


FORECASTING WIND DIRECTION IN ALOR SETAR USING MACHINE LEARNING TIME SERIES MODELS WITH TRIGONOMETRIC TRANSFORMATION

Nur Arina Bazilah Kamisan ^{1*}, Pow Jing Huei ², Muhammad Hisyam Lee ³

^{1,2,3}Department of Mathematical Science, Faculty of Science, Universiti Teknologi Malaysia
Johor Bahru, Johor, 81310, Malaysia

Corresponding author's e-mail: * nurarinabazilah@utm.my

Article Info	ABSTRACT
<p>Article History: Received: 25th August 2025 Revised: 7th January 2026 Accepted: 12th March 2026 Available online: 8th April 2026</p> <p>Keywords: Directional statistics; Prophet; Random Forest; Time Series; Wind Direction Forecasting.</p>	<p>Forecasting wind direction is inherently challenging due to its circular nature, where conventional numerical models often encounter discontinuities at the 0°/360° boundary. This study compares two modelling strategies for daily wind direction prediction in Alor Setar, Malaysia, using data from 2013–2017. A transformation-based approach and a direct numerical approach are compared for forecasting wind direction to assess their differences. In the transformation-based method, wind direction values are converted into sine and cosine components to preserve circularity, with predictions later reconstructed using inverse trigonometric functions. The direct approach predicts wind direction values without transformation. Three models, Prophet, Random Forest, and Holt-Winters, are applied under both strategies. Model performance is evaluated using time series plots, wind rose diagrams, and angular error metrics, including Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Results indicate that the Random Forest model is the best model for forecasting the wind direction in Alor Setar, and the transformation-based approach produces more accurate and stable predictions, effectively capturing directional continuity, while the direct approach yields higher angular errors and fails to replicate the observed wind direction distribution. To our knowledge, this is one of the first studies in Malaysia to systematically apply transformation-based approaches for wind direction forecasting. The findings highlight the practical importance of improved wind direction prediction for renewable energy optimization, aviation safety, and environmental monitoring.</p> <p> This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License.</p>

How to cite this article:

N. A. B. Kamisan, P. J. Huei, and M. H. Lee, "FORECASTING WIND DIRECTION IN ALOR SETAR USING MACHINE LEARNING TIME SERIES MODELS WITH TRIGONOMETRIC TRANSFORMATION", *BAREKENG: J. Math. & App.*, vol. 20, no. 3, pp. 2363-2374, Sep, 2026.

Copyright © 2026 Author(s)

Journal homepage: <https://ojs3.unpatti.ac.id/index.php/barekeng/>

Journal e-mail: barekeng.math@yahoo.com; barekeng_journal@mail.unpatti.ac.id

Research Article · **Open Access**

1. INTRODUCTION

Advances in machine learning have led to a major improvement in the accuracy of wind forecasting, which is crucial for many services and safety [1]. One of the important aspects of wind forecasting is for the proper operation and optimization of wind energy plants. However, due to the country's tropical climate, varied geography, and land use patterns, reliable wind forecasting is difficult. Malaysia has a tropical climate, which means it has high temperatures, high humidity, and a lot of rain. These climatic conditions can cause abrupt fluctuations in wind patterns, making wind predictions at specific places challenging [2]. Additionally, Malaysia's complex topography and land use patterns can also affect wind flow patterns, adding another layer of complexity to wind forecasting [3]. In wind forecasting, there are two aspects that need to be considered: wind speed and wind direction.

Wind direction forecasting is a crucial aspect of weather prediction, significantly impacting various sectors such as agriculture, aviation, and renewable energy. In Malaysia, a country characterized by diverse topography and weather patterns, accurate wind direction forecasting holds particular importance. Statistical models have been extensively used in wind direction forecasting, and currently, machine learning models are getting more attention due to their simplicity and effectiveness. The work by Yatiyana et al. [4] addresses the challenge of unpredictable wind power generation in electrical systems due to fluctuating wind patterns. To enhance energy supply efficiency, the study develops ARIMA-based models for predicting wind turbine output by focusing on wind speed and direction. The choice of ARIMA is motivated by its quick responsiveness. Using real data in case studies, the models achieve less than 5% error for wind speed and 16% for wind direction forecasting.

Accurate forecasting of wind direction is essential for a wide range of applications, including renewable energy integration [5], where optimal turbine alignment and energy output depend on anticipating prevailing winds; aviation operations in which runway selection, flight path planning, and safety protocols are influenced by wind orientation [6], maritime navigation where vessel routing, docking maneuvers, and fuel efficiency are impacted [7], and environmental monitoring particularly for tracking the dispersion of pollutants, wildfire smoke, or airborne particulates [8]. Unlike linear variables, wind direction is inherently circular, meaning that 0° and 360° denote the same orientation. This cyclical property introduces mathematical complexities, as conventional statistical and machine learning forecasting techniques are designed for linear scales, often resulting in discontinuities or inflated errors when values approach the $0^\circ/360^\circ$ boundary [9]. Addressing this circularity is therefore a critical step in developing accurate and robust wind direction prediction models.

