POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

EXECUTIVE SUMMARY OF THE THESIS

An Evaluation of Data-Driven Interpretable Methods to Detect Tropical Cyclones

LAUREA MAGISTRALE IN COMPUTER SCIENCE AND ENGINEERING - INGEGNERIA INFORMATICA

Author: MARCO ADRIANO FERRERO

Advisor: PROF. FRANCESCO AMIGONI

Co-advisors: PROF. LETIZIA TANCA, PROF. FEDERICO CERUTTI, DAVIDE AZZALINI

Academic year: 2023-2024

1. Introduction

Tropical Cyclones (TCs) are some of the most extreme weather phenomena that can occur on Earth, characterized by strong winds, torrential rains, and storm surges. They are defined as intense low-pressure systems that gain energy from warm water near the equator and rotate over tropical oceans. Their impact on coastlines and population centers can be devastating, causing valuable environmental and economic damages [15].

TCs forecasting aims to detect the occurrence and development of cyclonic activity in a particular area of our planet. Throughout the last century and up to the present day, climatologists and scientists have devised various tools to address this problem. However, due to the extreme complexity and irregularity of these phenomena, it is still difficult to fully resolve some of the open questions regarding the genesis, intensity, track, and extreme events associated with Tropical Cyclones [4]. In this context, Machine Learning (ML) techniques, which aim to learn relationships and patterns within data, have proven to be highly impactful in contributing to a deeper understanding of the topic [2]. In particular, the use of data collected over more than 50 years by the world's leading weather organizations has enabled the creation of increasingly powerful models capable of improving the predictive skills of forecasting systems. Although ML-based solutions can provide highly accurate predictive results, their behavior is often difficult to interpret. Usually, these models are very complex, and their dynamics are not transparent to the users. Even experts who use them may not fully understand the processes and evaluations that lead to a given prediction. This thesis aims to demonstrate the potential contributions of ML methods in developing interpretable systems, capable of expressing rules and explanations on the predictions that binds meteorological variables to the TCs detection.

2. Datasets

TCs are weather events that can form in various regions of the world, with slight variations and frequency depending on the basin in which they occur [7]. This study focuses on the South-West Indian Ocean basin, with geographical coordinates of 0°S, 30°S, 30°W, and 90°W, as the target area for forecasts.

The datasets considered for this work are categorized into three groups: data on large-scale global drivers, local variables related to the reference area, and data concerning the presence of TC in the target area. All the data involved are historical observation from 1980 to 2022.

2.1. Global Drivers

Global drivers are all those factors that influence global climate anomalies. Usually, they are seasonal or intraseasonal phenomena that occur regularly, and are often related to the increasing chances of observing TCs. For this study, the European Center for Medium-Range Weather Forecasts (ECMWF) provided datasets representing two well-known anomalies: the Madden-Julian Oscillation (MJO) and the El Niño-Southern Oscillation (ENSO) [3].

The first dataset describes an index to monitor the MJO, an intraseasonal (30-90 days) variability in the tropical atmosphere. It consists of large-scale coupled patterns in atmospheric circulation and deep convection that propagate slowly (5 m/s) eastward across the warm sea surfaces of the Indian and Pacific Oceans [16]. A pair of Principal Components (PCs) time series make up the index that describes this anomaly, under the name of real-time multivariate MJO series, RMM1 and RMM2 [14].

The second dataset for the global drivers observations describes the ENSO anomaly [10], a quasi-periodic event that alters wind and Sea Surface Temperatures (SSTs) in the Pacific Ocean, causing irregularities in tropical and subtropical areas. The data used to describe this phenomenon represent SSTs in specific Pacific and Indian Ocean areas.

2.2. Local Drivers

Local drivers data are related to all the observations for meteorological variables occurring in the target region. ERA5, implemented by the Copernicus Climate Change Service (C3S) within the ECMWF, is a highly advanced system that provides data on over 300 meteorological variables at various atmospheric levels[1]. This system is used to retrieve data on local meteorological variables, such as daily weather conditions for wind speed, surface pressure, and precipitation.

ERA5 records have one of the highest resolutions among systems of its kind, providing data every 0.25° displacement in latitude or longitude. As a reanalysis system, it represents in details the global atmosphere, land surface, and ocean waves from 1950 to the present [5]. Table 1 provides a detailed view of the considered variables for this work, and their corresponding atmospheric levels.

