures are present for low values of PCA2.

We show the raw images $\mathbf{S}_i$ for each point in Figure 3 (right) along with Saffir-Simpson intensity categories. Out method is able to identify the structural patterns which lead up to RI: In all three examples with positive importance

scores (B, C, F), cloud-top temperatures are cold and getting colder. Trajectory F also shows the formation of an eye, during which time the TC intensified by 40 kt. The examples with negative importance scores (A, D, E) have warmer cloud-top temperatures with little indication of cooling in the core. **This analysis corroborates other studies that indicate strong, deepening convection is crucial for RI, particularly near the TC core where energy is most likely to result in strengthening surface winds.[17]**
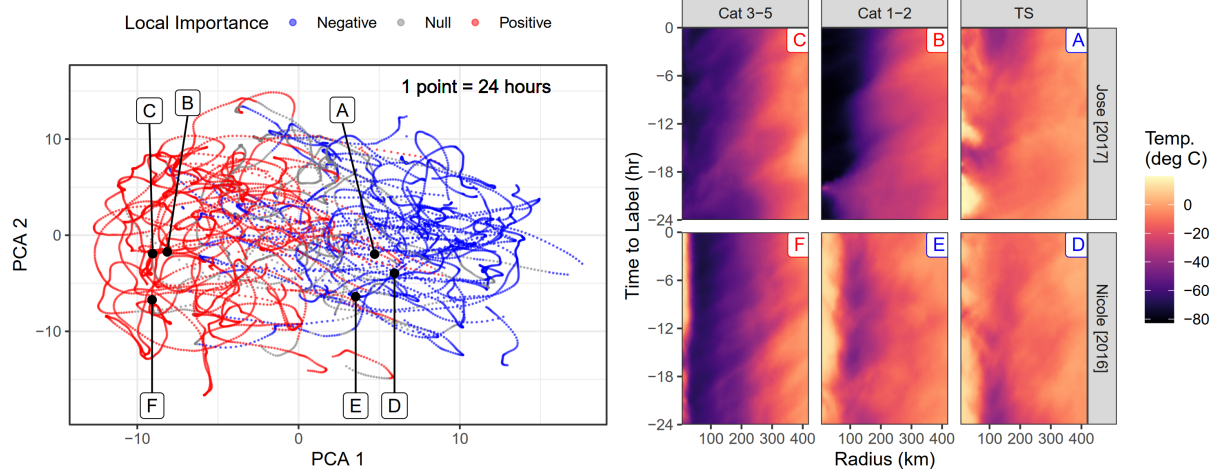
Figure 3: Differences in Distribution. (*Left*) PCA map, where each point, colored by its importance score, represents a 24h structural trajectory. In general, importance decreases from left to right, while current intensity increases from top to bottom. Note that PCA is only used for visualization purposes; the test is performed on the entire trajectory $\mathbf{S}_t$. (*Right*) Panels of selected trajectories or points for Hurricane Jose [2017] and Hurricane Nicole [2016]. The inset letter indicates the location and local importance in the PCA map. Positive cases tend to have cold cloud tops near the core, growing in coverage and depth of convection with time.