Many studies have focused on wind speed forecasting, including those in Malaysia [10]–[14], but very limited attention has been given to wind direction forecasting, particularly when using machine learning methods that properly account for its circular nature. Most existing works rely on linear models that encounter discontinuities at the $0^\circ/360^\circ$ boundary. To bridge this gap, this study systematically compares direct and transformation-based forecasting approaches using Prophet, Holt–Winters, and Random Forest models. The contribution of this paper lies in demonstrating how trigonometric transformations can improve predictive accuracy for circular data within a Malaysian context.

Similarly, Jing-Jing et al. [15] utilized a statistical model, which is circular-circular regression, a statistical technique for circular data like wind direction, to improve wind direction forecasts for turbine height. The model applies Mobius transformation to link predicted and observed wind directions, correcting biases in numerical weather prediction models. This approach aims to enhance accuracy in predicting wind direction, which is crucial for optimizing wind turbine performance. The study by Erdem & Shi [16] compares different methods for short-term wind forecasting. It finds that decomposing wind speed with ARMA models works best for predicting wind direction, while traditional ARMA models perform better for wind speed. However, vector autoregression (VAR) models show higher accuracy for wind direction and comparable results for wind speed, especially with moderate speed-direction correlations. Using a restricted VAR approach can create more concise models, but it needs careful construction for sustained forecasting power.

There are several studies by researchers in Malaysia, but they mostly focused on the pattern or behavior of the wind direction instead of forecasting the wind direction [17]–[20]. Wind direction forecasting might have received less attention, not only in Malaysia but in many other countries, due to the angular characteristics of the data, leading to difficulties in forecasting the data with linear models. Researchers might prioritize other weather aspects like rainfall prediction or severe weather tracking, which have more immediate societal impacts. Existing models providing acceptable accuracy for wind direction could also

reduce the urgency for extensive research. Limited funding, collaborations with international organizations, and perceived economic priorities might further influence this focus.

2. RESEARCH METHODS

To address the challenges posed by the circular nature of wind direction, two primary modelling strategies are the transformation method and the direct forecasting method. The transformation method involves representing wind direction θ as two orthogonal components on the unit circle using trigonometric functions $x = \cos \theta$, $y = \sin \theta$. This transformation maps each angular measurement to a point on the circumference of a unit circle, effectively removing the discontinuity between 0° and 360° . Forecasting models are then trained on these x and y components rather than the raw angular values. Once predictions are obtained, the original directional value is reconstructed using the two-argument arctangent function, and the atan2 function ensures correct placement in all quadrants, thereby preserving directional accuracy. This approach allows conventional statistical and machine learning models designed for linear data to process wind direction without encountering boundary discontinuities or inflated errors.

The direct forecasting method, in contrast, predicts wind direction in degrees (or radians) directly, without any prior transformation. This approach is straightforward and avoids the additional computational step of transforming and reconstructing the variable. However, it treats the angular variable as linear, which can lead to significant inaccuracies near the wrap-around point. For example, a prediction of 1° compared with an actual value of 359° would appear to have a large error (358°) in the linear sense, despite the true difference being only 2° . Consequently, while the direct method can be simpler to implement, it often produces unstable forecasts and larger angular errors, especially when wind direction fluctuates across the $0^\circ/360^\circ$ threshold.

In practice, the transformation method is widely regarded as more robust for circular variables like wind direction, as it maintains directional continuity and avoids misinterpretation of boundary-adjacent values. The direct method remains relevant in certain applications where computational simplicity is prioritized, or where directional variability is minimal, but it requires careful handling of error calculations to produce meaningful results. In this study, we evaluate and compare the effectiveness of both approaches using three popular forecasting models: Random Forest, Holt-Winters, and Prophet. The dataset comprises daily wind direction measurements in Alor Setar, Malaysia, with training data from 2013 to 2016 and testing data from 2017. The objective is to determine whether circular transformation improves forecasting accuracy and compare the prediction quality of selected methodologies.

The three models, Random Forest, Holt-Winters, and Prophet, represent a mix of machine learning and statistical methods commonly used in time series forecasting [27]. Random Forest was chosen for its ability to capture nonlinear dependencies [24], Holt-Winters for its effectiveness in seasonal decomposition, and Prophet for its robustness to missing data [26] and adaptable trend-seasonality modeling [27]. Although circular-specific models exist, we aimed to evaluate the performance of widely used forecasting models when applied with and without trigonometric transformation. Before modeling, missing values were checked and interpolated where necessary, and lag structures for Random Forest were optimised through validation experiments.

