OMuLeT: Online Multi-Lead Time Location Prediction for Hurricane Trajectory Forecasting

Ding Wang,[1] Boyang Liu,[1] Pang-Ning Tan,[1] Lifeng Luo[2]
[1]Department of Computer Science & Engineering, Michigan State University
[2]Department of Geography, Michigan State University
{wangdin1, liuboya2, ptan, lluo}@msu.edu

Abstract

Hurricanes are powerful tropical cyclones with sustained wind speeds ranging from at least 74 mph (for category 1 storms) to more than 157 mph (for category 5 storms). Accurate prediction of the storm tracks is essential for hurricane preparedness and mitigation of storm impacts. In this paper, we cast the hurricane trajectory forecasting task as an online multi-lead time location prediction problem and present a framework called OMuLeT to improve path prediction by combining the 6-hourly and 12-hourly forecasts generated from an ensemble of dynamical (physical) hurricane models. OMuLeT employs an online learning with restart strategy to incrementally update the weights of the ensemble model combination as new observation data become available. It can also handle the varying dynamical models available for predicting the trajectories of different hurricanes. Experimental results using the Atlantic and Eastern Pacific hurricane data showed that OMuLeT significantly outperforms various baseline methods, including the official forecasts produced by the U.S. National Hurricane Center (NHC), by more than 10% in terms of its 48-hour lead time forecasts.

Introduction

Hurricanes are one of the most powerful storms on Earth that have the potential to cause devastating losses and destruction along their paths. For example, the Galveston Hurricane of 1900 is considered the deadliest hurricane in United States, responsible for at least 8000 deaths (Blake, Landsea, and Gibney 2011). In 2005, Hurricane Katrina took away more than 1500 lives and caused at least $108 billion of property damages (Blake, Landsea, and Gibney 2011). Given their severe impact, accurate long-range prediction of hurricane tracks is critical to give ample time for emergency response teams to take appropriate actions that will minimize property damages and loss of human lives. Towards this end, dynamical models such as NOAA's Hurricane Weather Research and Forecasting (HWRF) system and U.S. Navy Global Environmental Model (NAVGEM) have been widely used as the primary tool for hurricane forecasting (Vickery, Skerlj, and Twisdale 2000; Davis et al. 2008). Although the skills of these models have improved steadily over the years, the forecast errors and variability among the model predictions still increase with lead time.

Ensemble forecasting seeks to better represent the range of forecast uncertainties by combining outputs generated by multiple dynamical models (Zhang and Krishnamurti 1999; Gneiting and Raftery 2005; Leutbecher and Palmer 2008). Each dynamical model can produce one or more ensemble member outputs by perturbing its initial conditions or model parameters. The ensemble mean or median are commonly used as the deterministic forecast from the ensemble. These estimates assume that every member is equally skillful, and thus, their predictions should be weighted equally. Such an assumption may not be realistic due to the inherent differences in the way the ensemble member outputs are generated. Thus in an operational forecast environment when such ensemble forecasts are issued on a regular basis, the weight of each member must be established based on their accuracy in predicting the hurricane tracks (DeMaria et al. 2005). However, determining the appropriate weights is not a trivial task as the skills of the models may vary from one hurricane to another. To overcome this challenge, the primary goal of this paper is to develop an online hurricane trajectory forecasting framework that can dynamically update the weights of the ensemble members based on their past and recent performance when verified against observations.

In the United States, the National Hurricane Center (NHC) is responsible for monitoring and providing official forecasts of hurricane trajectories and their intensities to the public. With a set of dynamical model forecasts as guidance, the official NHC forecasts are produced based on the experience and judgment of the forecasters. A secondary goal of this paper is to investigate the feasibility of using an online learning approach to generate forecasts that are equally or more skillful than the official forecasts reported by NHC.

In addition to the dynamical models, various statistical (DeMaria et al. 2005; Reich and Fuentes 2007) and hybrids of statistical-dynamical models (Kim and Webster 2010; Vecchi et al. 2011) have been developed. There has also been growing interest in recent years to apply machine learning methods to the hurricane trajectory forecasting problem (Lee and Liu 2000; Kordmahalleh, Sefidmazgi, and Homaifar 2016; Alemany et al. 2019). However, there are sev-

			Ensemble member forecasts at different lead times, τ					
			$\tau = 1$ (24 hrs)			$\tau = 2$ (48 hrs)		
Hurricane h_i	Time t	NHC Best track, $\mathbf{y}^{i,t}$	AEMI $\mathbf{x}_1^{i,t,1}$	AEMN $\mathbf{x}_2^{i,t,1}$	CLP5 $\mathbf{x}_3^{i,t,1}$	AEMI $\mathbf{x}_1^{i,t,1}$	AEMN $\mathbf{x}_2^{i,t,2}$	CLP5 $\mathbf{x}_3^{i,t,2}$
SANDY	1	[12.7; -78.7]	[14.6; -77.8]	N/A	[13.4; -80.7]	[18.4; -76.4]	N/A	[14.2; -82.3]
	2	[12.9; -78.1]	[15.7; -77.7]	N/A	[14.1; -79.2]	[19.8; -76.8]	N/A	[15.6; -79.9]
	3	[14.0; -77.6]	[17.8; -76.9]	N/A	[16.0; -77.5]	[22.7; -76.5]	N/A	[18.0; -77.9]
IRMA	1	[16.4; -32.5]	[18.5; -35.3]	[18.8; -35.3]	N/A	[19.1; -39.7]	[19.2; -40.0]	N/A
	2	[17.1; -34.2]	[18.4; -38.4]	N/A	N/A	[18.2; -42.9]	N/A	N/A
	3	[17.9; -36.1]	[18.4; -40.2]	[18.7; -40.5]	N/A	[17.5; -44.4]	[17.6; -45.0]	N/A

Table 1: Example of NHC best track data along with the forecasts generated by an ensemble of dynamical models (AEMI, AEMN, and CLP5) for hurricanes Sandy and Irma at 24 and 48-hour lead times. Let $\mathbf{x}_j^{i,t,\tau}$ be the τ-th lead-time forecast generated at time t for hurricane h_i by ensemble member j and $\mathbf{y}^{i,t}$ be its best track location. N/A denotes missing values.

eral limitations to these machine learning approaches. First, they are mostly based on auto-regressive or recurrent neural network models, using only historical observation data. Due to the inherent error propagation problem (Cheng et al. 2006) in such models, they are mostly suitable for short-range predictions. Second, due to the chaotic nature of the weather system and the varying conditions in the atmosphere and ocean temperature, the historical data alone may not be enough to train a reliable long-range forecasting model. By grounding the historical observations with multi-model ensemble forecasts, it may lead to more reliable predictions. Third, the previous methods are mostly designed for batch learning. Thus, they require the model to be re-trained from scratch whenever new observations become available. An online learning method is more appealing as it allows the model to be efficiently updated to fit the new observations while adapting to the concept drifts present in the data.

