

EMERGING TECH CONFERENCE – Edge Intelligence

Volume 03, 2024, pages 39-45

Proceedings of Emerging Tech Conference:
Edge Intelligence 2024

Short-term weather forecasting in
maritime environments exploiting a ML-based model at the Edge

Vasiliki Naskari^{1*}, Ioannis Masklavanos, Stavros Adam and Gregory Doumenis

¹Department of Informatics and Telecommunications, University of Ioannina, Arta, Greece

naskvasil@uoi.gr

Abstract

This study deals with short-term weather forecasting in maritime environments using machine learning. Data from autonomous ship sensors, such as, temperature, humidity, wind speed, electrostatic field, and sea conditions, were used to enhance forecast accuracy. A Radial Basis Function (RBF) model was developed to tackle nonlinear structure of the data. Data preprocessing and model development were performed using MATLAB, while model evaluation was conducted in Python. The dynamic modeling approach that was adopted demonstrated significant accuracy improvements over static models, indicating the potential for enhancing maritime safety and operational efficiency with automated, energy-efficient systems.

1 Introduction

1.1. Background

Marine weather forecasting is crucial for safeguarding seafarers and optimizing maritime operations. Accurate forecasts prevent accidents, guide navigation, and conserve shipping resources. Traditionally, expert forecasters manually compile detailed bulletins on wind, sea states, visibility, and significant weather events, usually covering large areas. Data science and machine learning advancements now enable potentially more accurate, automated forecasting processes. Such processes utilize analysis of data from sensors on ships, capturing real-time metrics like pressure, temperature, humidity, wind speed, and sea conditions (K.U. Jaseena, 2022), (EFOS, n.d.). Integrating these sensors with machine learning algorithms intends to boost the creation of automated weather prediction systems which process data with high accuracy and minimal delay (Aakash Parmar, 2017). Weather forecasting for maritime operations faces challenges from the unpredictable nature of wind and waves. Researchers and practitioners have developed various methods to tackle these challenges, categorized into physical, statistical, and machine learning approaches. Physical methods use meteorological data and physical laws to model weather conditions, often applying empirical models for wave and wind predictions (Foley Aoife M, 2012) (Kazeminezhad M.H., 2005). Statistical methods analyze historical data to find patterns, effective for trends but limited by environmental unpredictability (Zuluaga Carlos D., 2015) (Lydia M, 2016) (Poggi P., 2003). Machine learning approaches while adopting the statistical processing view, however, they process

** Corresponding author

large datasets to uncover complex, dynamic relationships, offering flexibility and superior accuracy over traditional methods (Mengning Wu, 2019). By integrating these diverse approaches, researchers strive to enhance forecast reliability and maritime safety (Özger Mehmet, 2007) (Chang G.W., 2017) (Berbić Jadran, 2017).

1.2. Problem Statement

Our aim is to develop a system able to provide short-term local weather predictions (nowcasts), with emphasis on abrupt/adverse phenomena. We differentiate from similar studies, by adding the Static Atmospheric Electric Field measurement in our dataset. Additionally, the study seeks to improve the entire data pipeline, from collection to analysis. By implementing a tailored data-cleaning technique and developing an advanced machine learning algorithm, we aim to enhance the accuracy and reliability of weather predictions, crucial for safe maritime operations.

2 Methodology

2.1. Data Collection

The initial phase of our analysis involved data cleaning to ensure the reliability and accuracy of our data, which was sourced from the Azure Databricks platform (Microsoft, n.d.). This platform provided a scalable environment essential for handling large datasets. We focused on key meteorological variables critical to our predictive model, including atmospheric pressure, humidity, temperature, electrostatic field, wind speed, weather conditions, and timestamps. During data cleaning, we checked each data point for validity and consistency, removing inconsistencies or irrelevant information that could affect our analysis. Measurement errors or outliers were corrected to maintain the dataset's integrity. For gaps in data, we used linear interpolation for short interruptions and removed any segments where data was missing for longer than five minutes to preserve dataset coherence. Categorical data were converted into numerical formats to facilitate detailed analysis.