In this paper, the wind direction values (in degrees) are first transformed into their corresponding sine and cosine components to account for the circular nature of the data. Specifically, each wind direction θ (in degrees) is converted using the following trigonometric functions in Eqs. (1) and (2):

$$\sin(\theta) = \sin\left(\frac{\pi \times \theta}{180}\right), \quad (1)$$

$$\cos(\theta) = \cos\left(\frac{\pi \times \theta}{180}\right). \quad (2)$$

This transformation maps the circular variable onto the unit circle, enabling the use of conventional regression and forecasting methods that assume linearity. Once the sine and cosine values are forecasted, the predicted wind direction can be reconstructed using the arctangent function in Eq. (3) [21]:

$$\hat{\theta} = \begin{cases} \left(\tan^{-1} 2 \left(\hat{\sin}(\theta), \hat{\cos}(\theta) \right) \times \frac{180}{\pi} \right) + 360, & \text{if } \tan^{-1} 2 \left(\hat{\sin}(\theta), \hat{\cos}(\theta) \right) \times \frac{180}{\pi} < 0, \\ \tan^{-1} 2 \left(\hat{\sin}(\theta), \hat{\cos}(\theta) \right) \times \frac{180}{\pi}, & \text{otherwise.} \end{cases} \quad (3)$$

This reconstruction ensures that the predicted values remain within the range $[0^\circ, 360^\circ)$, preserving the circular characteristics [22].

2.1 Random Forest Model

Random Forest (RF), proposed by [23], is an ensemble learning method that constructs multiple decision trees using bootstrap samples and aggregates their predictions to improve accuracy and robustness. While RF is originally designed for independent and identically distributed data, it can be adapted for time series forecasting by creating lagged input features to capture temporal dependencies [24]. To predict the value at time t , the model uses previous observations $t - 1, t - 2, \dots, t - p$ as predictors as in Eq. (4):

$$X_t = [y_{t-1}, y_{t-2}, \dots, y_{t-p}], y_t = f(X_t) + \epsilon_t. \quad (4)$$

Here, X_t denotes the feature vector comprising the previous p lagged values of the target variable, y_t is the forecasted value at time t , and f represents the nonlinear function approximated by the RF model.

RF builds an ensemble of decision trees $\{T_1, T_2, \dots, T_M\}$, each trained on a different bootstrap sample from the original dataset [25], where M is the total number of trees. During prediction, the forecasts from all trees are averaged to obtain the final output as in Eq. (5):

$$\hat{y}_t = \frac{1}{M} \sum_{m=1}^M T_m(X_t). \quad (5)$$

2.2 The Holt-Winters Model

The Holt-Winters exponential smoothing method is a time series forecasting technique that extends exponential smoothing to capture trends and seasonality [26]. It is particularly useful for data with regular patterns over time, such as seasonal effects. There are two main versions of the Holt-Winters method: additive and multiplicative. The additive model is suitable when seasonal variations are roughly constant in magnitude, whereas the multiplicative model is appropriate when seasonal fluctuations increase or decrease proportionally with the level of the series [27]. Since the data used in this study do not exhibit a trend component, the additive Holt-Winters model without trend was employed. This version focuses on modeling the level and seasonal components only. The model is defined by the following equations:

$$\text{Level: } L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1}), \quad (6)$$

$$\text{Trend: } b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}, \quad (7)$$

$$\text{Seasonal: } S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-s}, \quad (8)$$

$$\text{Forecast: } F_{t+1} = L_t + b_t + S_{t+1-s}, \quad (9)$$

where α = level smoothing parameter ($0 \leq \alpha \leq 1$), β = trend smoothing parameter ($0 \leq \beta \leq 1$), γ = seasonal smoothing parameter ($0 \leq \gamma \leq 1$), s = seasonal length, Y_t = actual data at period t , L_t = level smoothing at period t , b_t = trend smoothing at period t , S_t = seasonal smoothing at period t , and F_{t+1} = forecasting at period $(t + 1)$.

2.3 Prophet Model

Prophet is an open-source forecasting tool developed by Facebook designed to handle time series data with strong seasonal patterns and missing values [27]. It decomposes a time series into three main components: trend, seasonality, and holiday effects, allowing for flexible modeling of real-world time series. The general form of the Prophet model is expressed as in Eq. (10):

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t, \quad (10)$$

where $g(t)$ represents the trend component, $s(t)$ captures seasonality (daily, weekly, yearly), $h(t)$ accounts for effects of holidays or special events, and ϵ_t is the error term. The seasonal component $g(t)$ can be

modelled using a Fourier series to capture periodic effects. Prophet uses the following formulation in Eq. (11) for seasonal effects:

$$s(t) = \sum_{n=1}^N \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right) \right), \quad (11)$$

where P is the period, N is the number of Fourier terms, and a_n, b_n are coefficients learned from the data.