Case Study: Hurricane Nicole [2016]. We examine the full series $\{(\mathbf{S}_t, Y_t)\}_{t \geq 0}$ for Hurricane Nicole [2016] in Figure 4. When $I(\mathbf{s}_t) \gg 0$, we indeed find that TC intensity is increasing (center; points A, C, D). More interesting still, the trajectories *prior* to points A and C — the *onset* of RI — already begin to indicate that RI is likely; there are obviously markers of RI in the radial profiles which may be of use to forecasters. Future work will identify archetypal trajectories of this sort: $I(\mathbf{s}) \gg 0$ prior to intensification.
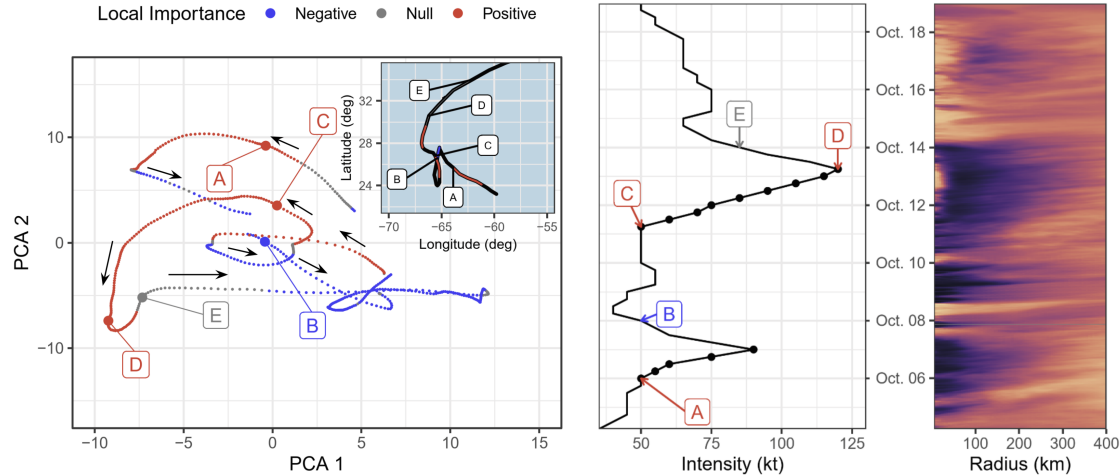
Figure 4: Case Study: Hurricane Nicole [2016]. (*Left*) Full set of trajectories $\{\mathbf{S}_t\}_{t \geq 0}$ for Hurricane Nicole [2016], with select trajectories labeled; again, each point represents a 24-hour trajectory. The physical track of the storm across the North Atlantic is shown inset, with the same trajectories labeled (here represented as segments of the track). (*Center*) Storm intensity over time, where solid points mark RI events. (*Right*) Full trajectory of Hurricane Nicole over 2 weeks. Note that the importance scores capture rapidly intensifying periods (A, C, D), collapsing convection (B), and the decay of the TC eye (E).

4

5 Conclusion

Our method successfully detects distributional differences in 24-hour "structural trajectories" (derived from GOES-IR images) leading up to periods of rapid intensification. A preliminary partitioning of trajectories by our proposed "local importance score" shows several known archetypal structural trajectories, laying the groundwork for forecasting tools. Further extending this framework to trajectories of other high-dimensional data promises to provide the insight necessary to understand and predict these increasingly deadly events.

6 Acknowledgments

References

[1] Mark DeMaria. A simplified dynamical system for tropical cyclone intensity prediction. *Monthly Weather Review*, 137(1):68–82, 2009.

[2] Mark DeMaria and John Kaplan. An updated statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic and Eastern North Pacific basins. *Weather and Forecasting*, 14(3):326–337, 1999.

[3] Mark DeMaria, Charles R. Sampson, John A. Knaff, and Kate D. Musgrave. Is tropical cyclone intensity guidance improving? *Bulletin of the American Meteorological Society*, 95(3):387–398, 2014.

[4] Vernon F. Dvorak. Tropical cyclone intensity analysis and forecasting from satellite imagery. *Monthly Weather Review*, 103(5):420–430, 1975.

[5] Liang Hu, Elizabeth A Ritchie, and J Scott Tyo. Short-term tropical cyclone intensity forecasting from satellite imagery based on the deviation angle variance technique. *Weather and Forecasting*, 35(1):285–298, 2020.

[6] John Janowiak, Bob Joyce, and Pingping Xie. NCEP/CPC L3 half hourly 4km global (60S - 60N) merged IR V1, 2020.

[7] John Kaplan and Mark DeMaria. Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Weather and forecasting*, 18(6):1093–1108, 2003.

[8] John Kaplan, Mark DeMaria, and John A. Knaff. A revised tropical cyclone rapid intensification index for the Atlantic and Eastern North Pacific basins. *Weather and Forecasting*, 25(1):220–241, 2010.

[9] John Kaplan, Christopher M Rozoff, Mark DeMaria, Charles R Sampson, James P Kossin, Christopher S Velden, Joseph J Cione, Jason P Dunion, John A Knaff, Jun A Zhang, et al. Evaluating environmental impacts on tropical cyclone rapid intensification predictability utilizing statistical models. *Weather and Forecasting*, 30(5):1374–1396, 2015.

[10] Ilmun Kim, Ann B Lee, and Jing Lei. Global and local two-sample tests via regression. *Electronic Journal of Statistics*, 13(2):5253–5305, 2019.

