JOHAN: A Joint Online Hurricane Trajectory and Intensity Forecasting Framework

Ding Wang
Michigan State University
East Lansing, Michigan, USA
wangdin1@msu.edu

Pang-Ning Tan
Michigan State University
East Lansing, Michigan, USA
ptan@msu.edu

ABSTRACT

Hurricanes are one of the most catastrophic natural forces with potential to inflict severe damages to properties and loss of human lives from high winds and inland flooding. Accurate long-term forecasting of the trajectory and intensity of advancing hurricanes is therefore crucial to provide timely warnings for civilians and emergency responders to mitigate costly damages and their life-threatening impact. In this paper, we present a novel online learning framework called JOHAN that simultaneously predicts the trajectory and intensity of a hurricane based on outputs produced by an ensemble of dynamic (physical) hurricane models. In addition, JOHAN is designed to generate accurate forecasts of the ordinal-valued hurricane intensity categories to ensure that their severity level can be reliably communicated to the public. The framework also employs exponentially-weighted quantile loss functions to bias the algorithm towards improving its prediction accuracy for high category hurricanes approaching landfall. Experimental results using real-world hurricane data demonstrated the superiority of JOHAN compared to several state-of-the-art learning approaches.

CCS CONCEPTS

• **Information systems** → **Data mining**; • **Applied computing** → *Forecasting*.

KEYWORDS

Online learning, Trajectory forecasting, Quantile regression

ACM Reference Format:
Ding Wang and Pang-Ning Tan. 2021. JOHAN: A Joint Online Hurricane Trajectory and Intensity Forecasting Framework. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '21), August 14–18, 2021, Virtual Event, Singapore.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3447548.3467400

1 INTRODUCTION

Hurricanes are tropical cyclones with maximum sustained wind speed (or intensity) of at least 64 knots or higher. The categorization of hurricane intensities in terms of their 1-minute maximum sustained wind speed, also known as the Saffir-Simpson scale, is shown in Table 1. High-intensity hurricanes, such as those of categories 3 or higher, have accounted for nearly 85% of hurricane-related damages according to the U.S. National Hurricane Center (NHC). For example, hurricane Harvey caused an estimated $125 billion of property damages and 107 confirmed deaths in 2017 [3, 21] while hurricane Florence caused $24.2 billion in damages and 54 deaths in 2018 [20, 23]. Due to its potential catastrophic impact, accurate long-term prediction of its path and intensity is critical to alert civilian population threatened by the imminent approach of a hurricane.

Storm category	Sustained Winds (kt)	Types of Damage Due to Hurricane Winds
-	<33	Tropical depression
-	34-63	Tropical storm
1	64-82	Very dangerous winds will produce some damage
2	83-95	Extremely dangerous winds will cause extensive damage
3	96-112	Devastating damage will occur
4	113-136	Catastrophic damage will occur
5	≥ 137	Catastrophic damage will occur

Table 1: Categorization of tropical cyclone intensity based on the Saffir-Simpson hurricane wind scale (SSHWS)[24]

Despite its importance, hurricane prediction is a notoriously hard problem due to the complex physical mechanisms governing the dynamics of a tropical cyclone, which include factors such as sea surface temperature and vertical wind shear. To address this issue, numerous physics-based models [14] have been developed over the years to provide forecast guidance on the trajectory and intensity of impending hurricanes. These models would generate their forecasts by solving the mathematical equations governing the physics of the atmosphere and ocean coupling. Despite the advances in these models for trajectory prediction, little improvements have been achieved for intensity prediction.

In recent years, there have been growing interests in applying machine learning techniques to improve the performance of hurricane prediction tasks [1, 17, 18, 25]. However, many of the existing works were developed for forecasting hurricane trajectories only, with very few of them designed to predict intensities or both. Furthermore, accurate forecasting of its ordinal category is often more important than the wind speed itself when communicating the severity of an impending hurricane to the public. Indeed, a prediction error of 60 mph may seem trifle for a category 0 tropical storm but is significant if a category 5 hurricane at 160mph was incorrectly predicted as a category 2 storm at 100mph. Furthermore, as shown in Table 2, current methods were mostly limited to short-range predictions (24 hours or less) using historical observations as predictors. These methods are also mostly trained in a batch learning mode, and thus, are incapable of modeling the non-stationary nature of hurricane trajectories and their intensities.

KDD '21, August 14–18, 2021, Virtual Event, Singapore

ACM ISBN 978-1-4503-8332-5/21/08…$15.00
https://doi.org/10.1145/3447548.3467400

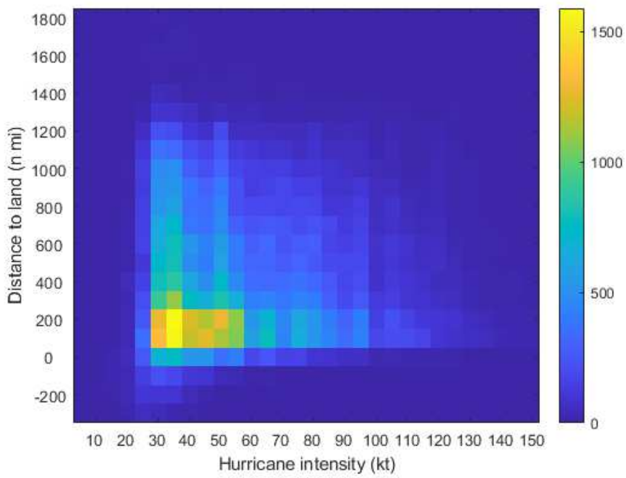


Figure 1: Heat map showing the relationship between hurricane intensity and its distance to nearest U.S. coastline. Negative distance indicate that the hurricane has made landfall.

To overcome these limitations, this paper presents a novel online learning framework called JOHAN (**J**oint **O**nline **H**urricane Tr**A**jectory and I**N**tensity Prediction) for long-term forecasting (up to 48 hours) of hurricane trajectory and intensity. By using an online learning approach, our model can be efficiently updated to fit new observations while adapting to concept drifts present in the non-stationary data. JOHAN employs outputs from an ensemble of dynamical (physical) models such as U.S. Navy Global Environmental Model (NAVGEM) [15] and Hurricane Weather Research and Forecasting system (HWRF) [14] to generate its forecasts. These dynamical models are designed to simulate future atmospheric conditions from the current conditions. However, the skills of these ensemble members (i.e., dynamical models) may vary from one hurricane to another. By training the model in an online fashion, our framework will be able to take into account the varying skills of the ensemble members over time.

There are several reasons for developing an algorithm that can predict the hurricane trajectory and intensity jointly. First, previous studies have shown the importance of using trajectory information for intensity prediction [8, 9]. As an illustration, Figure 1 shows the relationship between hurricane intensity and its distance to the nearest U.S. coastline using 6-hourly hurricane data between 1851 to 2020 from NHC. The plot suggests that hurricanes with higher intensities are more likely to be distributed at shorter distances to the coastline. This phenomenon has been observed in other recent studies [26]. For example, Wang and Toumi have noted that the distance at which the tropical cyclone hits its peak intensity has grown closer to the coastline, decreasing at a rate of 30km per decade. This suggests the utility of using location information from the trajectory to help improve the prediction accuracy for high category hurricanes. Furthermore, the plot also shows that most of the hurricanes lose their intensity after landfall, which is not surprising as their energy dissipates rapidly on land, causing a sharp drop in its intensity.

Finally, it is worth noting that not all predictions are equal in importance. Accurate prediction of high category hurricanes with potential for landfall is more critical than lower category hurricanes whose projected path is heading away from the coastline. This is because hurricanes approaching landfall have potential to cause more damaging impacts to civilian population from storm surges, high winds, inland flooding, etc. Unfortunately, high category hurricanes also tend to occur less frequently than the lower category ones, which leads to a class imbalance problem. To overcome these challenges, JOHAN uses an exponentially-weighted quantile loss function to bias its algorithm towards predicting more accurately high intensity hurricanes that are approaching landfall.