2.4 Error Measurement

Model evaluation is essential to assess the accuracy and reliability of forecasting methods. Two widely used error metrics are the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE) [28]. These metrics are commonly applied in model fitting, validation, selection, comparison, and forecast performance evaluation [29]. They quantify the differences between observed and predicted values, with lower values indicating better model performance. In this study, we have categorized them into the transformed and direct approaches. For the direct approach, the error will be calculated using the MAE in Eq. (12) and RMSE in Eq. (13). They are defined as:

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^N |Y_t - F_t|, \quad (12)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N Y_t - F_t}, \quad (13)$$

where N is the number of observations, Y_t is the actual value and F_t is the predicted value.

For transformed data, the error measurement will calculate the smallest difference between the actual and predicted angle by using the formula in Eq. (14):

$$\Delta\theta_t = \min(|Y_t - F_t|, 360^\circ - |Y_t - F_t|). \quad (14)$$

This accounts for the circular nature of wind direction, where, for example, a prediction of 5° compared to an actual value of 355° should yield an error of 10° , not 350° . This adjusted error ensures accurate evaluation for directional data. The traditional error metrics, such as MAE and RMSE, can then be applied to $\Delta\theta_t$.

3. RESULTS AND DISCUSSION

This study uses daily wind direction data collected in Alor Setar from 2013 to 2017 to examine the effectiveness of trigonometric transformations in forecasting wind direction. The in-sample data will be selected from 1st January 2013 until 31st December 2016. Whilst the out sample will be from 1st January 2017 until 31st December 2017. As wind direction is a circular variable, with 0° and 360° representing the same direction, visual and statistical analysis of the raw data is crucial before applying any forecasting models. In Fig. 1, the wind direction is illustrated using a time series plot (a) and a wind rose diagram (b).

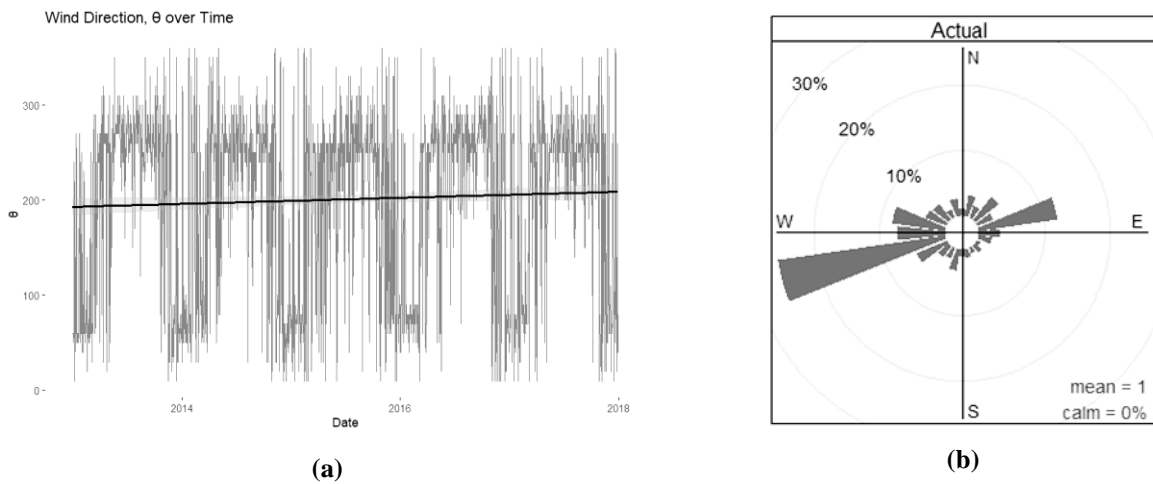


Figure 1. (a) Time Series Plot and (b) Wind Rose Diagram for Alor Setar Daily Wind Direction From 2013 to 2017

In Fig. 1 (a), we can see there is seasonality and a trend pattern contained in the data, but the wind rose diagram in Fig. 1 (b) could not capture or tell the seasonality pattern. To model and forecast using time series models, it is important to identify the characteristics of the data before analyzing the data to select a correct model for the modelling process.