Local Variables	Atmospheric Levels
Surface Pressure	-
Sea Surface Temperature	-
Temperature	550 hPa, 300 hPa 200 hPa
Relative Vorticity	850 hPa, 550 hPa 250 hPa
Air Density	-
Wind Gust Speed	-
Wind Speed	1000 hPa, 850 hPa 300 hPa
Precipitation	-

Table 1: Selected meteoreological variables and the corresponding atmospheric levels.

2.3. IBTrACS

The third dataset describes the historical occurrences of TCs in the area of interest. It is extracted from the International Best Track Archive for Climate Stewardship (IBTrACS) [6], a public dataset that provides information on the position, intensity, and size of TCs worldwide. The archive covers a period from the beginning of the previous century to the present day. It is then essential in the problem definition for labeling samples and implementing supervised methods.

3. Data Analysis

The ERA5 reanalysis system offers high-resolution data for the target area. However, the large data dimensionality poses a challenge when training simple and low-complexity models. Therefore, a preliminary analysis is necessary to determine how to approximate the source dataset while reducing its dimensionality. One of the best solutions is to segment the target region into a few sub-areas and evaluate the mean

[1] `https://climate.copernicus.eu/climate-reanalysis`

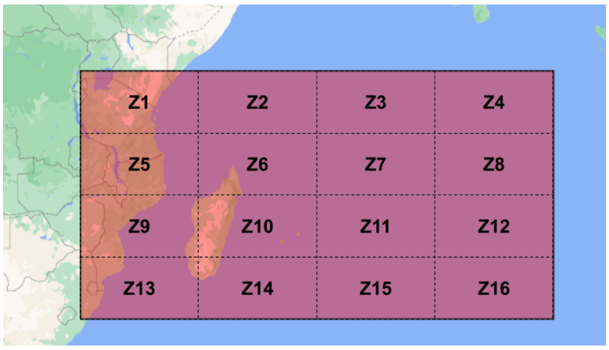

Figure 1: Segmentation of the target zone.

and standard deviation of the local drivers for each of them. In addition, selecting the most important features may help to reduce the computational effort and eliminate irrelevant information.
To address the segmentation issue, Kernel Density Estimation (KDE) [13] is applied. KDE is a technique that estimates the Probability Density Function (PDF) through data analysis. In this scenario, it evaluates how the meteorological variables distribute over the target region. PDFs are computed on different geographic segmentations of the basin, to find the best solution that could be a fair compromise between an approximation of the original ERA5 data distribution and a reasonable reduction of its dimensionality. Comparing the different PDFs obtained, the optimal solution for this work results in segmenting the target region into 16 zones, as it is displayed in Figure 1.
In addition to the region segmentation, features selection is considered to reduce the number of variables involved in the modeling process. Features selection is critical to address some of the traditional issues related to high dimensionality [9]. It enhances performance by removing redundant or unnecessary information beforehand and simplifies models, making them easier to train and interpret. The technique implemented in this work to address this challenge is the tree-based features selector. This method uses tree-based estimator construction to compute features importance based on the impurity evaluation.
Figure 2 displays the overall relevance of the meteorological variables resulted from the tree-based selection. From the evaluation of both variables and segmentation zones, wind speed, relative vorticity, surface pressure, and air density are the most relevant local variables, while the most significant zones are located in the central tropical belt, between 7.5°S and 22.5°S, and further offshore in the Indian Ocean.

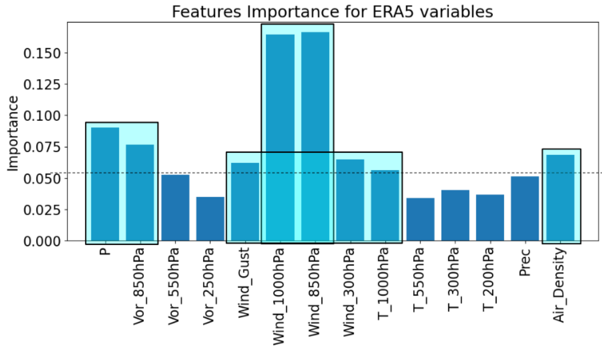

Figure 2: Tree-based features selection of local meteorological variables.

