

APPLICATION PAPER

Predicting cyclone landfall using mutual information and dilated recurrent neural network

Abhijit Mukherjee¹  and Pabitra Mitra^{2,*}

¹Centre of Excellence in Artificial Intelligence, Indian Institute of Technology, Kharagpur, India

²Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur, India

*Corresponding author. E-mail: pabitra@cse.iitkgp.ernet.in

Received: 01 October 2022; **Accepted:** 25 October 2022

Keywords: Cyclone landfall; deep learning; mutual information

Abstract

Cyclones are a severe storm system with a defined center, occurring in the tropical regions. Upon landfall, it causes massive damage to both lives and the economy. With the increase in frequency and intensity of tropical cyclones occurring over the years and growing coastal settlements, the study of cyclone landfall remains of paramount importance for disaster control and mitigation. Cyclones experience rapid changes, with various environmental factors modulating the trajectory and intensity. Thus predicting cyclone landfall demands a highly precise technique coupled with knowledge of environmental parameters. With the complexity and nonlinearity of the cyclone track data, determining parameters conducive for the landfall prediction of a cyclone remains crucial for precision and knowledge of the storm system. While numerous methods have been employed for detecting causal interactions among weather systems like Granger Causality and Transfer Entropy, each comes with its limitation and computational overhead. In this work, we investigate the where and when of a cyclone landfall by studying the influencing factors regulating the location and time of a cyclone landfall over the North Indian Ocean with mutual information (MI). We utilize dilated recurrent neural network with gated recurrent unit cells coupled with feature selection via MI criterion for predicting the cyclone landfall location and intensity between 12 and 36 training hours. The model efficacy is validated further on the landfall data of a recent devastating storm—Fani.

Impact Statement

This paper explores cyclone landfall prediction over the North Indian Ocean by predicting the latitude, longitude, and intensity. With the help of mutual information based feature selection and a deep neural network model trained over phases of a cyclone, this study is helpful in disaster management. The study also emphasizes upon the importance of feature selection, by highlighting some important parameters conducive for the prediction task and its accuracy. The efficacy of the proposed method is further validated by the data of a recent cyclonic storm Fani.

1. Introduction

A tropical cyclone (also referred to as a hurricane or typhoon) is a rotating storm system formed over warm waterbodies in the tropical regions across the globe. After intensification over warm ocean waters, numerous cyclones advance toward coastal land where upon arrival it wreaks massive damage

to the economy and lives, before losing its intensity and characteristic structure. US National Hurricane Centre defines cyclone landfall as the “intersection of the surface center of a tropical cyclone with a coastline.” An estimate of 16.75% increase in cyclone intensity over the North Indian Ocean (NIO) has been observed over the past four decades, rendering coastal settlements vulnerable to severe damage (Albert et al., 2021). With increasing settlements along the coastline and rising ocean temperature, determining cyclone trajectory and landfall is crucial for disaster response. The task of predicting cyclone trajectory and landfall has evolved over the years from statistical models to hybrid and data-driven approaches to enhance the accuracy of the prediction. But in-situ prediction of a cyclone has always been a difficult task owing to the highly fluctuant behavior of the storms. The trajectory of a storm is not only dependent on the past track data but also on its interaction with several external environmental factors and the local orography of the coastline. Owing to land-sea friction, landfalling cyclones demonstrate rapid asymmetric transition from a highly symmetric structure (Powell, 1987; Blackwell, 2000). Thus the precise prediction of a cyclone track on landfall still remains quite challenging. Several studies on climate-cyclone interactions have used time-delayed correlation and the discovery of causal links in climate-cyclogenesis has been conducted to address this. Although time-delayed correlation analysis has been the most commonly used technique, correlation does not imply causation.

While achieving high prediction accuracy is a goal in cyclone landfall study, knowledge of the environmental factors modulating the intensity and direction is also important. Thus in this study, we intend to address the task by incorporating a feature selection method with mutual information (MI) into the prediction model that not only improves the accuracy but also directs toward possible interaction among the variables ultimately governing the changing parameters of a cyclone system. While numerous feature selection methods have been utilized for multivariate time series analysis for reducing the computational cost of predictive modeling along with an increase in accuracy, namely wrapper and filter methods, the former comes with a computation overhead and filter methods utilize statistical techniques which prove much more efficient. Under reasonable assumptions, MI has proved to increase accuracy for regression tasks (Frénay et al., 2013). We have chosen MI, a filter method for detecting a concise set of features conducive to our forecasting task. While recurrent neural network (RNN) has shown the capability of modeling nonlinear temporal relations of a hurricane, their prediction accuracy can be highly improved by including causally linked predictors specific to the task. Also, using dilated recurrent neural network (DRNN) provides much better performance over conventional recurrent models with exponentially increased dilation, dilated recurrent skip connection, and flexibility of using any recurrent units as the building block. Thus we have used DRNN with gated recurrent unit (GRU) cells for the prediction model. Data from Fani, a recent devastating storm is kept aside while training the model and is used to check the efficacy of the proposed methodology.