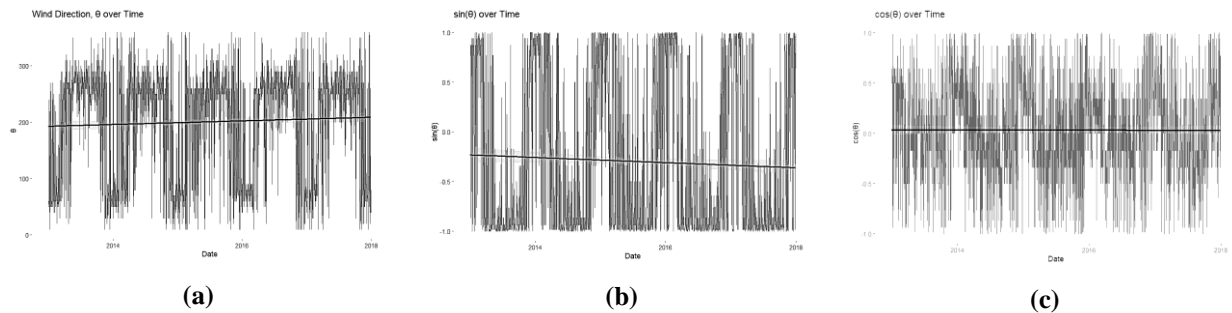


Figure 2. Time Series Plot for (a) Actual Wind Direction, (b) Wind Direction after Sine Transformation, and (c) Wind Direction after Cosine Transformation

Fig. 2 (a) displays the time series plots of the raw wind direction (θ), along with its sine and cosine transformations in Fig. 2 (b) and Fig. 2 (c). In Fig. 2 (a), a slight upward trend can be observed in the raw wind direction. Fig. 2 (b) shows a more noticeable downward trend in the $\sin(\theta)$ series, while Fig. 2 (c) indicates that the $\cos(\theta)$ values are relatively stable, without a clear trend pattern. Although long-term trends are not strongly pronounced across the three series, all exhibit noticeable seasonal behavior with a 365-day cycle, indicating the presence of annual patterns in wind direction.

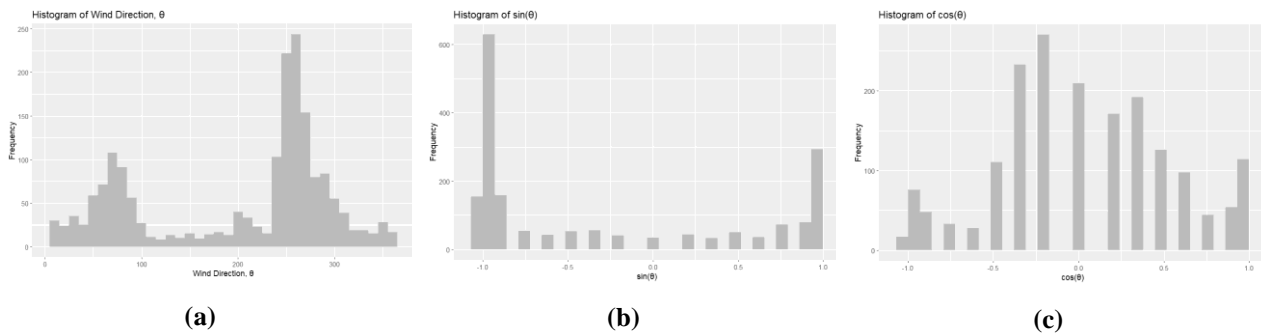


Figure 3. Histogram Plot for (a) Actual Wind Direction, (b) Sine Transformation, and (c) Cosine Transformation

Fig. 3 (a) represents histograms of the daily wind direction (θ), Fig. 3 (b) represents the transformation of $\sin(\theta)$, and Fig. 3 (c) represents the transformation of $\cos(\theta)$. In Fig. 3 (a), the distribution is evidently non-normal and multimodal, with two main peaks around 100° and 290° , indicating that wind tends to blow

predominantly from two main directions throughout the years. This pattern is consistent with seasonal shifts in monsoonal or prevailing winds affecting the region. In Fig.3 (b) and Fig. 3 (c), the histograms of $\sin(\theta)$ and $\cos(\theta)$ are bounded between -1 and 1, as expected, but still display skewness and non-symmetric shapes. For all three histograms confirming that the data does not follow a normal distribution.

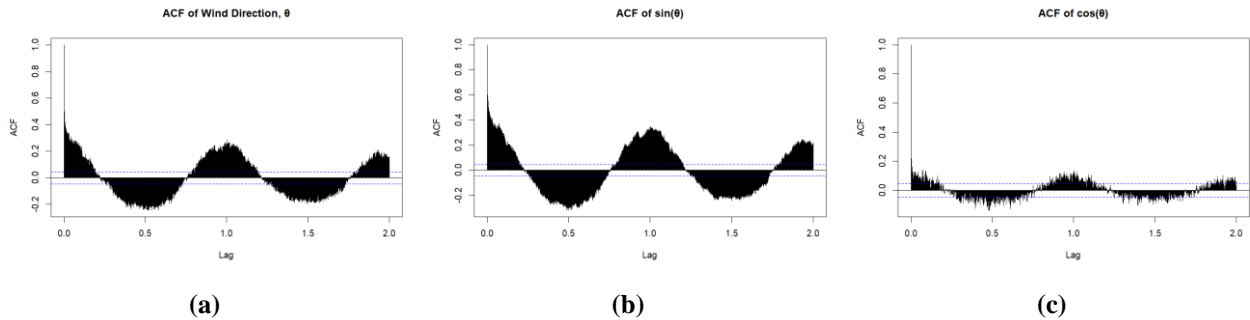


Figure 4. ACF Plot for (a) Actual Wind Direction, (b) Sine Transformation, and (c) Cosine Transformation