Designing an online learning algorithm for hurricane trajectory forecasting is a challenge for several reasons. First, the models trained for predicting the hurricane's location at different lead times must take into account the inherent autocorrelation along the trajectory. Furthermore, they are susceptible to the partially observed data problem inherent in multi-lead time forecasting tasks (Xu, Tan, and Luo 2014). For example, if the model is updated every six hours with new observation data and the forecast horizon (i.e., maximum lead time) is 48 hours, it is insufficient to revise only the latest model. Instead, we must also revise some of the earlier models, starting from 48 hours ago up to 6 hours ago, to avoid propagating the errors from the earlier models into future prediction. Another challenge is that the ensemble members available may vary from one hurricane to another (see Table 1). Due to the missing forecasts by some model members, the online algorithm must adaptively learn the weights in spite of the varying feature lengths. To address these challenges, we propose a novel framework called OMuLeT (**O**nline **Mu**lti-**Le**ad **T**ime Forecasting), which employs an online learning with restart strategy to incrementally update the weights of the ensemble members. OMuLeT also employs a novel weight renormalization scheme to handle the varying number of ensemble member forecasts. Experimental results using the Atlantic and Eastern Pacific hurricane data showed that OMuLeT can improve the 48-hour lead time official forecast of NHC by more than 10%.

Problem Formulation

We consider the problem of predicting hurricane trajectory using forecasts generated from a multi-model ensemble with multiple lead times. Consider a set of N hurricanes, $h_1 \leq h_2 \leq \cdots \leq h_N$, ordered by their start times. For the i-th hurricane, let $\mathbf{y}^{i,t} \in \mathbb{R}^2$ denote its location (latitude and longitude) at time t, where $t \in \{t_{i,1}, \cdots, t_{i,\Gamma_i}\}$ and Γ_i denotes the observed trajectory length for hurricane h_i. Furthermore, at each time t, our goal is to forecast the hurricane's location at a future time step $t + \tau$, where $\tau \in \{1, \cdots, T\}$ is the lead time and T is the forecast horizon.

Let m_i be the number of ensemble member forecasts available for hurricane h_i. The set of ensemble member forecasts available to predict the location of h_i at time $t + \tau$ is represented by the matrix $\mathbf{X}^{i,t,\tau} \in \mathbb{R}^{2\times m_i}$, while its ground truth location is given by $\mathbf{y}^{i,t+\tau} \in \mathbb{R}^2$. The hurricane trajectory data is given by a set of 2-tuples, $\{(\mathbf{X}^{i,t,\tau}, \mathbf{y}^{i,t+\tau})\}$, where the superscript i denotes the hurricane, t is the forecast generation time, and τ is the forecast lead time.

Varying Feature Length: One of the key characteristics of the multi-model ensemble hurricane trajectory data is that its feature length, i.e., number of ensemble member forecasts (m_i) associated with each hurricane, may vary from one hurricane to another . Specifically, although there are numerous ensemble member forecasts generated over the years, each hurricane has forecasts obtained from an average of only 19 ensemble members in our dataset. The unavailable ensemble members would create non-random missing patterns in the data. Imputing the missing values is not a viable solution due to the high missing rate. Instead, we propose an approach that can automatically handle the varying feature length by renormalizing the weights of the ensemble members.

Temporal Inconsistencies: Outputs from the dynamical models have varying degrees of temporal inconsistencies. First, the dynamical models can have different forecast time intervals. Some models generate their forecasts every 6 hours, while others every 12 hours. To address this problem, we perform interpolation to impute the missing values of the 12-hourly forecast intervals to obtain 6-hour forecasts for all ensemble members. Second, the forecast duration often varies among the ensemble members. For example, some members generate their forecasts for only a few days, while

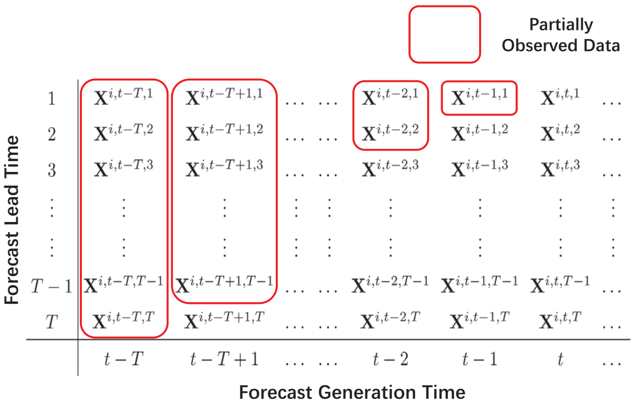


Figure 1: An illustration of the partially observed labeled data problem. The red rectangles denote the set of model forecasts for which ground truth are available at time t.

others may extend longer than a week. In addition, their forecast horizon are also different. For example, some models produce forecasts with a maximum lead time of 24 hours, while others may generate forecasts for a lead time up to 120 hours. A novel weight renormalization approach is proposed in this paper to address such temporal discrepancies.

Partially Observed Labeled Data: Another challenge is that ground truth values for the multi-lead time forecasts are only partially observed. This problem is illustrated in Figure 1. Let $\mathbf{X}^{i,t-1,1}$ be the set of ensemble member forecasts generated for hurricane h_i at time $t-1$ for the lead time $\tau = 1$. If the current time is t, then the ground truth value $\mathbf{y}^{i,t}$ will be available to verify the accuracy of the forecasts in $\mathbf{X}^{i,t-1,1}$. However, for the longer-range forecasts, $\mathbf{X}^{i,t-1,2}$, $\mathbf{X}^{i,t-1,3}$, $\cdots$, $\mathbf{X}^{i,t-1,T}$, the ground truth values have not been observed. In fact, the ground truth values are only available for any previous forecast $\mathbf{X}^{i,t-k,\tau}$ for which $\tau - k \leq 0$. This corresponds to the red rectangles shown in Figure 1. As time progresses to $t+1$, the true value for $\mathbf{y}^{i,t+1}$ will be known. Conventional online learning algorithms use the new observation $\mathbf{y}^{i,t+1}$ to update their latest models only. This is insufficient for multi-lead time forecasting as the new observation data may trigger a cascading effect since some of the earlier models from which the current models are obtained are also outdated. The models for various lead times generated at time t must be rolled-back all the way to time $t-T+1$ and updated again with the new observation data to alleviate the error propagation problem. This strategy is known as online learning with restart (Xu, Tan, and Luo 2014).