[11] Ilmun Kim, Aaditya Ramdas, Aarti Singh, and Larry Wasserman. Classification accuracy as a proxy for two-sample testing. *The Annals of Statistics*, 49(1):411–434, 2021.

[12] Philip J. Klotzbach, Steven G. Bowen, Roger Pielke, and Michael Bell. Continental U.S. hurricane landfall frequency and associated damage: Observations and future risks. *Bulletin of the American Meteorological Society*, 99(7):1359–1376, 2018.

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

[14] Trey McNeely, Ann B Lee, Kimberly M Wood, and Dorit Hammerling. Unlocking GOES: A statistical framework for quantifying the evolution of convective structure in tropical cyclones. *Journal of Applied Meteorology and Climatology*, 59(10):1671–1689, 2020.

[15] Timothy L Olander and Christopher S Velden. The advanced dvorak technique (adt) for estimating tropical cyclone intensity: Update and new capabilities. *Weather and Forecasting*, 34(4):905–922, 2019.

[16] Elizabeth A Ritchie, Kimberly M Wood, Oscar G Rodríguez-Herrera, Miguel F Piñeros, and J Scott Tyo. Satellite-derived tropical cyclone intensity in the North Pacific ocean using the deviation-angle variance technique. *Weather and forecasting*, 29(3):505–516, 2014.

[17] Robert Rogers, Paul Reasor, and Sylvie Lorsolo. Airborne doppler observations of the inner-core structural differences between intensifying and steady-state tropical cyclones. *Monthly Weather Review*, 141(9):2970–2991, 2013.

[18] Elizabeth R. Sanabia, Bradford S. Barrett, and Caitlin M. Fine. Relationships between tropical cyclone intensity and eyewall structure as determined by radial profiles of inner-core infrared brightness temperature. *Monthly Weather Review*, 142(12):4581–4599, 2014.

[19] Timothy J Schmit, Paul Griffith, Mathew M Gunshor, Jaime M Daniels, Steven J Goodman, and William J Lebair. A closer look at the ABI on the GOES-R series. *Bulletin of the American Meteorological Society*, 98(4):681–698, 2017.

[20] Kimberly M Wood and Elizabeth A Ritchie. A definition for rapid weakening of North Atlantic and Eastern North Pacific tropical cyclones. *Geophysical Research Letters*, 42(22):10–091, 2015.

A Algorithms

Algorithm 1: Test for distributional differences in labeled sequence data.

Require: type of test (MC=TRUE for MC test; MC=FALSE for permutation test); train data $\{Y_t, \mathbf{S}_t\}_{t \in \mathcal{T}_1}$ and regression method for estimating $m_{\text{post}}(\mathbf{s}) := \mathbb{P}(Y_t = 1 | \mathbf{S}_t = \mathbf{s})$; for MC test, additional train data $\{Y_t\}_{t \in \mathcal{T}_2}$ and regression method for estimating $m_{\text{seq}}(Y_{t-1}, \ldots, Y_{t-k}) := \mathbb{P}(Y_t = 1 | Y_{t-1}, \ldots, Y_{t-k})$; number of repetitions B; test points $\{\mathbf{s}_t\}_{t \in \mathcal{V}}$.

Ensure: p-value for testing $H_0 : p(\mathbf{s}_t | Y_t = 1) = p(\mathbf{s}_t | Y_t = 0)$, for all t and $\mathbf{s}_t$, local importance scores $\{I(\mathbf{s}_t)\}_{t \in \mathcal{V}}$

// Estimate underlying probability distributions

(1) Estimate $m_{\text{prior}} := \mathbb{P}(Y_t = 1)$ with class proportion $\widehat{m}_{\text{prior}} = \frac{1}{|\mathcal{T}_1|} \sum_{t \in \mathcal{T}_1} Y_t$.

(2) Regress Y_t on $\mathbf{S}_t$ using $\mathcal{T}_1$ to compute $\widehat{m}_{\text{post}}$.

(3)   **if** MC **then**
      Regress Y_t on $Y_{t-1}, \ldots, Y_{t-k}$ using $\mathcal{T}_2$ to compute $\widehat{m}_{\text{seq}}$.
   end if

// Compute test statistic and estimate its null distribution

(4) Compute test statistic $\lambda = \sum_{t \in \mathcal{V}} \lambda_t^2$, where $\lambda_t = \widehat{m}_{\text{post}}(\mathbf{s}_t) - \widehat{m}_{\text{prior}}$.