2 RELATED WORKS

Due to the complexity of modeling the dynamics of tropical cyclones, there have been growing interests in developing machine learning and deep learning techniques for the hurricane prediction problem. Table 2 reviews some of the existing works, which can be categorized in terms of the input features used, learning approaches, target variable to be predicted, and the forecast horizon (i.e., maximum lead time of the forecast).

First, existing methods typically use the historical trajectory data, climate/meteorological data, or outputs from physical models as input features for their prediction models. While historical data are more suitable for short-range (nowcasting) predictions [17], their performance tend to be poor since they do not capture the current and future environmental conditions that affect the hurricane's path and intensity. Methods utilizing meteorological data are usually based on deep learning techniques, such as generalized advesarial networks (GAN) [22] and convolutional LSTM (ConvLSTM) [16]. While these works are promising, their prediction errors are still relatively large since the models are typically trained using coarse-scale images (e.g., $0.5^\circ \times 0.5^\circ$). Methods that use physical model outputs tend to generate more reliable long-term forecasts since the dynamical models consider the current environmental conditions when simulating their future forecast scenarios [11, 25].

Second, most of the existing works focused on the trajectory prediction task only even though intensity forecasting is the more challenging problem. Although the regression and deep learning methods can be applied to hurricane intensity forecasting problem, they are not designed for predicting ordinal-valued categories, unlike the approach proposed in this paper. Third, current methods mostly employ a batch learning approach to train their models. This may not be feasible nor effective in an operational forecast environment, when a new hurricane is continuously tracked and the model needs to be periodically updated (say every 3 to 6 hours) to reflect the new trajectory and intensity information.

Finally, recent works have focused on using deep learning and online learning approaches for hurricane trajectory prediction problems. For deep learning, [17] used sparse RNN with a flexible topology to generate hurricane trajectory predictions. [12] proposed a Long Short-Term Memory (LSTM) network to predict typhoon tracks using historical observation data from 1949 to 2011 while [1] employed RNN over a grid system to handle the non-linearity of hurricane trajectory forecasting. For online learning, [25] presented a multi-lead time forecasting framework for hurricane trajectory prediction. They showed that ensemble forecasting using outputs from physical models significantly outperform batch methods such as LSTM trained on historical trajectory data.

Reference	Method	Input Features	Prediction Task	Lead Time (Forecast horizon)	Learning Mode
DeMaria et al. 2005	Linear regression	Historical data	Intensity	Multi-step (72 hrs)	Batch
Moradi Kordmahalleh et al. 2016	RNN	Historical data	Trajectory	Multi-step (12 hrs)	Batch
Cox et al. 2018	Association rule	Historical data	Trajectory	Multi-step	Batch
Mudigonda et al. 2017	ConvLSTM	Atmospheric data	Trajectory	Multi-step	Batch
Gao et al. 2018	LSTM	Historical data	Trajectory	Multi-step (72 hrs)	Batch
Alemany et al. 2019	RNN	Historical data	Trajectory	Multi-step (120 hrs)	Batch
Rüttgers et al. 2019	GAN	Atmospheric image	Trajectory	Single step (6 hrs)	Batch
Kim et al. 2019	ConvLSTM	Climate data	Trajectory	Multi-step (15 hrs)	Batch
Eslami et al. 2019	CNN	Physical model outputs	Trajectory & intensity	Multi-step	Batch
Wang et al. 2020	Online linear	Physical model outputs	Trajectory	Multi-step (48 hrs)	Online
Giffard-Roisin et al. 2020	Neural network	Historical data and atmospheric image	Trajectory	Multi-step (24 hrs)	Batch

Table 2: Literature review of recent works on tropical cyclone prediction.

3 PROBLEM STATEMENT

Our goal is to design an online learning framework for joint prediction of hurricane trajectory and its intensity (both ordinal category and continuous values). At first glance, knowing the category of a hurricane does not appear to add any new information about the hurricane intensity since the former is derived from latter value (see Table 1). Nevertheless, the information is indeed useful as it is possible for the predicted category error to be small even though the error in predicting the maximum sustained wind speed is large. For example, given a category 5 hurricane with maximum sustained wind speed of 140 knots. A model that predicts its intensity to be 100 knots will have a lower error than one that predicts its intensity to be 200 knots; yet, the former has a larger category error (since 100 knots is a category 3 hurricane) compared to the latter, which still predicts the correct category. Furthermore, a category 2 hurricane at 95 knots predicted as 115 knots has a lower intensity prediction error compared to one predicted as 60 knots even though the former has a larger error since the category 2 cyclone is incorrectly predicted as a major category 4 storm rather than category 1, which is closer to it. Thus, leveraging both ordinal category and real-valued intensity information can help improve the prediction framework.

Consider a set of hurricanes, $\{h_1, h_2, \dots, h_C\}$, ordered by their start times. Assuming there are n_i data points (time steps) associated with hurricane h_i, then $N = \sum_{i=1}^{C} n_i$ is the total number of time steps in the hurricane dataset. Let $\mathbb{X} = \{\mathcal{X}^1, \mathcal{X}^2, \dots, \mathcal{X}^N\}$ be the set of trajectory forecasts generated by an ensemble of dynamical models, where each $\mathcal{X}^t$ corresponds to the hurricane trajectory forecasts generated at time step t. Similarly, the intensity forecasts generated by the ensemble members can be denoted as $\tilde{\mathbb{X}} = \{\tilde{\mathbf{X}}^1, \tilde{\mathbf{X}}^2, \dots, \tilde{\mathbf{X}}^N\}$. Let $\tilde{n}_i$ be the accumulated number of data points from hurricane h_1 to h_i, i.e. $\tilde{n}_i = \sum_{j=1}^{i} n_j$. Thus, $\{(\mathcal{X}^j, \tilde{\mathbf{X}}^j) \mid \tilde{n}_{i-1} < j \le \tilde{n}_i\}$ is the set of trajectory and intensity data points associated with hurricane h_i. Assume T is the forecast horizon, i.e., maximum lead-time of the forecasting task. For each time step t, let $\mathcal{X}^t \in \mathbb{R}^{2\times m_t \times T}$ be the hurricane trajectory forecasts (latitude and longitude), where m_t is the number of ensemble member (dynamic model) forecasts available at time step t. The ensemble member trajectory forecasts for lead time τ at time t is denoted as $\mathbf{X}^{t,\tau} \in \mathbb{R}^{2\times m_t}$, with the corresponding ground truth location $\mathbf{y}^{t,\tau} \in \mathbb{R}^2$. Let $\mathbf{Y}^1, \mathbf{Y}^2, \dots, \mathbf{Y}^N$ be the ground truth locations for all lead times at each time step t, where $\mathbf{Y}^t \in \mathbb{R}^{2\times T}$. Similarly, let $\tilde{\mathbf{X}}^t \in \mathbb{R}^{T\times \tilde{m}_t}$ be the hurricane intensity forecasts at time step t, where $\tilde{m}_t$ is the number of ensemble member forecasts available at time step t. The ensemble member forecasts for lead time τ at time t is denoted as $\tilde{\mathbf{x}}^{t,\tau} \in \mathbb{R}^{\tilde{m}_t}$, with the corresponding ground truth intensity value $\tilde{y}^{t,\tau} \in \mathbb{R}$. Let $\tilde{\mathbf{y}}^1, \tilde{\mathbf{y}}^2, \dots, \tilde{\mathbf{y}}^N$ be the true intensity values for N time steps, where each $\tilde{\mathbf{y}}^t = [\tilde{y}^{t,1}\ \tilde{y}^{t,2} \cdots \tilde{y}^{t,T}]^T$ is a vector of intensity values for all lead times at time step t. Furthermore, let $\hat{y}^{t,\tau}$ and $\hat{\mathbf{y}}^t$ be the corresponding intensity categories associated with the real-valued intensities in $\tilde{y}^{t,\tau}$ and $\tilde{\mathbf{y}}^t$, respectively. The transformation from hurricane intensity values to their corresponding intensity categories is based on an ordered list of boundary values $-\infty < b_1 < \cdots < b_5 < \infty$. Specifically, the predicted intensity value $z^{t,\tau}$ is assigned to category $\hat{y}^{t,\tau}$ if $b_{\hat{y}^{t,\tau}} < z^{t,\tau} \le b_{\hat{y}^{t,\tau}+1}$.