Methodology

We consider an online model of the form

$$f(\mathbf{X}^{i,t,\tau}) = \mathbf{X}^{i,t,\tau}\mathbf{w}^{i,t,\tau} \tag{1}$$

for predicting the location of hurricane h_i at time $t+\tau$, where $\mathbf{w}^{i,t,\tau} \in \mathbb{R}^m$ is the weight vector associated with m ensemble member forecasts. Conventional online learning algorithms (Crammer et al. 2006) typically assume that the feature matrix $\mathbf{X}^{i,t,\tau}$ is either complete or has few missing values, which can be imputed during preprocessing. However, due to the varying feature length problem described in the previous section, some ensemble member forecasts are not available for the entire duration of a hurricane. Below, we describe our proposed approach to address this problem.

Weight Renormalization

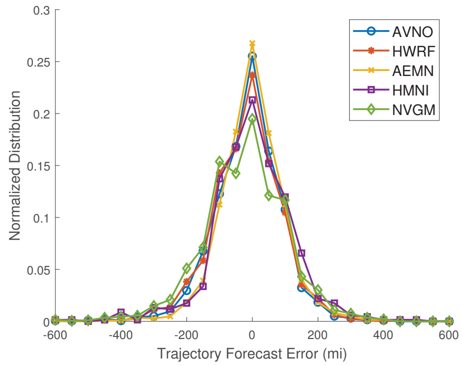
This section presents the weight renormalization approach employed by our online learning framework to overcome the varying feature length problem. Let $\boldsymbol{\mu} = \{\mu_1, \mu_2, \cdots, \mu_m\}$ be the set of all ensemble members and $\mathbf{M}_i \subseteq \boldsymbol{\mu}$ be the subset of members whose forecasts are available for hurricane h_i. Since $|\mathbf{M}_i| \ll m$, imputing the missing ensemble member forecasts is not an effective approach given the large amount of missing values present in the data. Instead, we present an online learning approach that uses only the ensemble member forecasts available for the given hurricane ($\mathbf{M}_i$) and update their weights accordingly when new observation data becomes available at each time step. Specifically, we assume the forecasts from each ensemble member follow a Gaussian distribution centered at the true location. To illustrate this, Figure 2 shows a normalized histogram of the trajectory forecast errors for 5 dynamical models when applied to more than 200 hurricanes in our dataset. Observe that the forecast error distribution indeed resembles that of a Gaussian distribution. We also assume that the weights of the ensemble members form an m-dimensional simplex, i.e.:

$$\Delta_m = \{w_1^i, w_2^i, \cdots, w_m^i \mid \forall i : \sum_j w_j^i = 1, w_j^i \geq 0\}.$$

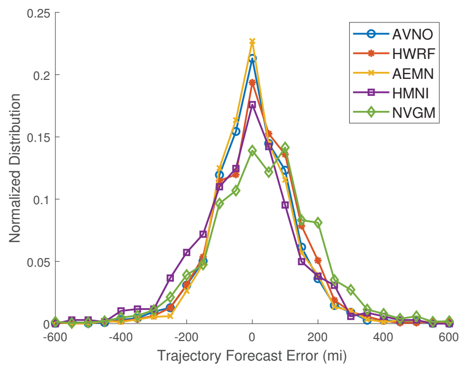
Given a hurricane h_i, our framework performs the following steps to incrementally update the weights:

1. We extract the subvector $\mathbf{w}_0^i \in \mathbb{R}^{m_i}$ from the full vector $\mathbf{w} \in \mathbb{R}^m$, whose elements contain only the weights of the m_i ensemble members in $\mathbf{M}_i$.
2. We normalize $\mathbf{w}_0^i$ to have unit sum as follows:
$$\mathbf{w}_0^i \leftarrow \mathbf{w}_0^i / c, \text{ where } c = \|\mathbf{w}_0^i\|_1 = \sum_j \mathbf{w}_{0,j}^i$$
3. At each time step $t = \{t_{i,1}, t_{i,2}, \cdots, t_{i,\Gamma_i}\}$:
   (a) We use the normalized weights to predict the location of the hurricane at lead time $t+\tau$, i.e., $f(\mathbf{X}^{i,t,\tau}) = \mathbf{X}^{i,t,\tau}\mathbf{w}^{i,t,\tau}$, where $\mathbf{w}^{i,t,\tau}$ is computed from $\mathbf{w}_0^i$ according to Eqn. (5).
   (b) After observing the ground truth location $\mathbf{y}^{i,t}$, we update $\mathbf{w}^{i,t,\tau}$ using the method described in Section .
4. After the last update at time $t = \Gamma_i$, the updated weights are renormalized to their original sum:
$$\mathbf{w}_0^i \leftarrow c\mathbf{w}_0^i \tag{2}$$
before being replaced into the full vector $\mathbf{w}$. This ensures that $\mathbf{w}$ remains a simplex after the weight update.

The preceding approach enables our framework to update only the weights of the ensemble members whose forecasts are available for hurricane h_i. The weights need to be



(a) Forecast errors along the latitude direction.



(b) Forecast errors along the longitude direction.

Figure 2: Normalized distribution of trajectory forecast errors (in 50 miles bin) for 5 different dynamical models along their latitude and longitude directions.

renormalized when replacing them back into the full vector $\mathbf{w}$. Below, we provide a formal justification for the weight renormalization approach. For brevity, we assume the hurricane location is a scalar variable, even though the theorem below can be extended to 2-dimensional location vectors.