(5) **for** $b \in \{1, 2, \ldots, B\}$ **do**
  - Draw new train labels $\{\widetilde{Y}_t\}_{t \in \mathcal{T}_1}$ under H_0:
      if MC **then**
        Initialize $\widetilde{Y}_1, \ldots, \widetilde{Y}_k$ to random sequence.
        Draw sequence of length $100 \times k$ from $\widetilde{Y}_t \sim \text{Binom}(\widehat{m}_{\text{seq}}(\widetilde{Y}_{t-k}, \ldots, \widetilde{Y}_{t-1}))$ for burn-in.
        Draw new labels $\widetilde{Y}_t \sim \text{Binom}(\widehat{m}_{\text{seq}}(\widetilde{Y}_{t-k}, \ldots, \widetilde{Y}_{t-1}))$, for $t \in \mathcal{T}_1 \setminus \mathcal{T}_0$.
      else
        Permute original labels $\{Y_t\}_{t \in \mathcal{T}_1}$.
      end if
  - Regress $\widetilde{Y}_t$ on $\mathbf{S}_t$ using $\mathcal{T}_1$ to compute $\widehat{m}_{\text{post}}^{(b)}$.
  - Recompute test statistic $\widetilde{\lambda}^{(b)} = \sum_{t \in \mathcal{V}} \widetilde{\lambda}_t^{(b)2}$ where $\widetilde{\lambda}_t^{(b)} = \widehat{m}_{\text{post}}^{(b)}(\mathbf{s}_t) - \widehat{m}_{\text{prior}}$.

// Compute approximate p-values

(6) Compute p-values according to

$$\widehat{p}_{\text{global}} = \frac{1}{B+1}\left(1 + \sum_{b=1}^{B}\left(\widetilde{\lambda}^{(b)} > \lambda\right)\right) \text{ and } \widehat{p}_t = \frac{1}{B+1}\left(1 + \sum_{b=1}^{B}\left(|\widetilde{\lambda}_t^{(b)}| > |\lambda_t|\right)\right).$$

(7) Define local importance scores by

$$I(\mathbf{s}_t) = \frac{\text{sign}(\lambda_t)}{\widehat{p}_t}$$

return $\widehat{p}_{\text{global}}, \{I(\mathbf{s}_t)\}_{t \in \mathcal{V}}$

Algorithm 2: Identifying observations that fall within rapid intensification events.

Require: Sequence of contiguous intensity (i.e. maximum wind speed) observations $\{W_i\}_{t=1}^T$ for a storm and an intensity change threshold c.

Initialize: $\mathbf{Z} = \mathbf{0}_{T,T}, Y = \mathbf{0}_T$

for t *in* $1 : T - 1$ **do**

    $\Delta_{t,1} = W_{t+1} - W_t$; `// lead-1 intensity change`

    if $t \leq T - 4$ **then**

        $\Delta_{t,4} = \max(W_t, \ldots, W_{t+4}) - W_t$; `// lead-4 intensity change`

    end

end

$\mathcal{A} = \{t : \Delta_{t,4} \geq c\}$; `// 24-hour windows containing RI`

$\mathcal{B} = \{t : \Delta_{t,1} > 0\}$; `// 6-hour windows containing intensification`

for t *in* $1 : T - 4$ **do**

    if $t \in \mathcal{A}$ **then** `// label 24-hour RI window`

        $Z_{t,t}, \ldots, Z_{t,t+4} = 1$;

        for h *in* $4 : 1$ **do**

            if $t + h - 1 \notin \mathcal{B}$ **then** `// trim non-intensification from end of event`

                $Z_{t,t+h} = 0$;

            else

                break;

            end

        end

        for h *in* $0 : 3$ **do**

            if $t + h \notin \mathcal{B}$ **then** `// trim non-intensification from start of event`

                $Z_{t,t+h} = 0$;

            else

                break;

            end

        end

    end

end

for t *in* $1 : T$ **do**

    $Y_t = \max(Z_{t,1}, \ldots, Z_{t,T})$; `// points only need to be valid for one start`

end

Output: $\{Y_i\}_{t=1}^T$

Note: *The above algorithm identifies rapid intensification. To identify rapid weakening instead, reverse the input sequence $\{W_i\}_{t=1}^T$ at initialization, then reverse the output sequence $\{Y_i\}_{t=1}^T$ at output.*