2. Previous Works

The problem of cyclone track prediction has been addressed typically by numerical models, statistical models, dynamic models, and hybrid models. But the dynamical, statistical, and numerical models are computationally expensive and demand huge computation power. With the recent increase in computation power of GPUs and TPUs along with more availability in tracking data from cyclones worldwide, the task of cyclone prediction has observed a more data-driven solution. Globenet (Hong et al., 2017) was designed to take 3D satellite imagery as input and predict the location of a single typhoon center by using complex convolutional neural network (CNN). Artificial neural networks were used to forecast the cyclone track from satellite images (Kovordányi and Roy, 2009). Principal component analysis (PCA) was used for cyclone eye detection (DeMaria et al., 2015). RNN was used to predict hurricane trajectory (Alemany et al., 2019). A fusion network using past track data and reanalysis data was used to predict the hurricane location (Giffard-Roisin et al., 2018). While forecasting cyclone track and eye detection have been studied, very few (Kumar et al., 2021) have addressed the issue of predicting the landfall location and

intensity of a cyclone, especially for the NIO basin which observes a significant amount of devastating cyclones each year.

Causal interaction among the units in a system is crucial to understand the underlying principles governing it. To this end, numerous studies and methods have been proposed over the years. Wiener proposed the precedence of the cause/driver over the effect/recipient. This idea of incorporating historical data in understanding the causal interaction and direction was later formalized by Clive Granger in the form of Granger Causality, built on the notion that a cause can help predict future effects beyond that which can be predicted purely based on their own previous value. But this found limited usage in systems with high nonlinearity. Thus transfer entropy was proposed that measures the reduction of signal uncertainty through probability distribution, by incorporating the historical values of both the cause and effect. But this process comes with huge computational overhead due to the curse of dimensionality. Thus MI forms a more robust yet lightweight alternative.

3. Background

A major factor in analyzing multivariate time series data is detecting temporal precedence among the features and predict the consequences of possible interaction among units in the system where interventional techniques like perturbation and manipulation seems infeasible. Granger Causality has been used extensively for this purpose, although the presence of nonlinearity renders the analysis questionable (Li et al., 2018). Filter-based alternatives like MI have been quite adept in estimating the relevance of a feature subset in predicting the target variable.

3.1. Mutual information

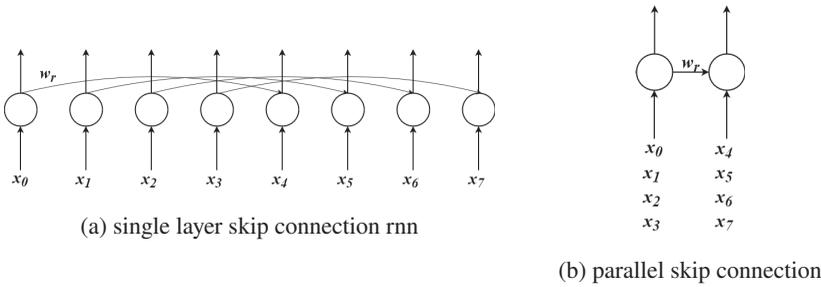
MI is a feature selection method based on information gain that gives the statistical dependence between two variables, that is the amount of information obtained from one random variable given another. It measures mutual dependence between two random variables and gives the reduction in uncertainty in learning about a variable, given another variable. Its value varies between 0 and 1, with a lower value indicating independence among the concerned variables. Given two random variables X and Y with their respective probability density function f_x and f_y and domains D_x and D_y , with the joint probability density function $f_{X,Y}$, the MI can be defined as

$$I(X;Y) = - \int_{D_x} \int_{D_y} f_{X,Y}(x,y) \log \frac{f_{X,Y}(X,Y)}{f_X(X)f_Y(Y)} dx dy. \quad (1)$$

This can be represented as $I(X;Y) = H(Y) - H(Y|X)$ where $H(X) \left(= - \int_{D_x} f_X(x) \log f_X(x) dx \right)$ is the entropy and $H(Y|X) \left(= - \int_{D_x} \int_{D_y} f_{X,Y}(x,y) \log \frac{f_{X,Y}(x,y)}{f_X(x)} dx dy \right)$ is the conditional entropy for Y given X (Frénay et al., 2013). MI has several benefits over other statistical techniques to understand the cooperative nature of a system such as: ability to detect any form of relationship between variables, both linear and nonlinear and also ease in interpreting as the amount of shared information between data sets.