Theorem 1. *Let y denotes the true hurricane location and $\{X_1, X_2, \cdots, X_M\}$ be i.i.d. random variables, representing the forecasts of M ensemble members. Assume each X_j is a random perturbation around y, i.e.:*

$$X_j = y + \epsilon(0, \sigma_j^2),$$

where $\epsilon(0, \sigma_j^2)$ is a Gaussian distribution with mean 0 and variance σ_j^2. Then, the best unbiased linear estimator (BLUE) for y is $z = \sum_j w_j X_j$, where $w_j = \frac{1}{\sigma_j^2} \sum_{j=1}^M \frac{1}{\sigma_j^2}$.

Proof. We consider a linear estimator of the form $z = \sum_j w_j X_j$. Since $\{X_1, X_2, \ldots, X_M\}$ are i.i.d. variables and z is unbiased. Therefore:

$$E(z) = \sum_{j=1}^M w_j E(X_j) = \sum_{j=1}^M w_j y = y \Rightarrow \sum_{j=1}^M w_j = 1 \quad (3)$$

The variance of the linear estimator is $Var(z) = \sum_{j=1}^M w_j^2 \sigma_j^2$. To find the $\mathbf{w}$ that minimizes the variance, subject to the constraint in Eq. (3), we consider the Lagrangian function $L = \sum_{j=1}^M w_j^2 \sigma_j^2 - \lambda(\sum_{j=1}^M w_j - 1)$. Taking its partial derivative w.r.t w_k and setting it to 0 yields

$$\frac{\partial L}{\partial w_k} = 2 w_k \sigma_j^2 - \lambda = 0 \Rightarrow w_k = \frac{\lambda}{2\sigma_k^2}$$

Following the constraint $\sum_{j=1}^M w_j = 1$, we can solve for λ and obtain:

$$w_j = \frac{1}{\sigma_j^2} \Big/ \sum_{k=1}^M \frac{1}{\sigma_k^2} \quad (4)$$

which completes the proof. □

The preceding theorem considers an estimator z computed from M i.i.d. variables. Let $\tilde{z}$ be another estimator computed using K of the i.i.d. variables in $\{X_j\}$. Without loss of generality, assume the subset corresponds to $X_1, X_2, \cdots, X_K$.

Corollary 1. *Let y be the true location of the hurricane and $\tilde{z} = \sum_{j=1}^K \tilde{w}_j X_j$ be a linear estimator of y, where each $X_j = y + \epsilon(0, \sigma_j)^2$. Then the best linear unbiased estimator (BLUE) for y using the K i.i.d. variables is $\tilde{z} = \sum_{j=1}^K \tilde{w}_j X_j$, where $\tilde{w}_j = c w_j$ and $c = \frac{1}{\sum_{j=1}^K w_j}$.*

The preceding corollary provides the normalization factor needed to re-scale the weights of the ensemble members.

Geographic Distance Loss Function

Instead of using a squared ℓ_2 (Euclidean) loss function, our framework considers the squared geographic distance to compute the error in location estimation. Let $\mathbf{z}^{i,t,\tau} = [z_1^{i,t,\tau}, z_2^{i,t,\tau}]$ be the predicted latitude and longitude position of hurricane h_i at time t for the lead time τ and $\mathbf{y}^{i,t+\tau} = [y_1^{i,t+\tau}, y_2^{i,t+\tau}]$ be the corresponding true location. The squared geographic distance between the predicted and true locations, $d[\mathbf{z}^{i,t,\tau}, \mathbf{y}^{i,t+\tau}]^2$, can be estimated as follows:

$$R_e^2 \left[(z_1^{i,t,\tau} - y_1^{i,t+\tau})^2 + (z_2^{i,t,\tau} - y_2^{i,t+\tau})^2 \cos^2 y_1^{i,t+\tau} \right],$$

where R_e is the radius of the earth. As R_e is a constant that can be absorbed into the regularizer term of an objective function, we can set $R_e = 1$ to simplify the notation. Furthermore, by transforming the coordinates of the location to

$$\tilde{\mathbf{y}}^{i,t+\tau} = [y_1^{i,t+\tau}, y_2^{i,t+\tau} \cos y_1^{i,t+\tau}]$$
$$\tilde{\mathbf{z}}^{i,t,\tau} = [z_1^{i,t,\tau}, z_2^{i,t,\tau} \cos y_1^{i,t+\tau}]$$

the geographic distance can be further simplified as follows:

$$d(\mathbf{z}^{i,t,\tau}, \mathbf{y}^{i,t+\tau})^2 = \|\tilde{\mathbf{z}}^{i,t,\tau} - \tilde{\mathbf{y}}^{i,t+\tau}\|^2$$

which is an ℓ_2 loss on the transformed coordinates.

OMuLeT: Proposed Framework

Our proposed framework, named OMuLeT, learns the weights for the ensemble members in an online fashion. Let m be the total number of ensemble members and m_i be the number of ensemble members whose forecasts are

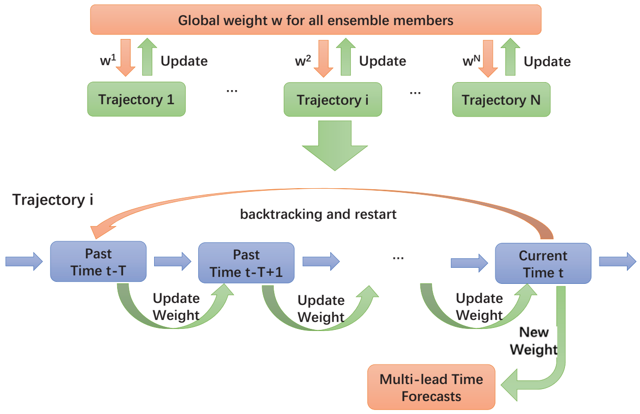


Figure 3: Proposed OMuLeT framework.

available for hurricane h_i. To predict the location of hurricane h_i at time $t+\tau$, we use the following linear estimator, $\mathbf{z}^{i,t,\tau} = \mathbf{X}^{i,t,\tau}\mathbf{w}^{i,t,\tau}$, where $\mathbf{X}^{i,t,\tau} \in \mathbb{R}^{2\times m_i}$, $\mathbf{w}^{i,t,\tau} \in \mathbb{R}^{m_i}$, and $\mathbf{1}_{m_i}^T\mathbf{w}^{i,t,\tau} = 1$. OMuLeT assumes the weight vector $\mathbf{w}^{i,t,\tau}$ can be decomposed into the following three factors:

$$\mathbf{w}^{i,t,\tau} = \mathbf{w}_0^i + \mathbf{u}^{i,t} + \mathbf{v}^{i,t,\tau} \tag{5}$$

$$\text{where} \quad \mathbf{1}_{m_i}^T\mathbf{w}_0^i = 1, \mathbf{1}_{m_i}^T\mathbf{u}^{i,t} = 0, \mathbf{1}_{m_i}^T\mathbf{v}^{i,t,\tau} = 0$$

The first term, $\mathbf{w}_0^i$, is a *global weight* that retains the weight information from past hurricanes. The second term, $\mathbf{u}^{i,t}$, is a *hurricane-specific* factor that modifies the weight to fit the current hurricane. The third term, $\mathbf{v}^{i,t,\tau}$, is a *lead time adjustment* factor to improve prediction at lead time τ.