4. Problem Definition

In this thesis, the TC forecasting problem is approached as a binary classification using supervised learning techniques. The models take as input the meteorological variables selected from global and local drivers, and produce a binary output that indicates the presence or absence of a cyclone. The data input consists of time series that describe all the local drivers of ERA5 for the 10 days preceding the forecast, as well as time series that describe the global drivers for the 30 days preceding the forecast. The use of time series allows for the inclusion of information about past observations leading up to a prediction. The problem is defined to perform forecasts with varying time horizons, ranging from one to ten days after the date corresponding to the input sample.
The source dataset is split into two subsets that are used in the training phase, in which algorithms run to learn from data, and in the testing phase, when the model implemented is tested with evaluation metrics. The training set is composed of all the samples covering the period of 1980-2011, while the testing set covers the 2012-2022 years.
The evaluation metrics used for the testing phase are some of the traditional metrics adopted in ML modeling. These are the accuracy, precision, recall, and false alarm rate.

5. Black-Box Models

The first category of ML solutions implemented are black-boxes, which are characterized by their opacity concerning the relationship that binds the input to the output. The black-box models implemented for TCs detection includes methods based on neural networks and ensemble techniques. The dataset used to train and evaluate these models is composed of all the features and timesteps in the timeseries described in the Section 2, and no features selection is applied.
The three models implemented are a Gradient Boosting (GB) classifier via the XGBoost implementation, a Long Short-Term Memory (LSTM) network, and a hybrid model, which combines LSTM autoencoders and GB method.
Figure 3 presents a comprehensive comparison of the predictive capabilities of the three models on the ten time horizons considered, reporting their precisions and false alarm rates. The combination of XGBoost and LSTM autoencoders proves to be a valuable solution for enhancing robustness and generalization. This hybrid model produces a deviation of around 10% on the Precision and False Alarm Rate metrics compared to the other two solutions, making it the most effective implementation in this category.

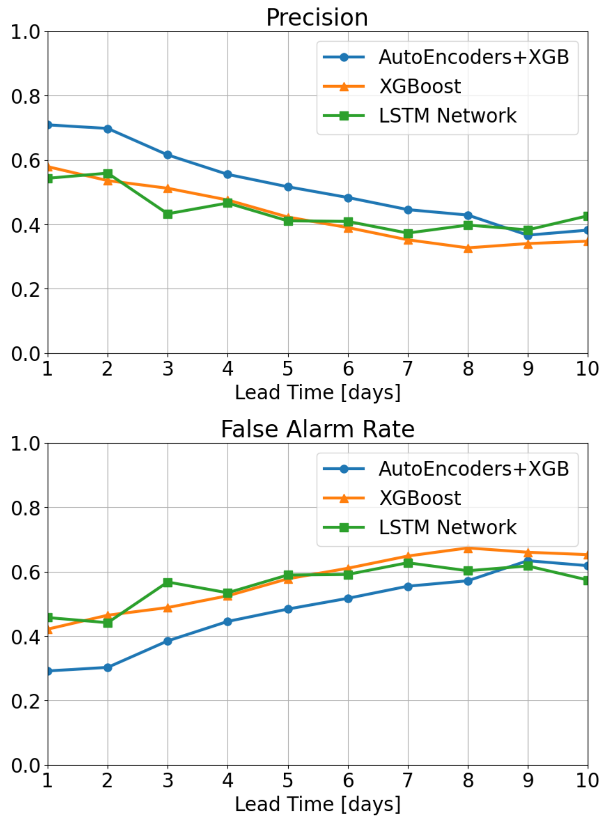

Figure 3: Precision and False Alarm Rate for the three black-box models implemented in the 10 forecasting horizons.

6. LIME

Local Interpretable Model-agnostic Explanations (LIME) [12] is a widely used technique for explaining black-box models. LIME analyzes the behavior of the model in a neighborhood of a specific instance prediction to provide interpretability. It is a model-agnostic method, meaning it can be applied to any predictive model without requiring an understanding of the learning processes or algorithms involved.
LIME is effective in describing predictions generated by black-box models. Specifically, by analyzing instances corresponding to correctly classified positive examples, we conducted a quantitative analysis to evaluate the influence of variables, sub-zones, and driver categories. The results indicate that global drivers have little influence on the model's classification. On the other hand, local variables such as wind speeds and relative vorticity have the strongest impact. Furthermore, the sub-zones that are most active correspond to those where the extreme event is present or nearby.