3.2. Dilated recurrent neural network

While variants of the RNN have been used traditionally for various sequential learning problems, their learning range of temporal dependencies is inhibited by gradient problems and has low computational efficiency. DRNN is a multilayer, cell-independent variant of RNN, powered by *dilated recurrent skip connection* that overcomes those problems while being computationally efficient and flexible in terms of the types of RNN cells stacked together as building blocks of the neural network (Chang et al., 2017). DRNN allows parallelization as demonstrated in the following figure: (a) demonstrates skip connection in a single layer recurrent neural cell and (b) demonstrates the parallel execution of the same skip connection that reduces the sequence length by four times.



As shown in the figure, the four cells $\{c_{4t}^{(l)}\}$, $\{c_{4t+1}^{(l)}\}$, $\{c_{4t+2}^{(l)}\}$ and $\{c_{4t+3}^{(l)}\}$ for the input subsequence $\{x_{4t}^{(l)}\}$, $\{x_{4t+1}^{(l)}\}$, $\{x_{4t+2}^{(l)}\}$ and $\{x_{4t+3}^{(l)}\}$ can be computed in parallel. This greatly improves the computational cost of training when compared to a regular RNN. The dilated skip connection which connects information flow between the layers by skipping certain timesteps can be represented as

$$c_t^{(l)} = f\left(x_t^{(l)}, c_{t-s^{(l)}}^{(l)}\right), \tag{2}$$

where $c_t^{(l)}$ is the cell at layer l at time t , $s^{(l)}$ is the skip length, $x_t^{(l)}$ is the layer l input at time t , and $f()$ is any RNN operation (like LSTM, GRU, or vanilla RNN). Here the dependency on previous cell state $c_{t-1}^{(l)}$ similar to that of a regular skip connection is omitted and instead a flexible dependence on a skipped previous cell state is included. This deals away with the conventional gradient problems faced by typical RNNs, extending the range of temporal dependence. In this setting, multiple recurrent layers are stacked that may comprise of vanilla RNN, LSTM or GRU, where the dilations are increased exponentially along the layers. Thus the dilation of the l th layer $s^{(l)}$ can be represented as

$$s^{(l)} = M^{l-1}, \tag{3}$$

for $l = 1, 2, \dots, L$, where M_0^l is the starting dilation.

The efficacy of a DRNN can further be quantified using *mean recurrent length* (Chang et al., 2017) which is a measure of the average dilation across different time spans within a cycle, if we take the cyclic graph representation of a RNN. In case of DRNN, the Mean Recurrent Length is quite small as compared to a regular skip connection RNN and this implies that the past information travels along fewer edges in case of a DRNN, hence encounters much less attenuation.

4. Data set

Regional Specialized Meteorological Centers across the world are tasked by the World Meteorological Organization to record and provide the tropical cyclone track along with the Best Track Data. This comprises the best location estimates of the storm track throughout the lifetime of a cyclone along with intensity, central pressure, and so forth. For this study, the data set has been obtained from RSMC New Delhi that looks over the storms occurring over NIO, comprising storm tracks from the year 1982 to June 2020. The data comprises essential features of the storm track at each point at three hourly intervals. Along with the original features, distance and direction have been derived from the data set as additional features (Alemany et al., 2019), crucial for tracking the trajectory behavior of each storm track. A major important factor affecting the trajectory of a cyclone is Sea Surface Temperature, which is appended to the data set (Kumar et al., 2021). In this study, we address the problem of cyclone landfall and thus only the data till the landfall have been considered. Data of *Fani*, one of the strongest cyclones occurring over the NIO since 1999 have been kept aside during the training phase to test proposed model.