The overall framework is shown in Figure 3. Given a hurricane h_i, we first extract the subvector $\mathbf{w}_0^i$ from $\mathbf{w}$ and normalize it to have unit sum. Both $\mathbf{u}^{i,0}$ and $\mathbf{v}^{i,0,\tau}$ are initialized to a vector of all zeros, $\mathbf{0}_{m_i}$. As new observation data becomes available, $\mathbf{u}^{i,t}$ and $\mathbf{v}^{i,t,\tau}$ are updated using the approach to be described below. Note that $\mathbf{w}_0^i$ remains unchanged throughout the time steps $t = \{t_{i,1}, t_{i,2}, \cdots, t_{i,\Gamma_i-1}\}$ and is only updated as follows after the entire trajectory has been processed:

$$\mathbf{w}_0^i \leftarrow \mathbf{w}_0^i + \rho\mathbf{u}^{i,\Gamma_i}. \tag{6}$$

The hyperparameter ρ controls the tradeoff between using the weights from current and past hurricanes. Finally, the updated weights in $\mathbf{w}_0^i$ are renormalized using Eqn. (2) before they are substituted back into the full vector $\mathbf{w}$.

Online Learning with Restart OMuLeT employs an online learning with restart strategy to address the partially labeled data problem, as illustrated in Figure 3. When the ground truth value $\mathbf{y}^{i,t}$ is observed at time t, it can be used to verify the accuracy of earlier forecasts, $\{\mathbf{X}^{i,t-1,1}, \mathbf{X}^{i,t-2,2}, \cdots, \mathbf{X}^{i,t-T,T}\}$. Let $\mathbf{W}^{i,t} = [\mathbf{w}^{i,t,1}, \mathbf{w}^{i,t,2}, \cdots, \mathbf{w}^{i,t,T}]$ be the weight matrix at time t for each of the T lead times. In conventional online learning, $\mathbf{W}^{i,t}$ is estimated from $\mathbf{W}^{i,t-1}$ using the observed $\mathbf{y}^{i,t}$. In turn, $\mathbf{W}^{i,t-1}$ was estimated from $\mathbf{W}^{i,t-2}$ using previous observation $\mathbf{y}^{i,t-1}$, and so on. This strategy may not be sufficient for multi-lead time prediction as $\mathbf{y}^{i,t}$ is used to modify $\mathbf{W}^{i,t-1}$ only, even though the data can help improve earlier estimates of $\mathbf{W}^{i,t-2}, \mathbf{W}^{i,t-3}, \cdots, \mathbf{W}^{i,t-T+1}$. For example, $\mathbf{y}^{i,t}$ can also be used to verify $\mathbf{X}^{i,t-2,2}$, and thus, improve the estimate of $\mathbf{w}^{i,t-2,2}$. More importantly, $\mathbf{y}^{i,t}$ can verify the accuracy of earlier forecasts $\mathbf{X}^{i,t-T,T}$, and thus, improve the estimate of the weights for the long-range model, $\mathbf{w}^{i,t-T,T}$. By utilizing $\mathbf{y}^{i,t}$ to update $\mathbf{w}^{i,t-T,T}$, this can help alleviate the error propagation problem encountered in multi-step ahead time series forecasting. The online learning with restart strategy employed by OMuLeT allows the algorithm backtracks to time $t-T+1$ and restarts its update for $\mathbf{w}^{i,t-T+1,T}$ to account for the new ground truth data available for $\mathbf{X}^{i,t-T,T}$. In turn, the updated weight matrix $\mathbf{W}^{i,t-T}$ is then used to update $\mathbf{W}^{i,t-T+1}$, taking into account the new ground truth data for $\mathbf{X}^{i,t-T+1,T-1}$. This process is repeated until the new weight matrix $\mathbf{W}^{i,t}$ is obtained.

Objective Function The weights of our online model are updated by solving the following optimization problem:

$$\begin{aligned}
&\min_{\mathbf{u}^{i,t},\{\mathbf{v}^{i,t,\tau}\}} \frac{1}{2}\sum_{\tau=1}^{T}\delta^{i,t,\tau}\gamma^{\tau}d\big[\mathbf{z}^{i,t,\tau},\mathbf{y}^{i,t+\tau}\big]^2 \\
&+ \frac{\omega}{2}\sum_{\tau=1}^{T-1}\big\|\mathbf{w}^{i,t,\tau+1}-\mathbf{w}^{i,t,\tau}\big\|^2 + \frac{\mu}{2}\big\|\mathbf{u}^{i,t}-\mathbf{u}^{i,t-1}\big\|^2 \\
&+ \frac{\nu}{2}\sum_{\tau=1}^{T}\big\|\mathbf{v}^{i,t,\tau}-\mathbf{v}^{i,t-1,\tau}\big\|^2 + \frac{\eta}{2}\sum_{\tau=1}^{T}\big\|\mathbf{v}^{i,t,\tau}\big\|^2 \\
&\text{s.t. } \forall\, t,\tau:\ \mathbf{1}_{m_i}^T\mathbf{u}^{i,t} = 0, \mathbf{1}_{m_i}^T\mathbf{v}^{i,t,\tau} = 0,
\end{aligned} \tag{7}$$

where $d[\cdot]$ is the geographic distance function described in Section while $\delta^{i,t,\tau}$ is an indicator function whose value is equal to 1 if $\mathbf{y}^{i,t+\tau}$ is observed at time t; otherwise its value is 0. The first term in the objective function represents the forecast error. The hyperparameter γ determines the relative importance of making accurate forecasts at different lead times τ. The second term ensures smoothness in the model parameters for different lead times whereas the third and fourth terms are designed to ensure the hurricane-specific factor $\mathbf{u}^{i,t}$ and lead time adjustment factor $\mathbf{v}^{i,t}$ do not change rapidly from their previous values at time $t-1$. The last term in the objective function imposes a sparsity constraint on the lead time adjustment factor. To implement the online learning with restart strategy, Eqn (7) must be solved $T+1$ times, starting from $t = t_c - T$ to $t = t_c$, where t_c is the current time.