Although LIME can be an excellent method to provide interpretability for many examples, it may have limitations. One aspect to consider is its strong instability. For instance, similar data points can produce significantly different descriptive rules [1]. This may be a consequence of the difficulty in correctly defining a neighborhood for a data instance, an issue that can occur with large datasets.

7. White-Box Models

White-boxes are Machine Learning techniques that offer transparency in the predictive processes. They provide clear explanations of how they work, enabling experts in the application domain to gain knowledge about the resulting predictions.

7.1. Dimensionality Reduction

White-box models are interpretable by construction. However, accurately classifying high-dimensional data can be challenging [11]. The reference dataset used in this work has high di-

mensionality, making it difficult to define suitable white-boxes and use all available data. Therefore, the source datasets must be reduced. The initial reduction involves redefining the target area for the TC classification task by adjusting the geographical distribution of samples. Specifically, the new target zone is the Z11 area. This sub-region is chosen based on the features selection and the analysis of TC distribution. According to the tree-based features selection, Z11 is one of the most expressive zones for detecting TCs in the entire South-West Indian Ocean basin.
In a second step, a further reduction of the local and global variables is carried out. Large-scale drivers are not considered anymore as their contribution to final predictions is too specific for particular seasonal conditions. On the other side, only a subset of local drivers is selected, according to the tree-based feature selection results. For each meteorological driver, we consider the mean and standard deviation of all data points in Z11 as valid contributors. Therefore, the dataset to train and evaluate white-box models consists of 16 meteorological contributors and one temporal indicator (Day of the Year), which specifies the number of the day in the year of each record. In addition, these forecasting methods do not consider the contribution of variables in past time windows, but only the observation in the current day. The new dataset has 17 features per record, significantly fewer than the original. This reduction is enough to improve the white-box models, describing predictions using a compact set of variables.

7.2. Models and Results

Two solutions are implemented to address TCs forecasting with white-box models: Decision Trees (DTs) and Bayesian Rule Lists (BRLs). These two methods are both effective in providing rules or decision paths that achieve classification in a intuitive way.
DTs are sequential methods that partition the dataset through a series of decision nodes. In this paper, their construction is implemented by the CART algorithm. BRLs combine Bayesian inference techniques and rule-based systems to produce a highly interpretable model. Rule-based techniques are highly interpretable due to the ability of rules to represent global behaviors through expressions close to the natural language. The rules generated by these models consist of IF-THEN statements, where the IF introduces one or more conditions, representing a partition of the source dataset defined [8].
In both cases the results produced on the test set indicate that these models hardly provide effective predictions for cyclone detection beyond 24 hours. For this reason further analyses on larger forecasting horizons are not taken in consideration. Table 2 reports evaluation metrics for the two white-box models implemented. Both the two implementations show similar predictive capabilities on the metrics considered.

Metrics	DT	BRL
Accuracy	0.98	0.98
Precision	0.80	0.78
Recall	0.66	0.70
False Alarm rate	0.20	0.22

Table 2: Evaluation metrics of the Decision Tree and Bayesian Rule Lists at 24h forecast horizon.

Table 3 shows the most effective rules that classify positive samples in the two models. In particular, the DT path in table covers 231 samples over the 346 positive classified examples in the training set. Similarly, the rules from the BRL in table classify the positive samples with a probability of 98.8%. When comparing the two rules, it appears that some variables are related to the same threshold values, such as Vor_850hPa_Mean and P_Std. There are still consistent rules in other cases, such as T_1000hPa_Std and Vor_850hPa_Std. Furthermore, the rules generated by the Bayesian model are typically longer and contain many of the same logical formulas. Conversely, those produced by the tree algorithm explore the feature space more extensively and are restricted in length by the hyperparameter that specifies the maximum depth of the tree.
It is relevant to note that white-box techniques are implemented after appropriate dimensionality reduction and do not represent large-scale generalizations. The target zone considered for DTs and BRLs is a limited portion of the entire South West basin of the Indian Ocean and could be strongly subject to bias. Then, it is dif-

ficult to state if the same rules and paths derived from white-boxes in this thesis can detect TCs in the entire basin. However, they can provide relevant insights into the drivers and threshold values that characterize this phenomenon in this specific condition.