5. Proposed Methodology

5.1. Feature selection

Usually feature selection methods can be categorized into filter approaches (using only the data for feature selection, like the PCA) and wrapper approaches (using an algorithm for feature selection, like recursive feature elimination [RFE]). Wrapper methods evaluate multiple models with various combinations of the features to find the optimal subset of features most conducive to the task. While this exhaustively searches the feature space, it comes with a huge computation overhead. On the other hand, Filter methods use statistical techniques to evaluate the feature space based on some criterion, which is less computationally expensive. Since Tropical Cyclone Best Track Data contains temporal ordering, it is critical to identify the most important predictors from the data set. In this task, we use MI as the feature selection method to exploit the connection between causation and prediction, in that we use the historical data of a storm to identify the feature subset that gives the maximum information gain while considering our prediction tasks, namely the latitude, longitude, and intensity of a storm at the time of landfall. From the data set, MI discovered for each prediction task is demonstrated in Table 1. Features having $MI > 0.18$ have been selected (values rounded off to three decimal places).

5.2. DRNN using GRU block

For the cyclone landfall study, we have dedicated three different networks (with DRNN blocks and fitted with MI) to determine the *latitude*, *longitude*, and *intensity* of a cyclonic storm over NIO for 12, 18, 24, and 36 training hours. We have used a three-layer DRNN network with 512, 256, and 128 GRU cells in each of the layers for determining the latitude and longitude of cyclone landfall. For predicting the intensity, we have used a three-layer DRNN with 2048, 1024, and 512 GRU cells in each of the layers.

Table 1. Pairwise mutual information between features and predictand.

| Features | Latitude | Longitude | Intensity |
|-----------|--------------|--------------|--------------|
| Season | 0.157 | 0.118 | 0.045 |
| Latitude | N/A | 0.555 | 0.221 |
| Longitude | 0.552 | N/A | 0.265 |
| Cino | 0.194 | 0.204 | 1.973 |
| ECP | 0.329 | 0.365 | 1.110 |
| MSSW | 0.250 | 0.260 | N/A |
| PD | 0.269 | 0.330 | 1.876 |
| Grade | 0.167 | 0.150 | 1.814 |
| OCI | 0.074 | 0.074 | 0.118 |
| Diameter | 0.071 | 0.095 | 0.142 |
| SST | 0.435 | 0.687 | 0.291 |
| Distance | 0.698 | 0.269 | 0.176 |
| Direction | 0.486 | 0.262 | 0.052 |

Note: Bold signifies selected features from the set of all features.

Abbreviations: Cino, T. No.; ECP, estimated central pressure; MSSW, maximum sustained surface wind; OCI, outermost closed isobar; PD, pressure drop; SST, sea surface temperature.

5.3. Methodology/steps

Step 1: Cyclone Best Track Data obtained from IMD.

Step 2: Additional features' *distance* and *direction* are calculated and added to the original data set.

Step 3: Filter the data set to keep only the cyclone track data having landfall and eliminate the data from cyclones that originate and dissipate over the ocean.

Step 4: Estimate the MI to select the important features for each of the target variables latitude, longitude, and intensity.

Step 5: A DRNN with GRU cells as building block is assigned for each of the three prediction tasks.

Step 6: Check the efficacy of the proposed method on the test data set and Fani data set.

5.4. Experimental setup

PyTorch¹ has been used for designing the model. For performing the experiments, the GPU available on Google Colab has been used which allocates one of Nvidia K80, P100, or P4 depending on availability. For the intensity prediction task, the model was run for 400 epochs, with a learning rate of 0.001. For predicting the latitude and longitude, the model was run for 250 epochs, with a learning rate of 0.001. The loss criterion for training the models is MSE loss. For the cyclone intensity prediction task, we have used *Adam* optimizer and for the prediction of latitude and longitude, *Adamax* optimizer has been used. The activation function chosen for predicting the intensity of a storm at landfall is *Swish* (Ramachandran et al., 2017), a gated variant of sigmoid activation function. For the prediction tasks of Latitude and Longitude, *ReLU* activation function has been used.

6. Results

We study the Tropical Cyclones over the NIO and investigate vital parameters like the *location* (latitude and longitude) and *intensity* of a cyclone, imperative for landfall study of cyclones, disaster management, and mitigation. For predicting each of these tasks, we have extracted different sets of features (from the historical data) that are conducive to the prediction task at hand. We have designed three models with different features as input and have used DRNN (with GRU blocks) for the prediction task. Results from recent work (Kumar et al., 2021) have been considered for a comparison study, where stacked LSTM (three layers) and Bi-LSTM (three layers) have been used for modeling similar prediction objectives. We report the results of the proposed method for latitude, longitude, and intensity for a time period of 12, 18, 24, and 36 training hours. The fivefold validation accuracy in terms of RMSE and MAE has been reported to demonstrate the efficacy of the proposed technique. Data from the recently occurred cyclone Fani have been kept separately while training the model and the model performance on the landfall prediction of this cyclone is also shown for each of the prediction tasks. While the results of Longitude and Intensity show a clear improvement in prediction accuracy, the proposed method gives a comparable performance for predicting the Latitude.