The Lagrange formulation for the problem is

$$\begin{aligned}
\mathcal{L} \;=\;& \frac{1}{2}\sum_{\tau=1}^{T}\delta^{i,t,\tau}\gamma^{\tau}\left\|\tilde{\mathbf{X}}^{i,t,\tau}\mathbf{w}^{i,t,\tau}-\tilde{\mathbf{y}}^{i,t+\tau}\right\|_2^2 \\
+\;& \frac{1}{2}Tr\left[\mathbf{V}^{i,t^T}(\omega\mathbf{L}+\eta\mathbf{I})\mathbf{V}^{i,t}\right] \\
+\;& \frac{\mu}{2}\left\|\mathbf{u}^{i,t}-\mathbf{u}^{i,t-1}\right\|^2 + \frac{\nu}{2}\left\|\mathbf{V}^{i,t}-\mathbf{V}^{i,t-1}\right\|_F^2 \\
-\;& \lambda\mathbf{1}_{m_i}^T\mathbf{u}^{i,t} - \sum_{\tau=1}^{T}\theta_\tau\mathbf{1}_{m_i}^T\mathbf{v}^{i,t,\tau}
\end{aligned} \tag{8}$$

where $\mathbf{V}^{i,t} = [\mathbf{v}^{i,t,1}, \mathbf{v}^{i,t,2}, \ldots, \mathbf{v}^{i,t,T}]$, λ, $\theta_1, \cdots, \theta_T$ are the Lagrange multipliers. The objective function can be

Input: $\rho, \gamma, \omega, \mu, \nu, \eta$
Output: Model parameters $\mathbf{w}$ and forecasts $\mathbf{z}$
Initialize: $\mathbf{w} = \mathbf{1}_m/m$;
for $i = 1, 2, \ldots, N$ **do**
  Extract $\mathbf{w}_0^i$ from $\mathbf{w}$
  Normalize: $\mathbf{w}_0^i \leftarrow \mathbf{w}_0^i/c$
  $\forall t, \tau : \mathbf{u}^{i,t} = \mathbf{0}_{m_i}, \mathbf{v}^{i,t,\tau} = \mathbf{0}_{m_i}$
  for $t = t_{i,1}, t_{i,2}, \cdots, t_{i,\Gamma_i}$ **do**
    Observe $\mathbf{y}^{i,t}$
    /* Backtracking and restart step */
    for $t' = t - T, t - T + 1, \ldots, t$ **do**
      Solve $\mathbf{A}^{i,t'} \varphi^{i,t'} = \mathbf{b}^{i,t'}$
      Solve $\mathbf{u}^{i,t'}$ and $\{\mathbf{v}^{i,t',\tau}\}$ using Eqn. (9)
    end
    /* Prediction step */
    for $\tau = 1, 2, \cdots, T$ **do**
      Compute $\mathbf{w}^{i,t,\tau}$ using Eqn. (5)
      $\mathbf{z}^{i,t+\tau} = f(\mathbf{X}^{i,t,\tau}) = \mathbf{X}^{i,t,\tau}\mathbf{w}^{i,t,\tau}$
    end
  end
  Update $\mathbf{w}_0^i$ based on Eq. (6)
  Renormalize: $\mathbf{w}_0^i = c\mathbf{w}_0^i$
  Substitute $\mathbf{w}_0^i$ back into the full vector $\mathbf{w}$
end

Algorithm 1: Proposed OMuLeT Framework

solved by taking the derivative of $\mathcal{L}$ with respect to the model parameters and setting them to zero. A closed-form solution for the Lagrange formulation is found by solving a system of linear equations: : $\mathbf{A}^{i,t}\varphi^{i,t} = \mathbf{b}^{i,t}$, where $\varphi^{i,t} = [\mathbf{\Delta} u^{i,t}, \mathbf{\Delta} v^{i,t,1}, \ldots, \mathbf{\Delta} v^{i,t,T}, \lambda, \theta_1, \ldots, \theta_T]^T$.

After obtaining $\varphi^{i,t}$, the weights $\mathbf{u}^{i,t}$ and $\mathbf{v}^{i,t,\tau}$ are updated as follows:

$$\begin{aligned} \mathbf{u}^{i,t} &= \mathbf{u}^{i,t-1} + \Delta\mathbf{u}^{i,t} \\ \mathbf{v}^{i,t,\tau} &= \mathbf{v}^{i,t-1,\tau} + \Delta\mathbf{v}^{i,t,\tau} \end{aligned} \tag{9}$$

The pseudocode of our framework is shown in Algorithm 1.

Experimental Evaluation

The hurricane best track (ground truth) data and NHC official forecasts are available from the NHC website[1], while the ensemble member forecasts were downloaded from the Hurricane Forecast Model Output website at University of Wisconsin-Milwaukee[2]. According to NHC, 46 models were used in the preparation of their official forecasts. However, only 28 of them were available at the UWM website, which will be used as ensemble members in our experiments.

Although the best track data dates back to 1851 for Atlantic and 1949 for Pacific oceans, both the NHC official forecasts as well as the ensemble member forecasts have a much shorter history. After fusing the best track with forecast data, our final dataset contains 212 tropical cyclones spanning the years 2012 to 2018. We performed linear interpolation to impute the missing values for ensemble members with 12-hourly forecasts to ensure they also generate 6-hourly forecasts. The maximum forecast lead time T is set to 48 hours. The hurricane data from 2012 to 2014 (84 tropical cyclones) are used for training and validation while those from 2015 to 2018 (128 tropical cyclones) are used for testing. Each trajectory has an average length of 24 time steps at 6 hourly intervals. In total, there are 2086 observations in the training and validation periods and 2946 observations in the test period. The hurricanes are divided into two groups, those originating from the Atlantic ocean and those from the Eastern Pacific ocean.