Decision Tree Paths
Vor_850hPa_Mean $\leq -1.32 \cdot 10^{-5} \wedge$
Vor_850hPa_Std $> 3.74 \cdot 10^{-5} \wedge$
Wind_300hPa_Std $> 7.54 \wedge$
P_Std $> 196.72 \wedge$
T_1000hPa_Std ≤ 0.42
Bayesian Rule Lists
P_Mean $\leq 101230 \wedge$
Wind_Gust_Std $> 2.11 \wedge$
Vor_850hPa_Std $> 3.26 \cdot 10^{-5} \wedge$
T_1000hPa_Std $\leq 0.92 \wedge$
Vor_850hPa_Mean $\leq -1.32 \cdot 10^{-5} \wedge$
Wind_850hPa_Mean $> 11.5 \wedge$
P_Std > 196

Table 3: Most significant classification rule for BRLs and DTs.

8. Conclusions

The importance of transparency in artificial intelligence systems, particularly in forecasting extreme weather events, is crucial for decision-making support. While data-driven approaches have improved forecasting accuracy, many current systems lack transparency, hindering their utility. This work aims to contribute to understand and improve interpretable methods for TCs forecasting, providing descriptions of the predictions that bind the input values to the TC presence.

The results obtained shows effectiveness in models forecasting. However, some of the experiments encountered several limitations and challenges, such as poor data quality, complex cyclone dynamics, and high dimensionality of meteorological data. LIME enhances explainability in black-box models, but it faces instability when involving complex data. Similarly, white-box models are able to provide interpretable rules or paths for an effective description, but they are difficult to adapt to large datasets.

New potential avenues for improving the interpretability of data-driven methods in TCs forecasting should be explored in future. These may include redefining the forecasting problem to focus on mature cyclone activity, exploring correlations between global meteorological drivers and TCs variations, and considering the implementation of state-of-the-art deep learning models like Deep Convolutional Networks and attention mechanisms to better learn from structured and time series data.

References

[1] David Alvarez-Melis and Tommi S Jaakkola. On the robustness of interpretability methods. *arXiv preprint arXiv:1806.08049*, 2018.

[2] Rui Chen, Weimin Zhang, and Xiang Wang. Machine learning in tropical cyclone forecast modeling: A review. *Atmosphere*, 11(7):676, 2020.

[3] Federica Maria De Luca. An empirical evaluation of ai-based methods for modeling the genesis of tropical cyclones. Master's thesis, Politecnico di Milano, 2023.

[4] Kerry Emanuel. 100 years of progress in tropical cyclone research. *Meteorological Monographs*, 59:15–1, 2018.

[5] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

[6] Kenneth R Knapp, Michael C Kruk, David H Levinson, Howard J Diamond, and Charles J Neumann. The international best track archive for climate stewardship (ibtracs) unifying tropical cyclone data. *Bulletin of the American Meteorological Society*, 91(3):363–376, 2010.

[7] Sally L. Lavender and John L. McBride. Global climatology of rainfall rates and lifetime accumulated rainfall in tropical cyclones: Influence of cyclone basin, cyclone intensity and cyclone size. *International Journal of Climatology*, 41(S1):E1217–E1235, 2021.

[8] Benjamin Letham, Cynthia Rudin, Tyler H. McCormick, and David Madigan. Interpretable classifiers using rules and Bayesian analysis: Building a better stroke prediction model. *The Annals of Applied Statistics*, 9(3):1350 – 1371, 2015.

[9] Jundong Li, Kewei Cheng, Suhang Wang, Fred Morstatter, Robert P. Trevino, Jiliang Tang, and Huan Liu. Feature selection: A data perspective. *ACM Computing Surveys*, 50(6), 2017.

[10] Michael J McPhaden, Stephen E Zebiak, and Michael H Glantz. Enso as an integrating concept in earth science. *Science*, 314(5806):1740–1745, 2006.

[11] Christoph Molnar. *Interpretable Machine Learning*. Leanpub, 2 edition, 2022.

[12] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, page 1135–1144, 2016.

[13] Stanisław Węglarczyk. Kernel density estimation and its application. In *Proceeding of the ITM web of conferences*, volume 23, page 00037. EDP Sciences, 2018.

[14] Matthew C Wheeler and Harry H Hendon. An all-season real-time multivariate mjo index: Development of an index for monitoring and prediction. *Monthly weather review*, 132(8):1917–1932, 2004.

[15] Joseph A. Zehnder. "tropical cyclone". encyclopedia britannica, 2024.

[16] Chidong Zhang. Madden-julian oscillation. *Reviews of Geophysics*, 43(2), 2005.