Fig. 4 (a) shows the autocorrelation function (ACF) plots of daily wind direction (θ), Fig. 4 (b) shows the ACF of $\sin(\theta)$ and Fig. 4 (c) shows the ACF of $\cos(\theta)$. The plots exhibit smooth, wave-like ACF patterns due to their inherent periodicity, with consistent oscillations between positive and negative correlations. These structured patterns confirm the presence of regular, predictable cycles in the data and support the use of seasonal components. The trend pattern in Fig. 2 does not show an extreme trend; therefore, to confirm whether the data contains a trend or not, the ADF test is used to help make the decision. The result of the ADF test is shown in Table 1 below.

Table 1. ADF Test for (a) Actual Wind Direction, (b) Sine Transformation and (c) Cosine Transformation

Variable	ADF Statistics	<i>p</i> -value	Conclusion
θ	-5.6433	Less than 0.01	Stationary
$\sin(\theta)$	-5.3155	Less than 0.01	Stationary
$\cos(\theta)$	-7.6361	Less than 0.01	Stationary

Based on the *p*-values obtained in Table 1, the null hypothesis is rejected for all three variables. This indicates that θ , $\sin \theta$ and $\cos \theta$ are stationary, with no significant trend observed in any of the series.

3.1 Forecasting Results: Direct vs Transformed Method

After the data is checked for trend and seasonality, three models are selected for modelling and forecasting, and tested for direct data and transformed data. For direct data, the data is trained to predict the wind direction angle (θ) directly. Three forecasting models were used for this purpose: Random Forest (RF), Holt-Winters Exponential Smoothing (HW), and Prophet. The transformed data are applied with the same three models, RF, HW, and Prophet models. After obtaining the forecasts for $\sin(\theta)$ and $\cos(\theta)$, the predicted wind direction angle was reconstructed using the $\tan^{-1} 2(\sin \theta, \cos \theta)$ function. The performance of both approaches is evaluated and compared in the following discussion to determine which method yields more accurate wind direction forecasts.

The in-sample forecast is compared using a time series plot, and the plots are obtained as in Fig. 5 below. The original data is represented by the black line, the Holt-Winters forecast is represented by the blue line, the Prophet forecast is represented by the green line, and the Random Forest forecast is represented by the red line.

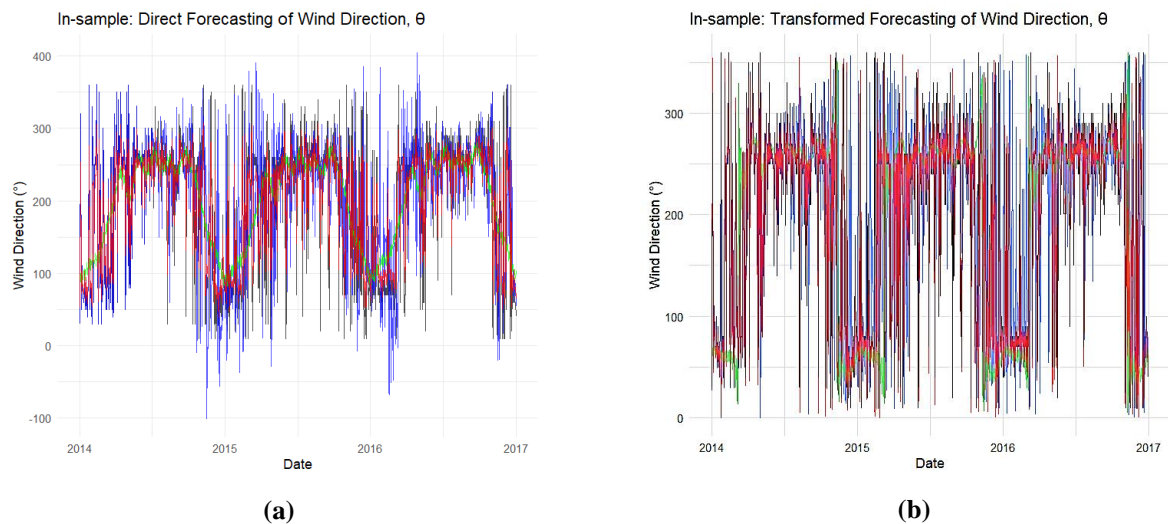
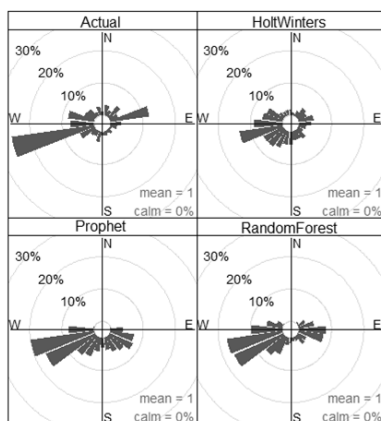