We compared OMuLeT against the following methods:

1. **LSTM**: Following the approach used in (Alemany et al. 2019)[3], we train an LSTM model (Hochreiter and Schmidhuber 1997) on the best track data from 1851 (for Atlantic ocean) and 1949 (for Pacific ocean) to 2014. We report the test results for the period between 2015 until 2018.
2. **Ensemble mean (EM)**: This corresponds to taking the average of all the ensemble member forecasts.
3. **NHC**: This is the gold standard, corresponding to the official forecasts generated by NHC.
4. **Passive-Aggresive (PA)**(Crammer et al. 2006): A well-known online learning algorithm that updates the weights of its linear model based on the following equation:
$$\mathbf{w}^{t+1} = \mathbf{w}^t + sign(\mathbf{y}^t - \mathbf{z}^t)\tau^t\mathbf{x}^t \tag{10}$$
5. **ORION**(Xu, Tan, and Luo 2014): A recently developed online learning algorithm for multi-lead time prediction.

For a fair comparison, the baseline methods such as PA and ORION also use the weight renormalization strategy to deal with the varying feature length problem. The source code and data used in our experiments are available on our website[4].

Experimental Results

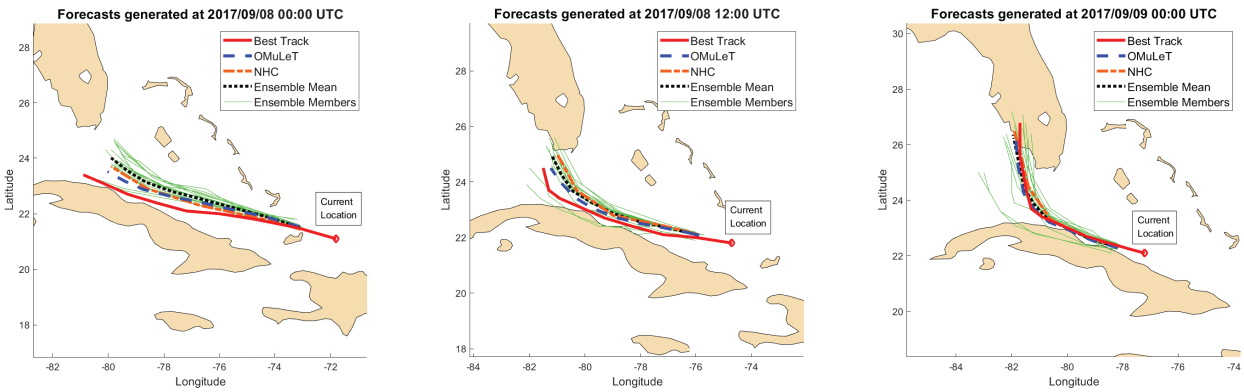
Table 2 summarizes the forecast errors of the different methods, in terms of their average geographic distance (in miles) between the true and predicted locations. There are several interesting conclusions can be drawn from the results shown in the table. First, the LSTM results were significantly worse than other methods despite using a longer history of hurricane trajectory data to train the model. This is not surprising as the historical tracks do not capture the varying atmospheric conditions and ocean temperatures that affect the path of the hurricanes. Furthermore, the parameters of the LSTM model are fixed after training unlike the online learning models such as PA, ORION, and OMuLeT that continuously update its parameters with new observation data.

Second, the performance of ensemble mean is comparable to the NHC official forecasts, which validates the skills of

[1]https://www.nhc.noaa.gov
[2]http://derecho.math.uwm.edu/models
[3]We directly predict the trajectory location instead of the grid from their paper.
[4]https://github.com/cqwangding/OMuLeT.

Location	Atlantic and Pacific ocean				Atlantic ocean only				Pacific ocean only			
Lead Time	12	24	36	48	12	24	36	48	12	24	36	48
$LSTM$	142.89	211.42	305.91	537.29	175.62	257.16	361.63	594.01	113.21	161.93	259.31	385.17
EM	26.09	40.20	53.77	68.49	27.74	42.74	57.85	75.39	24.67	38.06	50.33	62.74
PA	26.09	40.07	53.60	68.01	27.68	42.57	57.77	74.87	24.66	38.03	50.29	62.67
$ORION$	25.50	39.04	51.97	66.27	26.77	41.08	55.55	72.26	24.51	37.46	49.27	61.48
NHC	26.58	40.97	54.41	68.47	28.83	44.03	58.66	75.94	24.65	38.39	50.83	62.25
$OMuLeT$	**24.94**	**37.08**	**48.78**	**59.08**	**26.53**	**38.77**	**51.85**	**63.61**	**23.73**	**36.03**	**45.85**	**55.80**

Table 2: Comparison of mean geographic distance error (in miles) for various hurricane trajectory forecasting methods.



(a) (b) (c)

Figure 4: Comparison of 48-hour forecasts for Hurricane Irma from 2017/09/08 to 2017/09/09 by different methods.

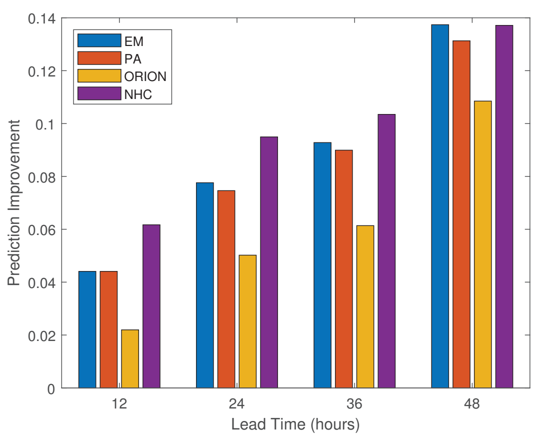


Figure 5: Percentage of forecast improvement of OMuLeT compared to the baseline methods.

the ensemble members. This is not surprising as the skills of the dynamical models have continuously improved over the years and the ensemble members considered in this study are a subset of those used in the NHC official forecasts.

Third, existing online algorithms such as PA and ORION do not significantly improve the prediction error of ensemble mean even though their weights are updated continuously. PA updates only its latest model whenever new observation data becomes available unlike ORION and OMuLeT, which employ the online update with restart strategy to revise some of their earlier models. Furthermore, ORION was designed for multi-lead time forecasting at a single location. Extending the approach to modeling different hurricanes is not effective as it fails to retain the weight information from past hurricanes, unlike the weight decomposition approach used in OMuLeT (see Eqs. (5)) and (6)).