The model performance in predicting the latitude is demonstrated in Table 2. Similarly, the longitude prediction for cyclone landfall is shown in Table 3. A crucial parameter in landfall study of a cyclone is the prediction of the storm intensity during landfall. The results obtained for predicting the intensity (maximum sustained wind speed) are shown in Table 4.

The discovered factors modulating the trajectory and intensity of a cyclonic storm possess some latent connection while some can be attributed to physical laws, like the *pressure drop* and *intensity* are related by Dvorak technique ($V_{\max} = 14.2 \times \sqrt{\Delta P}$, where V_{\max} is the maximum wind velocity and $\Delta P = P_0 - P_c$, with P_0 being the pressure of outmost closed isobar and P_c being the pressure at cyclone

¹ <https://pytorch.org/>.

Table 2. RMSE, MAE results for landfall latitude for different read times.

| Measure | Predictand | Method | T (Hours) | | | | |
|---------|---------------------------|----------|------------------|-------|-------|--------|------|
| | | | 4(12) | 6(18) | 8(24) | 12(36) | |
| RMSE | Fivefold cross validation | Latitude | Stacked-LSTM | 0.95 | 0.67 | 0.52 | 0.39 |
| | | | MI + DRNN | 1.16 | 0.77 | 0.65 | 0.49 |
| | Cyclone Fani | Latitude | Stacked-LSTM | 0.33 | 0.16 | 0.11 | 0.13 |
| | | | MI + DRNN | 0.24 | 0.19 | 0.14 | 0.12 |
| MAE | Fivefold cross validation | Latitude | Stacked-LSTM | 0.52 | 0.33 | 0.24 | 0.19 |
| | | | MI + DRNN | 0.60 | 0.40 | 0.30 | 0.21 |
| | Cyclone Fani | Latitude | Stacked LSTM | 0.27 | 0.11 | 0.08 | 0.10 |
| | | | MI + DRNN | 0.17 | 0.13 | 0.10 | 0.10 |

Note: Bold signifies the result of our proposed method.

Table 3. RMSE, MAE results for landfall longitude for different read times.

| Measure | Predictand | Method | T (Hours) | | | | |
|---------|---------------------------|-----------|------------------|-------|-------|--------|------|
| | | | 4(12) | 6(18) | 8(24) | 12(36) | |
| RMSE | Fivefold cross validation | Longitude | Stacked-LSTM | 1.30 | 0.85 | 0.72 | 0.50 |
| | | | MI + DRNN | 1.28 | 0.83 | 0.73 | 0.47 |
| | Cyclone Fani | Longitude | Stacked-LSTM | 0.60 | 0.36 | 0.19 | 0.23 |
| | | | MI + DRNN | 0.44 | 0.26 | 0.23 | 0.22 |
| MAE | Fivefold cross validation | Longitude | Stacked-LSTM | 0.75 | 0.46 | 0.37 | 0.29 |
| | | | MI + DRNN | 0.67 | 0.44 | 0.36 | 0.28 |
| | Cyclone Fani | Longitude | Stacked LSTM | 0.41 | 0.26 | 0.14 | 0.19 |
| | | | MI + DRNN | 0.34 | 0.21 | 0.17 | 0.18 |

Note: Bold signifies the result of our proposed method.

center). Similarly, the impacts of the cyclone diameter, grade, pressure drop, and estimated central pressure on the intensity are evident from the characteristics of a cyclonic storm.

7. Conclusion

This study of tropical cyclone landfall prediction over the NIO demonstrates the importance of feature selection for better predictive capacity and more interpretable results that can be conducive for further inquiry on cyclone trajectory and landfall. Also, the efficacy of DRNN in processing sequential signals and time series data, requiring less number of parameters has been shown in this study. While *Causality* in its purest sense is a deeply philosophical paradigm, MI provides the “predictive causality” critical for analyzing the impacts of various environmental factors on a storm trajectory and intensity. The demonstrated method can be used to discover important factors in cyclone Best Track Data and study the trajectory of landfalling cyclonic storms.