4 METHODOLOGY

At each time step t, we use the set of ensemble member forecasts for trajectory $\mathbf{X}^{t,\tau} \in \mathbb{R}^{2\times m_t}$ and intensity $\tilde{\mathbf{x}}^{t,\tau} \in \mathbb{R}^{\tilde{m}_t}$ to generate the trajectory and intensity predictions for lead time τ. The real-valued trajectory prediction $\mathbf{z}^{t,\tau} \in \mathbb{R}^2$ and intensity prediction $\tilde{z}^{t,\tau} \in \mathbb{R}$ are computed by linear predictors as follows:

$$\begin{aligned} \mathbf{z}^{t,\tau} &= f^{t,\tau}(\mathbf{X}^{t,\tau}) = \mathbf{X}^{t,\tau}\mathbf{w}^{t,\tau} \\ \tilde{z}^{t,\tau} &= \tilde{f}^{t,\tau}(\tilde{\mathbf{x}}^{t,\tau}) = \tilde{\mathbf{x}}^{t,\tau}\tilde{\mathbf{w}}^{t,\tau} \end{aligned} \tag{1}$$

where $\mathbf{w}^{t,\tau} \in \mathbb{R}^{m_t}, \tilde{\mathbf{w}}^{t,\tau} \in \mathbb{R}^{\tilde{m}_t}$ are the learned weight vectors associated with ensemble member forecasts for the trajectory and intensity models, respectively. The weight vectors are updated simultaneously in an online fashion whenever new observation data becomes available. One major challenge in using the ensemble member forecasts is that a significant proportion of the ensemble members may not generate any forecasts at a given time step t, which is why the number of ensemble members, m_t, varies from one hurricane to another. This is known as the varying feature length problem [25], which can be addressed using the weight re-normalization technique described in [25].

4.1 Proposed JOHAN Framework

The novelty of JOHAN is its ability to jointly predict the hurricane trajectory and intensity. The framework consists of a pair of weight updating components for hurricane trajectory and intensity prediction. Both components employ an exponentially-weighted quantile loss to improve their prediction performance for close-to-land hurricanes and high category hurricanes.

To learn the tasks jointly, our framework is trained to minimize the following objective function in an online learning fashion:

$$
\begin{aligned}
\mathcal{L} = \quad & \mathcal{L}_{tra}(\Theta, \xi) + \mathcal{L}_{int}(\tilde{\Theta}, \tilde{\xi}) \\
\text{s.t.} \quad & \xi^{t,\tau} = \begin{cases} g(\tilde{y}^{t,\tau}), \text{if } \tilde{y}^{t,\tau} \text{ is available} \\ g(\tilde{z}^{t,\tau}), \text{otherwise} \end{cases} \\
& \tilde{\xi}^{t,\tau} = \begin{cases} \tilde{g}(\mathbf{y}^{t,\tau}), \text{if } \mathbf{y}^{t,\tau} \text{ is available} \\ \tilde{g}(\mathbf{z}^{t,\tau}), \text{otherwise} \end{cases}
\end{aligned} \tag{2}
$$

where $\mathcal{L}_{tra}$ corresponds to the loss function for trajectory prediction while $\mathcal{L}_{int}$ is the loss for intensity forecasting. Θ and $\tilde{\Theta}$ are the model parameters associated with the hurricane trajectory and intensity prediction tasks, respectively. The quantile parameters ξ and $\tilde{\xi}$ are needed to bias the model towards predicting more accurately hurricanes that are close to the coastline or those with high categories. Unlike traditional quantile loss, ξ and $\tilde{\xi}$ are not user-specified hyperparameters but are automatically updated in an online fashion. Specifically, the quantile loss terms are updated to reflect the significant threat of a hurricane using the functions $g(\cdot)$ and $\tilde{g}(\cdot)$. Recall that $\mathbf{y}^{t,\tau}$ is the true hurricane location and $\tilde{y}^{t,\tau}$ is the true intensity at time t for lead time τ. However, since $\mathbf{y}^{t,\tau}$ and $\tilde{y}^{t,\tau}$ may not available during model update, we use the model predictions $\mathbf{z}^{t,\tau}$ and $\tilde{z}^{t,\tau}$ to approximate them when calculating the quantile parameters. Details of the quantile functions are given in Section 4.1.3. The objective function can be solved using standard quadratic programming solvers.

4.1.1 $\mathcal{L}_{tra}$ with Distance Quantile Regression. As hurricanes can cause severe damages in civilian populated areas, it is imperative to accurately identify hurricanes that are approaching landfall. Therefore, we would like to bias the model towards learning hurricanes with potential to strike the land. This can be done by encouraging hurricane forecasts that are more likely to make landfall. The possibility of hurricane landfall can be measured by the distance between its current location to the nearest coastline. Specifically, we introduce a distance loss decomposition to evaluate the model performance by taking into account its predicted distance to the coastline. For every ground truth location $\mathbf{y}$, we can find its corresponding projected point $\mathbf{p}$ to the nearest coastline. A unit normal vector to the coastline can be calculated as $\mathbf{n} = \frac{\mathbf{p}-\mathbf{y}}{\|\mathbf{p}-\mathbf{y}\|}$. Given a predicted location $\mathbf{z}$, its distance loss is defined as $\mathbf{d} = \mathbf{z} - \mathbf{y}$. The distance loss vector $\mathbf{d}$ can be decomposed into a parallel, $\mathbf{d}_{\|} = \mathbf{d} \cdot \mathbf{n}$, and a perpendicular component, $\mathbf{d}_{\perp} = \mathbf{d} - \mathbf{d}_{\|}$, as shown in Figure 2. With the definition of distance loss decomposition, the square loss $\frac{1}{2}\|\mathbf{z} - \mathbf{y}\|_2^2$ can be expressed equivalently as follows

$$
\begin{aligned}
\mathcal{L}_{tra} = \quad & \frac{1}{2}\left(\zeta^2 + \zeta^{*2} + (\mathbf{z} - \mathbf{y})_{\perp}^2\right) \\
\text{s.t.} \quad & (\mathbf{z} - \mathbf{y})_{\|} = \zeta - \zeta^*, \\
& \zeta \geq 0, \zeta^* \geq 0
\end{aligned} \tag{3}
$$

In order to encourage predictions with shorter distances to the coastline, Eqn. (3) can be further extended to accommodate the quantile loss in Eqn. (4). Note that Eqn. (4) is equivalent to Eqn. (3) by setting the quantile parameter ξ to 0.5.