Figure 5. Time Series Plot for in-Sample Forecasting of (a) Direct Data and (b) Transformed Data

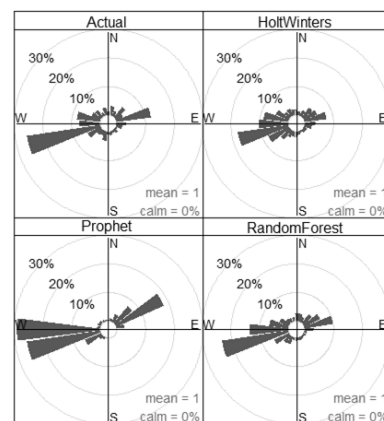
Fig. 5 shows the time series comparison between the actual wind direction and the direct and transformed forecasts from all models using the training data. All models appear to successfully capture the seasonal patterns of wind direction. The Holt-Winters and Random Forest models appear to have more fluctuation around the data compared to the Prophet model in Fig. 5 (a), but in Fig. 5 (b), the blue line seems to have more overfitting compared to the red and green lines. The advantage of the time series plot is that it can show the exact forecast according to the date. However, the time series plot does not give a clear idea of which model can follow the data accurately, since directional data is not linear data; it is important to use a wind rose diagram to see if the assumption is true.

In-sample : Wind Rose - All Models (Direct Method)



(a)

In-sample : Wind Rose - All Models (Transformed Method)



(b)

Figure 6. Wind Rose Diagram of in-Sample Forecasting for (a) Direct and (b) Transformed Data

Fig. 6 (a) and Fig. 6 (b) represent wind rose diagrams for in-sample predictions using the selected forecasting method for direct and transformed data. As the wind speed is fixed at 1 m/s, the diagrams purely illustrate the distribution of wind directions. The actual data reveal a multimodal distribution, primarily concentrated near the south–southwest (S–SW) and north–northeast (N–NE) directions. Both data show that the Random Forest model follows closely, resembling the actual directional distribution in direction and occurrence, while the Holt-Winters model shows more noticeable deviation, suggesting possible bias in its directional forecasts.

Comparing the results when using direct data and transformed data, the wind rose diagram for transformed data in Fig. 6 (b) looks more accurate compared to when modelled using the direct data. The forecast for all models generally can capture the angle and frequencies from the actual data, compared to Fig. 6 (a), the forecasts are more in the west-south-east (W-S-E) directions. To support this finding, the MAE and RMSE are calculated, and the results are shown in Table 2 below.

Table 2. Error Measurement of in-Sample Forecasting for Model Comparison between Direct and Transformed Data

Model	MAE		RMSE	
	Direct	Transformed	Direct	Transformed
Holt-Winters	51.6877	45.7851	72.2244	66.7747
Prophet	56.1345	45.0118	73.1086	66.3829
Random Forest	23.5302	6.5559	31.5213	9.0223

Table 2 displays the in-sample error measurements, MAE and RMSE, for the direct and transformed forecasting. Among the three models, Random Forest demonstrates the best performance, achieving the lowest MAE and RMSE, indicating higher accuracy and lower prediction error. In contrast, Holt-Winters and Prophet models exhibit higher error values, with Holt-Winters slightly outperforming Prophet. This suggests that the machine learning approach of Random Forest is more effective at capturing complex patterns in wind direction compared to the statistical models. However, since these results are solely based on in-sample performance, they do not reflect the models' ability to generate new data. Therefore, out-sample evaluation is necessary to draw more reliable conclusions about forecasting performance.