Fourth, OMuLeT consistently outperforms all the baseline methods irrespective of the forecast lead time. Figure 5 illustrates the forecast improvement of OMuLeT over NHC and other baselines for different lead times. Observe that the forecast improvement of OMuLeT over the baseline methods continues to grow with increasing lead times. More importantly, it outperforms the official NHC forecasts by more than 10% for the 48-hour lead time forecast.

Figure 4 shows an example of the trajectory forecasts for Hurricane Irma from September 8 to September 9, 2017. Observe that OMuLeT's 48-hour forecasts are closest to the best track compared to the baseline methods, especially in Figure 4(a) and 4(b). Despite the large variability among the ensemble member forecasts, OMuLeT was able to assign the appropriate set of weights to the ensemble members, which led to more accurate forecasts.

Figure 6 shows the dynamic weights of the ensemble members learned using OMuLeT. The plot suggests that the Global Forecast System (AVNO) generally has higher weights than others. Other models, such as the Hurricane Weather Research and Forecast system (HWRF) and U.K. Met Office Global Model (EGRI), have also become increas-

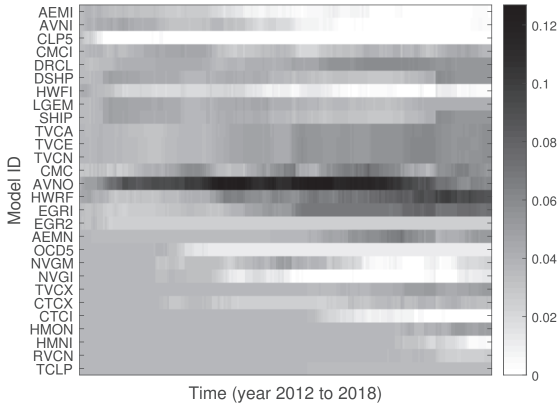


Figure 6: Global weight changes over time

ingly skillful in recent years. This result shows the advantage of using an online learning strategy to continuously adapt the weights of the ensemble members based on their relative performance for different hurricanes.

Conclusions

This paper presents a novel online learning framework called OMuLeT for multi-lead time hurricane trajectory forecasting. Unlike existing methods, OMuLeT uses multi-model ensemble member outputs to train its model. It also employs novel weight renormalization and update strategies to address the various modeling challenges. Experimental results showed that our framework significantly outperforms various baseline methods, including NHC official forecasts, especially for long-range forecasting. OMuLeT thus provides a promising approach for early warning systems.

For future work, we plan to investigate alternative approaches such as Bayesian model average (Raftery et al. 2005) for merging the ensemble member forecasts. However, such an approach will need to be modified to handle the multi-lead time trajectory forecasting problem with vast amount of missing values.

Acknowledgments

This research is partially supported by the National Science Foundation under grant #IIS-1615612. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Alemany, S.; Beltran, J.; Perez, A.; and Ganzfried, S. 2019. Predicting hurricane trajectories using a recurrent neural network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 468–475.

Blake, E. S.; Landsea, C. W.; and Gibney, E. J. 2011. *The deadliest, costliest, and most intense United States tropical cyclones from 1851 to 2010 (and other frequently requested hurricane facts)*. National Weather Service, National Hurricane Center.

Cheng, H.; Tan, P.-N.; Gao, J.; and Scripps, J. 2006. Multistep-Ahead Time Series Prediction. In *Proceedings of the 10th Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 765–774.

Crammer, K.; Dekel, O.; Keshet, J.; Shalev-Shwartz, S.; and Singer, Y. 2006. Online passive-aggressive algorithms. *Journal of Machine Learning Research* 7:551–585.

Davis, C.; Wang, W.; Chen, S. S.; Chen, Y.; Corbosiero, K.; DeMaria, M.; Dudhia, J.; Holland, G.; Klemp, J.; Michalakes, J.; et al. 2008. Prediction of landfalling hurricanes with the advanced hurricane WRF model. *Monthly weather review* 136(6):1990–2005.

DeMaria, M.; Mainelli, M.; Shay, L. K.; Knaff, J. A.; and Kaplan, J. 2005. Further improvements to the statistical hurricane intensity prediction scheme (SHIPS). *Weather and Forecasting* 20(4):531–543.

Gneiting, T., and Raftery, A. E. 2005. Weather forecasting with ensemble methods. *Science* 310(5746):248–249.

Hochreiter, S., and Schmidhuber, J. 1997. Long short-term memory. *Neural computation* 9(8):1735–1780.

Kim, H.-M., and Webster, P. J. 2010. Extended-range seasonal hurricane forecasts for the north atlantic with a hybrid dynamical-statistical model. *Geophysical Research Letters* 37(21).

Kordmahalleh, M. M.; Sefidmazgi, M. G.; and Homaifar, A. 2016. A Sparse Recurrent Neural Network for Trajectory Prediction of Atlantic Hurricanes. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 957–964.

Lee, R. S., and Liu, J. N. 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):680–689.

Leutbecher, M., and Palmer, T. N. 2008. Ensemble forecasting. *Journal of Computational Physics* 227(7):3515–3539.

Raftery, A. E.; Gneiting, T.; Balabdaoui, F.; and Polakowski, M. 2005. Using bayesian model averaging to calibrate forecast ensembles. *Monthly weather review* 133(5):1155–1174.

Reich, B. J., and Fuentes, M. 2007. A multivariate semiparametric Bayesian spatial modeling framework for hurricane surface wind fields. *The Annals of Applied Statistics* 1(1):249–264.

Vecchi, G. A.; Zhao, M.; Wang, H.; Villarini, G.; Rosati, A.; Kumar, A.; Held, I. M.; and Gudgel, R. 2011. Statistical–dynamical predictions of seasonal North Atlantic hurricane activity. *Monthly Weather Review* 139(4):1070–1082.

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

Xu, J.; Tan, P.-N.; and Luo, L. 2014. ORION: Online Regularized multI-task regressiON and its application to ensemble forecasting. In *IEEE international conference on data mining*, 1061–1066.

Zhang, Z., and Krishnamurti, T. 1999. A perturbation method for hurricane ensemble predictions. *Monthly Weather Review* 127(4):447–469.