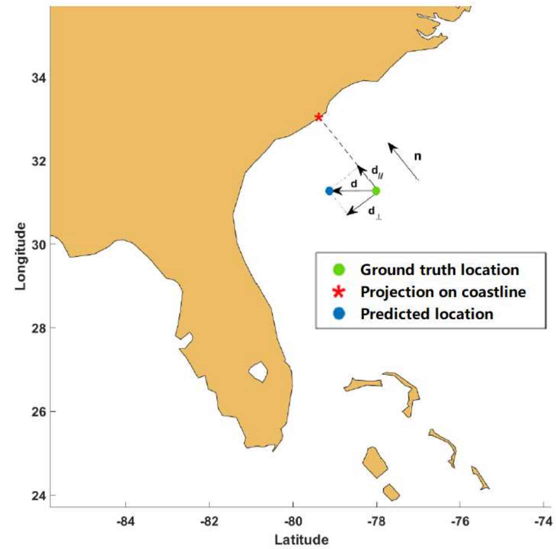


Figure 2: Decomposition of the distance loss vector. The green circle is the true location and the red star is its projected nearest coastline. The unit vector n points in the direction towards the land. The blue circle is the predicted location. The vector directed from the green circle to the blue circle is the distance loss vector d, which can be decomposed into a parallel $\mathbf{d}_{\|}$ and a perpendicular component $\mathbf{d}_{\perp}$.

$$
\begin{aligned}
\mathcal{L}_{tra} = \quad & (1-\xi)\zeta^2 + \xi\zeta^{*2} + \frac{1}{2}(\mathbf{z} - \mathbf{y})_{\perp}^2 \\
\text{s.t.} \quad & (\mathbf{z} - \mathbf{y})_{\|} = \zeta - \zeta^*, \\
& \zeta \geq 0, \zeta^* \geq 0
\end{aligned} \tag{4}
$$

We assume that the weight vectors $\mathbf{w}^{t,\tau}$ in Eqn. (1) can be decomposed into the following factors:

$$
\begin{aligned}
& \mathbf{w}^{t,\tau} = \mathbf{u}^t + \mathbf{v}^{t,\tau} \\
\text{s.t.} \quad & \mathbf{1}_{m_t}^T \mathbf{u}^t = 1, \ \ \mathbf{1}_{m_t}^T \mathbf{v}^{t,\tau} = 0
\end{aligned} \tag{5}
$$

where $\mathbf{u}^t$ is the shared weight vectors for all lead times while $\mathbf{v}^{t,\tau}$ is adjustment to the weight vectors associated with the different lead times τ. For brevity, we denote $\mathbf{W}^t = [\mathbf{w}^{t,1}, \mathbf{w}^{t,2}, \cdots, \mathbf{w}^{t,T}]$ as the weight matrix for all T lead times at time step t. To extend the preceding formulation to an online multi-lead time forecasting setting, the weight matrix $\mathbf{W}^t$ is updated by minimizing the following objective function at each time step t:

$$
\begin{aligned}
\mathcal{L}_{tra} = \quad & \sum_{\tau=1}^{T} \delta^{t,\tau}\gamma^{\tau}\left((1-\xi)\zeta^{t,\tau 2} + \xi\zeta^{*t,\tau 2} + \frac{1}{2}(\mathbf{z}^{t,\tau} - \mathbf{y}^{t,\tau})_{\perp}^2\right) \\
+ \quad & \frac{\omega}{2}\sum_{\tau=1}^{T-1}\left\|\mathbf{w}^{t,\tau+1} - \mathbf{w}^{t,\tau}\right\|^2 + \frac{\mu}{2}\left\|\mathbf{u}^t - \mathbf{u}^{t-1}\right\|^2 \\
+ \quad & \frac{\nu}{2}\sum_{\tau=1}^{T}\left\|\mathbf{v}^{t,\tau} - \mathbf{v}^{t-1,\tau}\right\|^2 + \frac{\eta}{2}\sum_{\tau=1}^{T}\left\|\mathbf{v}^{t,\tau}\right\|^2 \\
\text{s.t.} \quad & \forall\, t, \tau: \ \mathbf{1}_{m_t}^T \mathbf{u}^t = 1, \ \mathbf{1}_{m_t}^T \mathbf{v}^{t,\tau} = 0, \\
& (\mathbf{z}^{t,\tau} - \mathbf{y}^{t,\tau})_{\|} = \zeta^{t,\tau} - \zeta^{*t,\tau}, \\
& \zeta^{t,\tau} \geq 0, \zeta^{*t,\tau} \geq 0
\end{aligned} \tag{6}
$$

where $\delta^{t,\tau}$ is an indicator function whose value is 1 if $\mathbf{X}^{t,\tau}$ and $\mathbf{y}^{t,\tau}$ values are both available; otherwise its value is 0. In the objective function, the first term represents the forecast errors for all the lead times. The quantile parameter ξ determines the importance of making location predictions with shorter distance to coastlines. The hyperparameter γ determines the relative importance of making accurate predictions at different lead times. The second term ensures that the estimated model parameters would vary smoothly at different lead times, thus preserving the temporal autocorrelation of the predicted intensities. The third and fourth terms guarantee that the shared weight vector $\mathbf{u}^t$ and lead time adjustment weight vectors $\mathbf{v}^{t,\tau}$ are close to their values at previous time step. The last term penalizes large values in the lead time adjustment weight vectors. ω, μ, ν, η are hyperparameters that determine the relative importance of each term in the objective function. If $\xi \approx 1$, then the parallel distance loss will be ignored if $(\mathbf{z}^{t,\tau} - \mathbf{y}^{t,\tau})_{\parallel} > 0$. This means that optimizing $\mathcal{L}_{tra}$ will lead to models that are more biased toward predicting closer distance to the coastline.

4.1.2 $\mathcal{L}_{int}$ with Quantile Ordinal Regression. Our goal is to also generate accurate long range predictions of hurricane real-valued intensity and category. Here, we use the ϵ-insensitive loss to measure the intensity prediction error. Compared to mean square loss, the ϵ-insensitive loss is more robust as it provides a margin of tolerance ϵ [2, 10] when learning the regression function.

$$\begin{aligned} \epsilon - \text{insensitive loss:}\quad \mathcal{L}_{int} = \quad & \zeta + \zeta^* \\ \text{s.t.}\quad & z - y \le \epsilon + \zeta \\ & z - y \ge -\epsilon - \zeta^* \\ & \zeta \ge 0, \zeta^* \ge 0 \end{aligned} \tag{7}$$

Second, in order to communicate the severity of an impending hurricane to the public, accurate prediction of its category is just as important as the wind speed itself. As noted in Section 3, intensity prediction alone is insufficient because it ignores the effect of its prediction error on the predicted category. Therefore, we introduce an ordinal loss to ensure the model is focused more on data points located near the boundary between two ordinal categories. Analogous to the loss function defined for support vector ordinal regression [5], the ordinal loss function is defined as follows:

$$\begin{aligned} \text{Ordinal loss:}\quad \mathcal{L}_{int} = \quad & \zeta + \zeta^* \\ \text{s.t.}\quad & z - b_{\hat{y}+1} \le -1 + \zeta \\ & z - b_{\hat{y}} \ge 1 - \zeta^* \\ & \zeta \ge 0, \zeta^* \ge 0 \end{aligned} \tag{8}$$

Our intensity prediction loss can be measured by combining Eqns. (7) and (8). In addition, to penalize models that incorrectly predict high category hurricanes, the formulation can be extended to accommodate the quantile loss as $(1-\xi)\zeta + \xi\zeta^*$.

$$\begin{aligned} \mathcal{L}_{int} = \quad & (1-\xi)\zeta + \xi\zeta^* \\ \text{s.t.}\quad & z - b_{\hat{y}+1} \le -1 + \zeta \\ & z - b_{\hat{y}} \ge 1 - \zeta^* \\ & z - y \le \epsilon + \zeta \\ & z - y \ge -\epsilon - \zeta^* \\ & \zeta \ge 0, \zeta^* \ge 0 \end{aligned} \tag{9}$$

Similar to the trajectory model, we assume that the intensity weight vector $\tilde{\mathbf{w}}^{t,\tau}$ can be decomposed into the following factors:

$$\begin{aligned} & \tilde{\mathbf{w}}^{t,\tau} = \tilde{\mathbf{u}}^t + \tilde{\mathbf{v}}^{t,\tau} \\ \text{s.t.}\quad & \mathbf{1}_{\tilde{m}_t}^T \tilde{\mathbf{u}}^t = 1, \mathbf{1}_{\tilde{m}_t}^T \tilde{\mathbf{v}}^{t,\tau} = 0 \end{aligned} \tag{10}$$

where $\tilde{\mathbf{u}}^t$ is the shared weight vectors for all lead times while $\tilde{\mathbf{v}}^{t,\tau}$ is adjustment to the weight vectors associated with the different lead times τ. For brevity, we denote $\tilde{\mathbf{W}}^t = [\tilde{\mathbf{w}}^{t,1}, \tilde{\mathbf{w}}^{t,2}, \cdots, \tilde{\mathbf{w}}^{t,T}]$ as the weight matrix for all T lead times at time step t.