Subsequently, we conducted a correlation analysis to identify relationships between the variables using Pearson's and Spearman's correlation coefficients. This step aimed to detect both linear and non linear relationships within the data. The initial correlation results were weak, prompting further investigations using dynamic time warping, which confirmed the presence of non-linear relationships critical for developing models to accurately predict complex weather patterns. The data, summarized in Table 1, include:

Parameter	Unit	Measurement Range
Static Atmospheric Electric Field	KV/m	-20 - 20
Air Humidity	%	0 - 100
Air Temperature	°C	-30 - 60
Air Pressure	hPa	800 – 1100
Wind True Speed	m/sec	0 - 20
Weather Condition	-	Discrete Categories
Timestamp	-	Continuous Time

Table 1: Timeseries characteristics

This structured approach ensures an analysis framework, leveraging both traditional statistical techniques and advanced machine learning in a MATLAB environment to enhance the accuracy and depth of our weather forecasting models.

2.2. Systemic approach

To address the challenges of analyzing time series data and predicting the evolution of phenomena captured by these series, various techniques from classical statistics, such as linear and non-linear regression, have been proposed. However, our initial evaluations using these methods did not yield satisfactory results. In contrast, the field of machine learning offers numerous effective approaches for predicting system behavior, particularly when non-linear relationships exist between input data and the phenomena under study. Techniques such as Artificial Neural Networks, Bayesian Models, kernel machines, and Gaussian processes have shown particular promise. These advanced techniques, regardless of their origin—statistical learning or biological learning—have demonstrated exceptional performance in handling data analysis problems, especially those where data is infected with noise, errors, or missing values. Bayesian methods, although effective in many applications, did not seem adequate for our case due to unknown a priori distributions of input data.

Neural networks have been successfully applied to various problems, including pattern recognition, classification, regression, and dynamic system modeling. Our literature review on best practices for using neural networks revealed that radial basis function networks (RBF networks) effectively meet our requirements. These networks are feed-forward architectures similar to Multi layer Perceptrons (MLP) but with the added capability of processing inputs based on kernel functions.

2.3. Experimental Setup

Upon completing the data pre-processing stage, the next phase involved selecting the appropriate approach for applying the predictive algorithm to forecast weather conditions in maritime contexts. Two approaches were considered: a static approach and a dynamic approach. In the static approach, the predictive model is trained once using the available data. This method offers several advantages such as computational efficiency, as it minimizes the need for repeated model adjustments, thereby conserving resources and time. Suitability for scenarios where the underlying data distribution remains relatively stable over time, reducing the need for frequent updates. However, the static approach has notable limitations. It is highly sensitive to concept drift, where the underlying data distribution changes over time, rendering the model outdated and less accurate. This limitation prevents the model from adapting to new patterns and trends that may emerge in the data, thus compromising its robustness in dynamic environments. The inability to respond to evolving information significantly impairs the effectiveness of static models, especially in areas characterized by constant change.

Conversely, the dynamic approach involves periodic retraining of the predictive model to incorporate new data. This methodology's primary strength lies in its adaptability to changing data distributions and evolving patterns. By regularly updating the model with new information, the dynamic approach ensures that the predictive model remains aligned with the underlying data, thereby enhancing its predictive accuracy.

3 Case Study: Weather Condition Forecast

3.1. Introduction to the Case Study

Accurate weather forecasts are essential for maritime operations, as they directly affect navigation safety, route planning, and overall efficiency. The "weather condition" metric was selected for prediction due to its comprehensive nature, which helps in making more informed maritime decisions. This case study focuses on demonstrating how dynamic predictive models can be used to forecast weather conditions at sea, comparing their accuracy to traditional methods, and exploring how automated forecasts can improve maritime safety and operational efficiency.

The analysis method used is highly flexible and can be applied to any subset of weather data. The process involved a rolling window approach, where each window contained five minutes of data. We calculated key signal statistics of the electrostatic field for each window, including arithmetic mean, sign of the mean, signal energy, variance, skewness and curvature. These statistics were then combined with correlated variables from other weather data.