Table 4. RMSE, MAE results for landfall intensity for different read times.

| Measure | Predictand | Method | <i>T</i> (Hours) | | | | |
|---------|---------------------------|-----------|------------------|-------|-------|--------|------|
| | | | 4(12) | 6(18) | 8(24) | 12(36) | |
| RMSE | Fivefold cross validation | Intensity | Stacked-LSTM | 9.35 | 7.84 | 7.31 | 6.19 |
| | | | MI + DRNN | 8.29 | 7.40 | 5.97 | 6.56 |
| | Cyclone Fani | Intensity | Stacked-LSTM | 2.66 | 1.72 | 3.43 | 5.53 |
| | | | MI + DRNN | 2.54 | 1.12 | 0.59 | 1.27 |
| MAE | Fivefold cross validation | Intensity | Stacked-LSTM | 5.17 | 4.01 | 4.24 | 3.87 |
| | | | MI + DRNN | 4.04 | 3.10 | 2.50 | 3.31 |
| | Cyclone Fani | Intensity | Stacked LSTM | 2.03 | 1.37 | 2.64 | 4.10 |
| | | | MI + DRNN | 1.22 | 0.55 | 0.30 | 1.01 |

Note: Bold signifies the result of our proposed method.

Author Contributions. Conceptualization: A.M., P.M.; Data curation: A.M., P.M.; Funding acquisition: P.M.; Methodology: A.M., P.M.; Supervision: P.M. All authors approved the final submitted draft.

Competing Interests. The authors declare none.

Data Availability Statement. The Best Track Data of cyclones is available on RSMC, IMD site at <https://rsmcnewdelhi.imd.gov.in/>.

Ethics Statement. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Funding Statement. This research was supported by the Department of Science and Technology (DST), Government of India under the project “Machine Learning for Climate Change Analysis (MXA)” at SRIC, IIT Kharagpur.

Provenance. This article is part of the Climate Informatics 2022 proceedings and was accepted in *Environmental Data Science* on the basis of the Climate Informatics peer review process.

References

- Albert J, Krishnan A, Bhaskaran PK and Singh KS** (2021) Role and influence of key atmospheric parameters in large-scale environmental flow associated with tropical cyclogenesis and ENSO in the North Indian Ocean basin. *Climate Dynamics* 58, 17–34.
- Alemanly S, Beltran J, Perez A and Ganzfried S** (2019) Predicting hurricane trajectories using a recurrent neural network. *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 468–475.
- Blackwell KG** (2000) The evolution of hurricane Danny (1997) at landfall: Doppler-observed eyewall replacement, vortex contraction/intensification, and low-level wind maxima. *Monthly Weather Review* 128(12), 4002–4016.
- Chang S, Zhang Y, Han W, Yu M, Guo X, Tan W, Cui X, Witbrock M, Hasegawa-Johnson M and Huang TS** (2017) Dilated recurrent neural networks. *arXiv preprint arXiv:1710.02224*.
- DeMaria R, Chirokova G, Knaff J and Dostalek J** (2015) P189 machine learning algorithms for tropical cyclone center fixing and eye detection.
- Frénay B, Doquire G and Verleysen M** (2013) Is mutual information adequate for feature selection in regression? *Neural Networks* 48, 1–7.
- Giffard-Roisin S, Yang M, Charpiat G, Kégl B and Monteleoni C** (2018) Deep learning for hurricane track forecasting from aligned spatio-temporal climate datasets.
- Hong S, Kim S, Joh M and Song S** (2017) *Globenet: Convolutional neural networks for typhoon eye tracking from remote sensing imagery*. *arXiv preprint, arXiv:1708.03417*.
- Kovordányi R and Roy C** (2009) Cyclone track forecasting based on satellite images using artificial neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing* 64(6), 513–521.
- Kumar S, Biswas K and Pandey AK** (2021) Prediction of landfall intensity, location, and time of a tropical cyclone. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 14831–14839.

- Li S, Xiao Y, Zhou D and Cai D** (2018) Causal inference in nonlinear systems: Granger causality versus time-delayed mutual information. *Physical Review E* 97(5), 052216.
- Powell MD** (1987) Changes in the low-level kinematic and thermodynamic structure of Hurricane Alicia (1983) at landfall. *Monthly Weather Review* 115(1), 75–99.
- Ramachandran P, Zoph B and Le QV** (2017) Swish: A self-gated activation function. *arXiv preprint arXiv:1710.05941* 7, 1.