Putting it together, the weight matrix $\tilde{\mathbf{W}}^t$ for intensity prediction is trained to minimize the following objection function:

$$\begin{aligned} \mathcal{L}_{int} = \quad & \frac{1}{2}\sum_{\tau=1}^{T} \tilde{\delta}^{t,\tau}\tilde{\gamma}^{\tau}\left((1-\tilde{\xi})\zeta^{t,\tau} + \tilde{\xi}\zeta^{*t,\tau}\right) \\ + \quad & \frac{\tilde{\omega}}{2}\sum_{\tau=1}^{T-1}\left\|\tilde{\mathbf{w}}^{t,\tau+1} - \tilde{\mathbf{w}}^{t,\tau}\right\|^2 + \frac{\tilde{\mu}}{2}\left\|\tilde{\mathbf{u}}^t - \tilde{\mathbf{u}}^{t-1}\right\|^2 \\ + \quad & \frac{\tilde{\nu}}{2}\sum_{\tau=1}^{T}\left\|\tilde{\mathbf{v}}^{t,\tau} - \tilde{\mathbf{v}}^{t-1,\tau}\right\|^2 + \frac{\tilde{\eta}}{2}\sum_{\tau=1}^{T}\left\|\tilde{\mathbf{v}}^{t,\tau}\right\|^2 \\ \text{s.t.}\quad & \forall\, t,\tau:\ \mathbf{1}_{\tilde{m}_t}^T \tilde{\mathbf{u}}^t = 1,\ \mathbf{1}_{\tilde{m}_t}^T \tilde{\mathbf{v}}^{t,\tau} = 0, \\ & \tilde{z}^{t,\tau} - b_{\hat{y}^{t,\tau}+1} \le -1 + \zeta^{t,\tau} \\ & \tilde{z}^{t,\tau} - b_{\hat{y}^{t,\tau}} \ge 1 - \zeta^{*t,\tau} \\ & \tilde{z}^{t,\tau} - \tilde{y}^{t,\tau} \le \epsilon + \zeta^{t,\tau} \\ & \tilde{z}^{t,\tau} - \tilde{y}^{t,\tau} \ge -\epsilon - \zeta^{*t,\tau} \\ & \zeta^{t,\tau} \ge 0, \zeta^{*t,\tau} \ge 0 \end{aligned} \tag{11}$$

where $\tilde{\delta}^{t,\tau}$ is an indicator function whose value is 1 if $\tilde{\mathbf{x}}^{t,\tau}$ and $\tilde{y}^{t,\tau}$ values are both available; otherwise its value is 0. In the objective function, the first term represents the forecast errors for all the lead times. The hyperparameter $\tilde{\xi}$ determines the importance of making accurate predictions for high category hurricanes. The meanings of other terms in the objective function are similar to Equation (6).

4.1.3 Quantile Parameter Update. As described in the previous section, the quantile parameters are updated in an online fashion. In general, we want the quantile parameter for trajectory prediction to be large if the hurricane category is high; and the quantile parameter for intensity prediction to be large if the hurricane location is close to coastline. In our framework, we use a sigmoid function $\sigma(x) = 1/(1+e^{-x})$ to determine the quantile parameters.

For $\mathcal{L}_{tra}$, the parameter $\xi^{t,\tau}$ is calculated from the function $g(\mathbf{x})$ given in Eqn. (2) as follows:

$$g(\mathbf{x}) = \begin{cases} 0.5, & \text{for } x < \theta \\ \sigma([x-\theta]/c), & \text{otherwise} \end{cases} \tag{12}$$

where $\mathbf{x}$ is the ground truth or predicted intensity with current weight vector, θ is a hyperparameter that decides when the quantile loss is not needed, and c is a scaling factor. When the hurricane intensity is low, $g(\mathbf{x}) = 0.5$, and thus, the first loss term in $\mathcal{L}_{tra}$ reduces to the squared loss function. For high intensity hurricanes, $g(\mathbf{x}) \approx 1$. In this case, the framework gives higher weights for models that predict locations with shorter distance to the coastline.

For $\mathcal{L}_{int}$, the parameter $\tilde{\xi}^{t,\tau}$ is calculated from the function $\tilde{g}(\cdot)$ defined in Eqn. (2) as follows:

$$\tilde{g}(\mathbf{x}) = \begin{cases} 0.5, & \text{for } d_{coast}(\mathbf{x}) > \tilde{\theta} \\ \sigma([\tilde{\theta} - d_{coast}(\mathbf{x})]/\tilde{c}), & \text{otherwise} \end{cases} \tag{13}$$

where $\mathbf{x}$ is either the ground truth or predicted location for the current weight vector, $d_{coast}(\cdot)$ is a function that computes distance to the nearest coastline, $\tilde{\theta}$ is a hyperparameter that determines when the quantile loss is not needed, and $\tilde{c}$ is the scaling factor. When the hurricane is far from the coastline, we have $g(\mathbf{x}) = 0.5$. The first loss term of $\mathcal{L}_{int}$ is thus reduced to an ℓ_1-loss. When the hurricane is very close the land, $g(\mathbf{x}) \approx 1$. In this case, the framework assigns higher weights to models that predict higher intensities.

Error propagation is a challenge for multi-lead time forecasting. It is insufficient to update $\mathbf{W}^t$ and $\tilde{\mathbf{W}}^t$ from $\mathbf{W}^{t-1}$ and $\tilde{\mathbf{W}}^{t-1}$ alone as the previous weights are outdated without using the new observation. To address this problem, we apply the backtracking and restart strategy described in [25]. The overall framework with backtracking and restart strategy is illustrated in Figure 3. The pseudocode for the proposed framework is described in Algorithm 1. At each time step t, the newly available ground truth value can be determined as $\{\mathbf{y}^{t-T,T}, \mathbf{y}^{t-T+1,T-1}, \ldots, \mathbf{y}^{t-1,1}\}$ for trajectory forecasts and $\{\tilde{y}^{t-T,T}, \tilde{y}^{t-T+1,T-1}, \ldots, \tilde{y}^{t-1,1}\}$ for intensity forecasts. Therefore, to make all the previous learned weight vectors up to date, we updated the weight vectors from time step $t - T$ until $t - 1$ with quantile parameters generated by Eqn. (12) and (13). Then, the multi-lead time predictions for both trajectory and intensity can be produced given the ensemble of model outputs and updated model weight vectors.

```
Input: Θ, Θ̃, ξ, ξ̃
Output: Model parameters w, w̃ and forecasts z, z̃
Initialize: w = 1_m/m, w̃ = 1_m/m;
for t = 1,2, ..., N do
    if t is the first time step of a trajectory then
        Extract u^t from w, ũ^t from w̃
        Normalize:
          u^t ← u^t/|u^t|, v^{t,τ} ← 0, ũ^t ← ũ^t/|ũ^t|, ṽ^{t,τ} ← 0
    end
    Observe y^t, ỹ^t
    /* Backtracking and restart step */
    for t' = t − T, t − T + 1, ..., t − 1 do
        Trajectory predictions
          z^{t',τ} = f^{t',τ}(X^{t,τ}) = X^{t',τ} w^{t',τ}
        Intensity predictions z̃^{t',τ} = f̃^{t,τ}(x̃^{t',τ}) = x̃^{t',τ} w̃^{t',τ}
        Calculate ξ^{t,τ} and ξ̃^{t,τ} using Eq. 12 and 13
        Update u^{t'+1}, v^{t'+1,τ} by minimizing L_tra
        Update ũ^{t'+1}, ṽ^{t'+1,τ} by minimizing L_int
    end
    /* Prediction step */
    for τ = 1, 2, ..., T do
        Compute w^{t,τ} and w̃^{t,τ} for all lead times
        Trajectory predictions z^{t,τ} = f^{t,τ}(X^{t,τ}) = X^{t,τ} w^{t,τ}
        Intensity predictions z̃^{t,τ} = f̃^{t,τ}(x̃^{t,τ}) = x̃^{t,τ} w̃^{t,τ}
    end
    if t is the last time step of a trajectory then
        Substitute u^t back into the full vector w
        Substitute ũ^t back into the full vector w̃
    end
end
```

Algorithm 1: Proposed JOHAN Framework

5 EXPERIMENTS

We performed experiments using real-world hurricane trajectory and intensity data from various sources. The ground truth hurricane trajectory and intensity data along with the official forecasts are obtained from the National Hurricane Center (NHC) website[1], while the ensemble member forecasts are obtained from the Hurricane Forecast Model Output website at University of Wisconsin-Milwaukee[2]. We collected 6-hourly hurricane trajectory and intensity data from the year 2012 to 2020, which contains 336 tropical cyclones. Each tropical cyclone has an average length of 21.9 time steps (data points), which gives a total of 7364 data points. There are 27 trajectory forecast models and 21 intensity forecast models used in our experiments, which are a subset of the models used by NHC in the preparation of their official forecasts. The data from 2012 to 2017 (208 tropical cyclones) are used for training and validation, while those from 2018 to 2020 (128 tropical cyclones) are used for testing.