To prepare the data for forecasting, we shifted the historical weather condition data by 60 positions (equivalent to 60 minutes). This transformation helps train the model to predict future weather conditions, adjusted to the time interval defined by the shift. A representation of this process is illustrated in Figure 1.

		Time				
		n	n+1	n+2	n+3	...
Air Relative Humidity	
Air Temperature	
Air Pressure	
Wind True Speed	
Static Atmospheric Electric Field Statistics	Power
	Mean
	Sign
	Variance
	Skewness
	Curvature
Weather Condition	
Shifted Weather Condition		Weather Condition(n+60)	Weather Condition(n+61)	Weather Condition(n+62)	Weather Condition(n+63)	

Figure1: Preprocessed Data Table

Given the challenges discussed in Section 2.2, along with the test results and data characteristics, we chose to develop a Radial Basis Function (RBF) model with a Kernel filter. This model was selected for its ability to handle non-linear relationships and provide accurate predictions, making it ideal for forecasting weather conditions in dynamic maritime environments.

4 Results

In the static approach, we forecasted the weather condition metric 60 minutes into the future. This method yielded an accuracy of **57.32%**. While this accuracy highlights the model’s ability to provide general trends, it lacks the precision of dynamic models, as reflected by the comparative results.

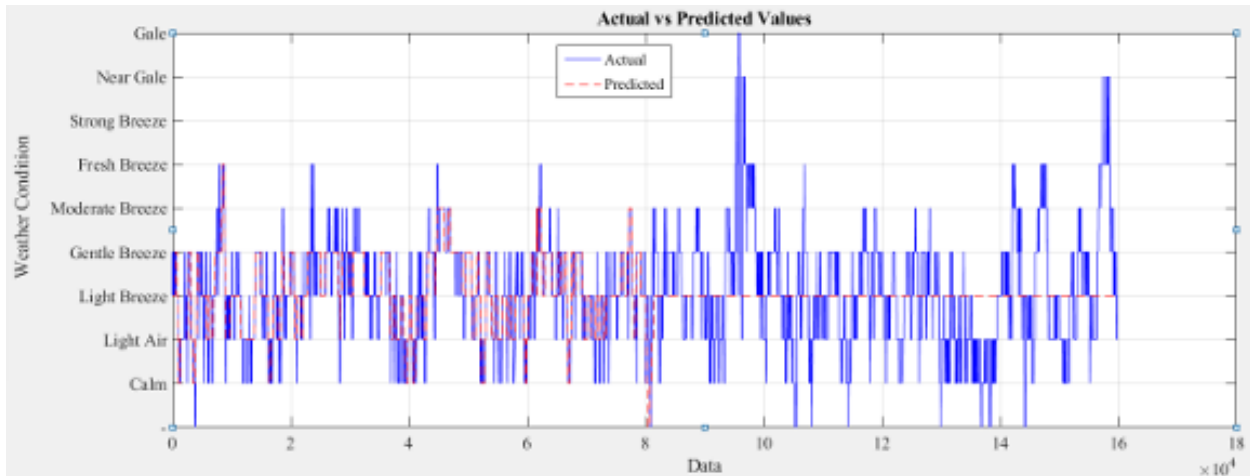


Figure2: Accuracy Results for The Static Model

1st dynamic approach (Prediction 60min Forward for a 20min Forecast Period)

In this dynamic approach, we predicted the weather condition 60 minutes ahead, focusing on a forecast window of 20 minutes. This method significantly improved prediction accuracy, achieving **76.98%**. The model was better at capturing fluctuations in the weather conditions over shorter forecast periods.