5.1 Baseline and Evaluation Metrics

We compared JOHAN against the following baseline methods:

(1) **Ensemble mean**: This method computes the mean value over all ensemble members for each given lead time.
(2) **Persistence**: This method assumes that the intensity and moving speed of the hurricane at the next time step is the same as current time step.
(3) **Passive-Aggressive(PA)** [7]: This is a well-known online regression algorithm.
(4) **ORION** [27]: This is an online multi-task learning algorithm for ensemble forecasting.
(5) **OMuLeT** [25]: This is a recently developed online learning algorithm for ensemble forecasting.
(6) **NHC**: This corresponds to the official forecasts generated by NHC, which is the gold standard.

For a fair comparison, the baseline methods (PA, ORION, OMuLeT, and JOHAN) also apply the backtracking and restart strategy to update their weights. Hyperparameters of the methods were tuned by minimizing the following mean distance error (MDE) for trajectory forecasts and macro-averaged mean absolute error (MAE) for intensity forecasts on the validation set.

$$\text{MDE} = \frac{1}{N} \sum_{t,\tau} \left[dis\left(\mathbf{z}^{t,\tau}, \mathbf{y}^{t,\tau}\right)\right] \tag{14}$$

$$\text{macro-MAE} = \frac{1}{6} \sum_{i=0}^{5} \left(\frac{1}{N_i} \sum_{\hat{y}^{t,\tau}=i} \left| \hat{z}^{t,\tau} - \hat{y}^{t,\tau} \right| \right) \tag{15}$$

[1] https://www.nhc.noaa.gov
[2] http://derecho.math.uwm.edu/models

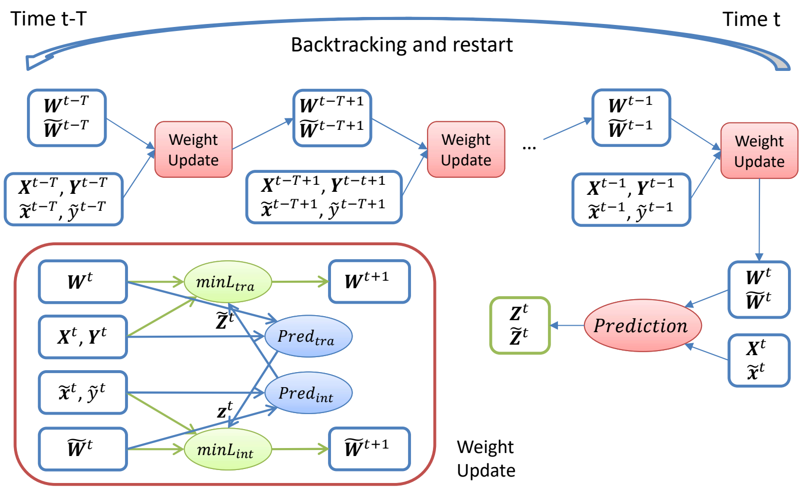


Figure 3: Proposed JOHAN framework.

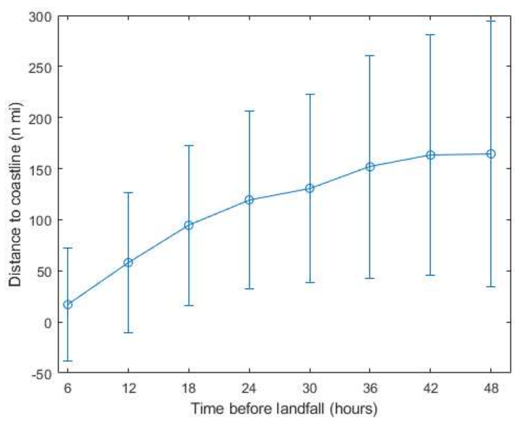


Figure 4: Distance to coastline for hurricanes from year 2012 to 2017 at different time before landfall.

where N is total number of trajectory forecasts in the validation set and N_i is the number of data points of category i in the validation set. We use MDE on the trajectory test data to evaluate the location prediction error, and MAE on the intensity test data to evaluate the error in both real- and ordinal-valued predictions. We also use F1-score to evaluate accuracy of the ordinal category predictions.

For JOHAN framework, the hyperparameters $\mathbf{\Theta}$ and $\tilde{\mathbf{\Theta}}$ are tuned with the fixed quantile parameters $\xi = \tilde{\xi} = 0.5$. The threshold θ for the function $g(\cdot)$ is set to 34 knots (kt), which is the lower bound for intensity of a tropical storm. As there are very few time steps with intensity more than 64 kt, the hyperparameter c is calculated by solving $g(\mathbf{x}) = \sigma([x - \theta]/c) = 0.9$ with $x = 64$ kt. For the function $\tilde{g}(\cdot)$, Figure 4 shows the distance to coastline for landfall hurricanes at different hours before landfall. The threshold $\tilde{\theta}$ is set to 300 nautical miles (n mi) where the hurricanes are unlikely to strike the land in 48 hours. The hyperparameter $\tilde{c}$ is calculated by solving $\tilde{g}(\mathbf{x}) = \sigma([\tilde{\theta} - d_{coast}(\mathbf{x})]/\tilde{c}) = 0.9$ with $x = 200$ n mi which is the distance from the coastline for hurricanes that might reach landfall after 24 hours. The source code and data used in our experiments are available at https://github.com/cqwangding/JOHAN.

5.2 Experimental Results

Table 3 summarizes the trajectory and intensity forecast errors of JOHAN and other baseline methods with lead times from 12 hours to 48 hours. For trajectory prediction, there are several interesting conclusions that can be drawn. First, the performance of ensemble mean is comparable to the NHC official forecast, which validates the advantage of using an ensemble of physical model outputs as predictors. Second, persistence performs much worse than other baselines. This is reasonable since persistence assumes that the moving speed of hurricane is unchanged. Third, OMuLeT and JOHAN generate comparable results and both outperform other baselines, including the official forecasts from NHC. This is not surprising as both methods employ similar strategies to update their trajectory prediction models. However, as will be shown below, OMuLeT is inferior to JOHAN in terms of its trajectory prediction error for hurricanes within 200 n mi from the coastline.

For intensity forecasts, the performance of ensemble mean is significantly worse than NHC as shown in Table 3. This is not surprising as intensity forecasting is still a very challenging problem for the dynamical and statistical models in the ensemble. Second, both persistence and PA perform poorly, much more so than other methods. Fourth, JOHAN generally has significantly better performance especially at longer lead times with results comparable to the official forecasts of NHC. This is impressive considering the fact that the ensemble members used in JOHAN is only a subset of the models used by NHC to generate their official forecasts.

JOHAN is designed to maximally identify threatening hurricanes with potential to strike landfall. Table 4 summarizes the trajectory and intensity forecast errors for hurricanes within 200 n mi to coastline with an intensity of at least 64 kt for different prediction methods at varying lead times (from 12 to 48 hours). The relative performance of all the baseline methods is similar to Table 3. The

	Trajectory error (in n mi)				Intensity error (in kt)			
Method	12	24	36	48	12	24	36	48
Ensemble Mean	23.30	36.34	50.22	65.03	6.742	8.692	9.899	11.036
Persistance	34.84	88.89	155.87	229.63	7.717	13.797	18.246	22.102
PA	23.30	36.34	50.23	64.80	6.547	10.563	13.294	15.268
ORION	23.37	36.36	50.21	65.00	5.792	7.941	9.388	10.551
OMuLeT	**22.20**	**34.94**	**48.07**	62.10	**5.632**	7.923	9.285	10.513
JOHAN	22.25	35.01	48.13	**62.08**	5.732	**7.700**	**8.948**	**10.026**
NHC	24.59	38.49	52.17	65.74	5.019	7.730	8.959	10.140

Table 3: Trajectory and intensity forecast errors for different methods at varying lead times from 12 to 48 hours.

	Trajectory error (in n mi)				Intensity error (in kt)			
Method	12	24	36	48	12	24	36	48
Ensemble Mean	15.80	28.47	41.21	52.09	13.274	17.325	20.246	20.382
Persistance	34.28	87.62	159.38	228.91	13.121	22.547	29.194	32.762
PA	15.81	28.48	41.22	51.98	8.765	16.042	24.275	24.656
ORION	16.02	28.61	41.29	52.15	**8.483**	**13.171**	16.806	17.555
OMuLeT	15.33	28.28	39.65	49.67	9.064	14.435	17.562	18.317
JOHAN	**15.28**	**28.15**	**39.28**	**49.06**	8.963	13.367	**16.270**	**16.554**
NHC	16.69	29.47	42.83	54.07	7.962	13.585	16.097	17.587

Table 4: Trajectory and intensity forecast errors for hurricanes within 200 n mi to coastline with intensity at least 64 kt for different methods at varying lead times from 12 to 48 hours.

	F1-score						
Category	macro-F1	0	1	2	3	4	5
Ensemble Mean	0.390	0.917	0.449	0.309	0.305	0.168	0.190
Persistence	0.396	0.859	0.327	0.247	0.234	0.330	0.378
PA	0.415	0.902	0.437	0.310	0.274	0.339	0.226
ORION	0.491	0.922	**0.530**	**0.377**	0.349	**0.418**	0.350
OMuLeT	0.494	**0.923**	0.502	0.367	**0.386**	0.380	0.404
JOHAN	**0.499**	0.922	0.499	0.357	0.385	0.380	**0.450**
NHC	0.540	0.924	0.554	0.411	0.390	0.462	0.502

Table 5: Comparison of F1-score, precision and recall for various hurricane intensity forecasting methods at different categories.

trajectory prediction of JOHAN outperforms all other baselines for near land hurricanes. This suggests that JOHAN is capable of utilizing the relationship between high intensity hurricanes and distance to the coastline to improve its prediction. For intensity predictions, JOHAN still maintains the best predictive performance at longer lead times.

In addition, we also evaluated the hurricane category prediction for the different methods by computing their F1-scores. The results are shown in Table 5. First, ensemble mean performs poorly for high category predictions, which is not surprising as the different ensemble members are not equally skillful. Second, the overall macro-F1 performance of persistence is similar to ensemble mean and much worse than other baselines. Third, PA is better than ensemble mean but still worse than both ORION and OMuLeT. The macro-F1 score of JOHAN is slightly higher than ORION and OMuLeT, though JOHAN has better performance for the higher categories. Table 6 summarizes the F1-scores for hurricanes within 200 n mi from the coastline, which clearly demonstrates the superiority of JOHAN compared to other baselines for high category hurricanes.

5.3 Ablation Study

A key aspect of JOHAN is its ability to update the quantile parameter in an online fashion. To investigate the advantages of utilizing a varying quantile parameter, we consider two variations of our

	F1-score						
Category	macro-F1	0	1	2	3	4	5
Ensemble Mean	0.378	0.919	0.414	0.234	0.318	0.188	0.195
Persistance	0.394	0.854	0.309	0.182	0.236	0.389	0.392
PA	0.454	0.899	0.412	0.281	0.298	0.395	0.436
ORION	**0.505**	**0.927**	**0.509**	**0.330**	0.375	0.431	0.459
OMuLeT	0.493	0.923	0.465	0.290	**0.402**	0.435	0.445
JOHAN	0.496	0.923	0.462	0.274	**0.402**	**0.447**	**0.467**
NHC	0.516	0.920	0.511	0.328	0.362	0.442	0.534

Table 6: Comparison of F1-score, precision and recall for hurricanes within 200 n mi to coastline with various hurricane intensity forecasting methods at different categories.

framework. JOHAN-NQ removes the quantile loss for the first term in Eqn. (6) and (11) and reduced them to squared loss and ℓ_1-loss. JOHAN-Q uses a fixed quantile parameter for all its weight updates.

Table 7 summarizes the trajectory and intensity forecast errors for the two variations of JOHAN framework while Table 8 summarizes their corresponding forecast errors for near land high intensity hurricanes. It is clear that increasing the fixed quantile parameters would lead to better performance for near land, high intensity hurricanes (Table 8) but at the expense of decreasing overall prediction accuracy (Table 7). JOHAN, with its varying quantile parameters, manages to maintain consistently accurate predictions in both scenarios.

	Trajectory (n mi)				Intensity (kt)			
Method	12	24	36	48	12	24	36	48
JOHAN	22.25	35.01	48.13	62.08	5.732	7.700	8.948	10.026
JOHAN-NQ	22.20	34.94	48.07	62.10	5.700	7.654	8.880	9.840
JOHAN-Q(ξ=0.6)	22.21	34.94	48.08	62.07	5.696	7.647	8.877	9.872
JOHAN-Q(ξ=0.7)	22.23	34.94	48.07	62.06	5.704	7.659	8.897	9.939
JOHAN-Q(ξ=0.8)	22.27	34.96	48.08	62.09	5.722	7.704	8.964	10.064
JOHAN-Q(ξ=0.9)	22.33	35.01	48.11	62.14	5.749	7.781	9.073	10.248

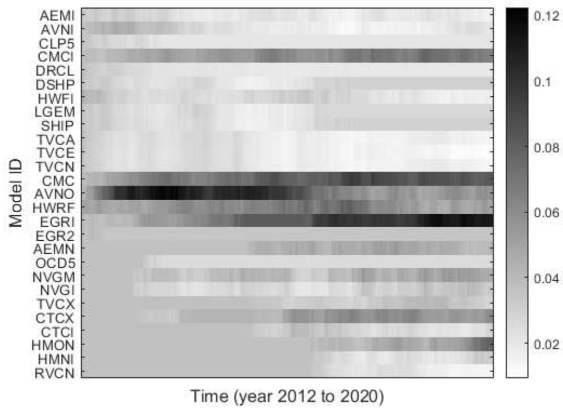
Table 7: Trajectory and intensity forecast errors for different variations of JOHAN.

	Trajectory (n mi)				Intensity (kt)			
Method	12	24	36	48	12	24	36	48
JOHAN	15.28	28.15	39.28	49.06	8.963	13.367	16.270	16.554
JOHAN-NQ	15.33	28.28	39.65	49.67	9.176	13.642	16.791	16.854
JOHAN-Q(ξ=0.6)	15.30	28.23	39.53	49.50	9.047	13.488	16.537	16.743
JOHAN-Q(ξ=0.7)	15.28	28.18	39.43	49.33	8.985	13.398	16.346	16.630
JOHAN-Q(ξ=0.8)	15.27	28.15	39.35	49.16	8.950	13.362	16.256	16.606
JOHAN-Q(ξ=0.9)	15.27	28.14	39.29	49.00	8.986	13.412	16.256	16.671

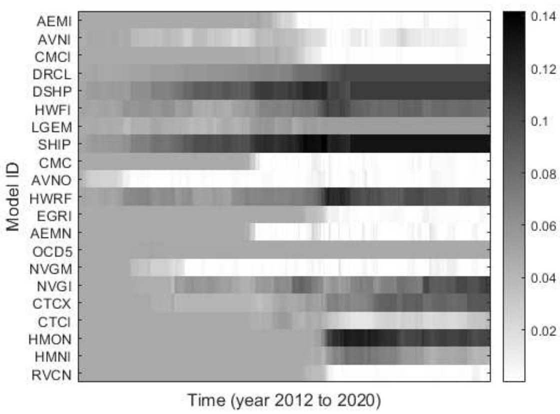
Table 8: Trajectory and intensity forecast errors for hurricanes within 200 n mi to coastline and at least 64 kt using different variations of JOHAN.