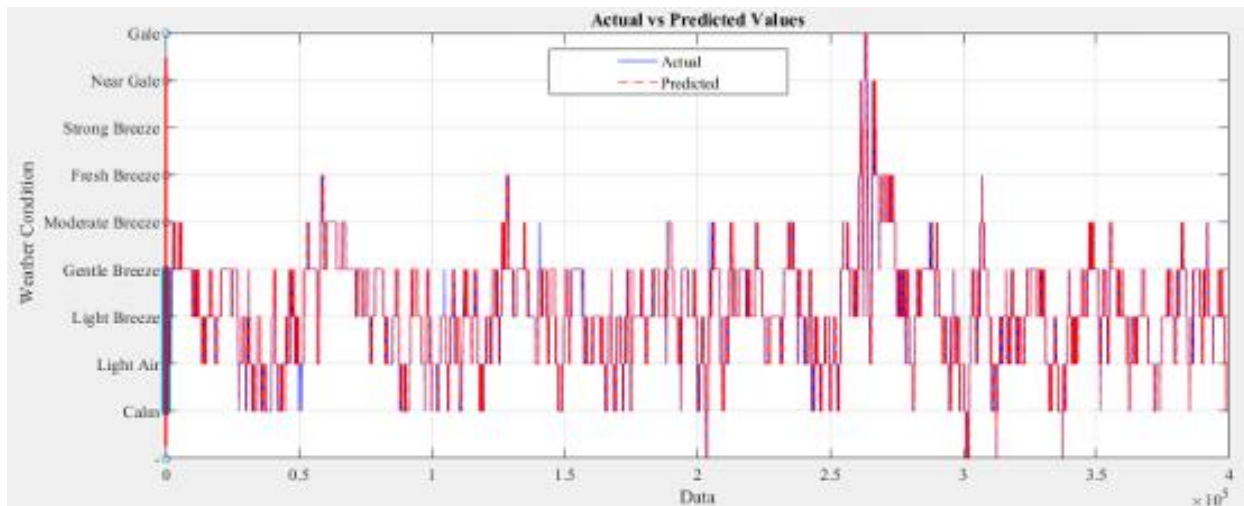


Figure3: Accuracy results for the 1st dynamic approximation model

2nd dynamic approach (Prediction 60min Forward for a 15min Forecast Period)

Further refining the forecast window to 15 minutes led to the highest accuracy of **83.17%**. This model demonstrated a strong capability in accurately forecasting short-term weather conditions, making it the

most effective approach in our study.

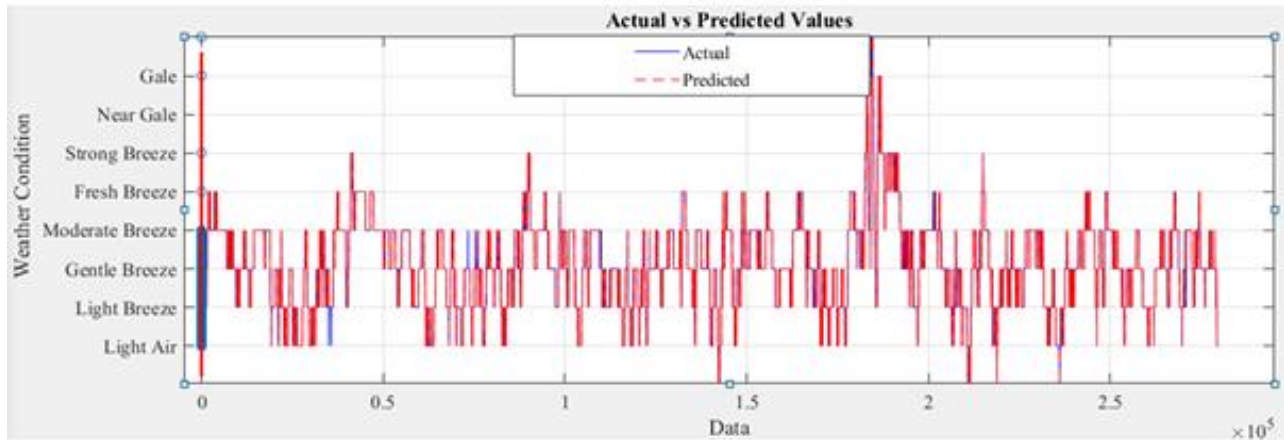


Figure4: Accuracy results for the 2nd dynamic approximation model

The model was tested in Python, and the following metrics were measured: accuracy, mean squared error (MSE), root mean squared error (RMSE), and correlation coefficient. Below is a comparative table of the results for the static approach and the two dynamic models, with the dynamic models showing considerable improvements across all evaluation metrics.