5.4 Comparison of Model Weights

The time-varying model weights generated by JOHAN are shown in Figure 5 for trajectory and intensity forecasts, respectively. Despite the shared information between the trajectory and intensity components of the framework, it is clear that the best models for trajectory prediction are not exactly the same as the best models for intensity prediction. This result agrees with previous findings in [11]. For example, AVNO was found to be one of the best models for trajectory prediction but its weight for intensity prediction is close to 0. Compared to trajectory forecasting, hurricane intensity forecasting is still a very challenging problem [4].



(a) Trajectory model weights



(b) Intensity model weights

Figure 5: Trajectory and intensity model weights in JOHAN changes over time.

6 CONCLUSION

This paper presents a novel framework called JOHAN for predicting long-range hurricane trajectory and intensity simultaneously. The framework employs a novel time-varying quantile loss function to improve its accuracy in predicting high category hurricanes with the potential for near landfall. Experimental results confirmed the efficacy of JOHAN, especially in terms of accurately predicting near landfall, high category hurricanes.

7 ACKNOWLEDGMENTS

This research is partially supported by the National Science Foundation under grant #IIS-2006633. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

REFERENCES

[1] Sheila Alemany, Jonathan Beltran, Adrian Perez, and Sam Ganzfried. 2019. Predicting hurricane trajectories using a recurrent neural network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 468–475.

[2] Mariette Awad and Rahul Khanna. 2015. Support vector regression. In *Efficient learning machines*. Springer, 67–80.

[3] Eric S Blake and David A Zelinsky. 2018. Tropical Cyclone Report: Hurricane Harvey. *National Hurricane Center* (2018).

[4] D Brown. 2017. Tropical Cyclone Intensity Forecasting: Still a Challenging Proposition. *National Hurricane Center* (2017).

[5] Wei Chu and S Sathiya Keerthi. 2007. Support vector ordinal regression. *Neural computation* 19, 3 (2007), 792–815.

[6] Taylor S Cox, Calvin SH Hoi, Carson K Leung, and Caden R Marofke. 2018. An accurate model for hurricane trajectory prediction. In *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, Vol. 2. IEEE, 534–539.

[7] Koby Crammer, Ofer Dekel, Joseph Keshet, Shai Shalev-Shwartz, and Yoram Singer. 2006. Online passive-aggressive algorithms. *Journal of Machine Learning Research* 7, Mar (2006), 551–585.

[8] Mark DeMaria and John Kaplan. 1994. A statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic basin. *Weather and Forecasting* 9, 2 (1994), 209–220.

[9] Mark DeMaria, Michelle Mainelli, Lynn K Shay, John A Knaff, and John Kaplan. 2005. Further improvements to the statistical hurricane intensity prediction scheme (SHIPS). *Weather and Forecasting* 20, 4 (2005), 531–543.

[10] Harris Drucker, Christopher J Burges, Linda Kaufman, Alex Smola, and Vladimir Vapnik. 1996. Support vector regression machines. *Advances in neural information processing systems* 9 (1996), 155–161.

[11] Ebrahim Eslami, Yunsoo Choi, Yannic Lops, and Alqamah Sayeed. 2019. A Deep Learning Driven Improved Ensemble Approach for Hurricane Forecasting. In *2019 ESIP Winter Meeting*.

[12] Song Gao, Peng Zhao, Bin Pan, Yaru Li, Min Zhou, Jiangling Xu, Shan Zhong, and Zhenwei Shi. 2018. A nowcasting model for the prediction of typhoon tracks based on a long short term memory neural network. *Acta Oceanologica Sinica* 37, 5 (2018), 8–12.

[13] Sophie Giffard-Roisin, Mo Yang, Guillaume Charpiat, Christina Kumler Bonfanti, Balázs Kégl, and Claire Monteleoni. 2020. Tropical cyclone track forecasting using fused deep learning from aligned reanalysis data. *Frontiers in Big Data* 3 (2020), 1.

[14] SG Gopalakrishnan, Qingfu Liu, Timothy Marchok, Dmitry Sheinin, Naomi Surgi, Robert Tuleya, Richard Yablonsky, and Xuejin Zhang. 2010. Hurricane Weather Research and Forecasting (HWRF) model scientific documentation. *L Bernardet Ed* 75 (2010), 7655.

[15] Timothy F Hogan, Ming Liu, James A Ridout, Melinda S Peng, Timothy R Whitcomb, Benjamin C Ruston, Carolyn A Reynolds, Stephen D Eckermann, Jon R Moskaitis, Nancy L Baker, et al. 2014. The navy global environmental model. *Oceanography* 27, 3 (2014), 116–125.

[16] Sookyung Kim, Hyojin Kim, Joonseok Lee, Sangwoong Yoon, Samira Ebrahimi Kahou, Karthik Kashinath, and Mr Prabhat. 2019. Deep-hurricane-tracker: Tracking and forecasting extreme climate events. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 1761–1769.

[17] Mina Moradi Kordmahalleh, Mohammad Gorji Sefidmazgi, and Abdollah Homaifar. 2016. A Sparse Recurrent Neural Network for Trajectory Prediction of Atlantic Hurricanes. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*. ACM, 957–964.

[18] Raymond ST Lee and James NK Liu. 2000. Tropical cyclone identification and tracking system using integrated neural oscillatory elastic graph matching and hybrid RBF network track mining techniques. *IEEE Transactions on Neural Networks* 11, 3 (2000), 680–689.

[19] Mayur Mudigonda, Sookyung Kim, Ankur Mahesh, Samira Kahou, Karthik Kashinath, Dean Williams, Vincen Michalski, Travis O'Brien, and Mr Prabhat. 2017. Segmenting and tracking extreme climate events using neural networks. In *Deep Learning for Physical Sciences (DLPS) Workshop, held with NIPS Conference*.

[20] NCEI. 2019. *Assessing the U.S. Climate in 2018*. National Centers for Environmental Information (NCEI).

[21] NHC. 2018. Costliest U.S. Tropical Cyclones Tables Updated. *National Hurricane Center* (2018).

[22] Mario Rüttgers, Sangseung Lee, Soohwan Jeon, and Donghyun You. 2019. Prediction of a typhoon track using a generative adversarial network and satellite images. *Scientific reports* 9, 1 (2019), 1–15.

[23] Stacy Stewart and Robbie Berg. 2019. Tropical Cyclone Report: Hurricane Florence. *National Hurricane Center* (2019).

[24] Harvey Thurm Taylor, Bill Ward, Mark Willis, and Walt Zaleski. 2010. The Saffir-Simpson hurricane wind scale. *Atmospheric Administration: Washington, DC, USA* (2010).

[25] Ding Wang, Boyang Liu, Pang-Ning Tan, and Lifeng Luo. 2020. OMuLeT: Online Multi-Lead Time Location Prediction for Hurricane Trajectory Forecasting. In *Proceedings of 34th AAAI Conference on Artificial Intelligence*.

[26] Shuai Wang and Ralf Toumi. 2021. Recent migration of tropical cyclones toward coasts. *Science* 371, 6528 (2021), 514–517.

[27] Jianpeng Xu, Pang-Ning Tan, and Lifeng Luo. 2014. ORION: Online Regularized multI-task regressiON and its application to ensemble forecasting. In *ICDM, 2014*. IEEE, 1061–1066.