Approach	Static	20min (1st dynamic approach)	15min (2nd dynamic approach)
Accuracy (%)	49.4972	76.98	83.17
MSE	1.1145	0.2555	0.1912
RMSE	1.0557	0.5054	0.4372
Cor. Coefficient	0.493648	0.9014	0.9276

Table 2: Performance evaluation

4.1. Discussion

The study revealed that dynamic predictive models, which incorporate periodic retraining, significantly outperform static models in both accuracy and adaptability. Their ability to integrate new data ensures precise predictions, even in changing weather conditions, making them highly effective for maritime operations. Additionally, the low energy requirements of these models make them ideal for use in resource-limited maritime environments. This energy efficiency allows continuous operation without straining ship power systems. Accurate weather forecasting not only enhances safety by providing timely warnings but also optimizes route planning, reducing fuel consumption and operational costs. The combination of these factors supports more efficient, resilient, and cost effective maritime logistics.

5 Conclusions

This paper presents an application of dynamic, non-linear predictive modeling on a weather dataset employing (static) atmospheric electric field measurements. The application of RBF based predictive models showed notable improvements particularly in terms of accuracy and adaptability. The execution of the model is particularly efficient in terms of time highlighting its potential for integration into energy-

efficient forecasting systems. Using more than 18 months of historical data, the system exhibited over 83% nowcast accuracy. However, a key limitation is the model's inability to accurately track fast and abrupt weather changes.

Thus, having the dynamic model perform well in the short term/low variance conditions, future research should explore hybrid models that incorporate large historic datasets and real-time data assimilation for greater accuracy. Our work could set the stage for advancing towards hybrid predictive models, leveraging the density of measuring stations via federated learning. The system architecture, designed to support both edge and cloud computing, enhances scalability and real-time forecasting capabilities. This flexibility will further strengthen the accuracy and efficiency of maritime weather forecasting, paving the way for more robust and reliable systems in the future.

References

- Aakash Parmar, K. M. (2017). Machine Learning Techniques For Rainfall Prediction: A Review. International Conference on Innovations in information Embedded and Communication Systems (ICIIECS).
- Berbicć Jadran, O. E. (2017, July). Application of neural networks and support vector machine for significant wave height prediction. *Oceanologia*, pp. 331-349.
- Chang G.W., L. H. (2017). An improved neural network-based approach for short-term wind speed and power forecast. *Renewable Energy*, pp. 301-311.
- EFOS. (n.d.). Retrieved from <https://efos.site/>
- Foley Aoife M, L. P. (2012, January). Current methods and advances in forecasting of wind power generation. *Renewable Energy*, pp. 1-8.
- K.U. Jaseena, B. C. (2022, June). Deterministic weather forecasting models based on intelligent predictors: A survey. *Journal of King Saud University - Computer and Information Sciences*, pp. 3393-3412.
- Kazeminezhad M.H., E.-S. A. (2005, October). Application of fuzzy inference system in the prediction of wave parameters. *Ocean Engineering*, pp. 1709-1725.
- Lydia M, S. K. (2016, March 15). Linear and non-linear autoregressive models for short-term wind speed forecasting. *Energy Conversion and Management*, pp. 115-124.
- Mengning Wu, C. S. (2019). Prediction of short-term wind and wave conditions for marine operations using a multi-step-ahead decomposition-ANFIS model and quantification of its uncertainty. *Ocean Engineering*.
- Microsoft. (n.d.). Azure Databricks. Retrieved from <https://azure.microsoft.com/en-us/products/databricks>.
- Özger Mehmet, Ş. Z. (2007, December). Prediction of wave parameters by using fuzzy logic approach. *Ocean Engineering*, pp. 460-469.
- Poggi P., M. M. (2003, December). Forecasting and simulating wind speed in Corsica by using an autoregressive model. *Energy Conversion and Management*, pp. 3177-3196.
- Zuluaga Carlos D., Á. M. (2015, October). Short-term wind speed prediction based on robust Kalman filtering: An experimental comparison. *Applied Energy*, pp. 321-330.