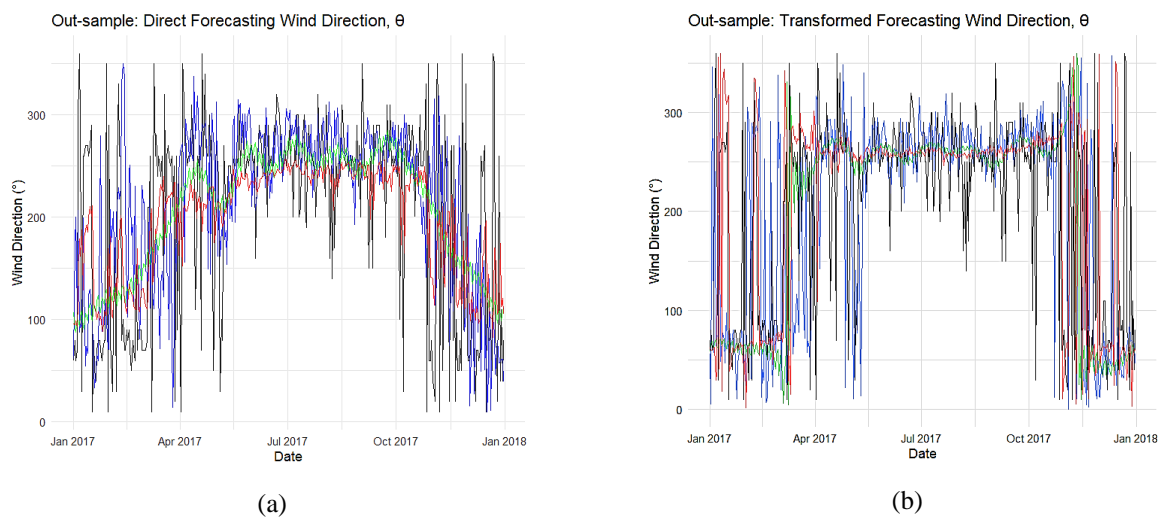


Figure 7. Time Series Plot for Out-Sample Forecasting of (a) Direct Data and (b) Transformed Data

Fig. 7 represents the time series plot of comparison between the out-sample forecast using the direct wind direction data and the transformed data. The Holt-Winters method appears to consistently overestimate the wind direction, while the Prophet method tends to underestimate it. Among the models, Random Forest demonstrates the most accurate performance in capturing the wind direction pattern. However, it still occasionally underestimates the values, indicating some room for improvement in generalization.

Out-sample : Wind Rose - All Models (Direct Method)

Out-sample : Wind Rose - All Models (Transformed Method)

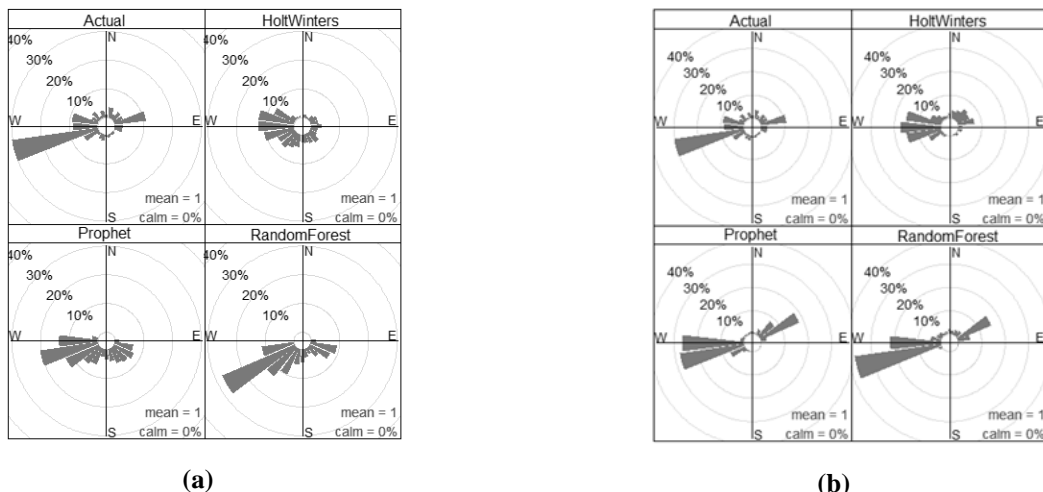


Figure 8. Wind Rose Diagram of Out Sample Forecasting for (a) Direct and (b) Transformed Data

Fig. 8 displays wind rose diagrams representing out-sample predictions from all models using direct and transformed data. With wind speed fixed at 1 m/s, these diagrams focus solely on illustrating the distribution of wind directions. The actual testing data in Fig. 8 (a) show a directional pattern like the training set, primarily concentrated around the S–SW and N–NE directions. Among the models, Random Forest best replicates this distribution, closely matching the observed patterns. Conversely, the Holt-Winters model shows greater divergence. Compared to the direct method, the transformed data forecasts in Fig. 8 (b) better capture the directional patterns for all models. Each model shows improved alignment with the actual training wind direction. Among them, Random Forest demonstrates the most accurate representation, closely matching the actual wind direction distribution.

Table 3. Error Measurement of Out-Sample Forecasting for Model Comparison between Direct and Transformed Data

Model	MAE		RMSE	
	Direct	Transformed	Direct	Transformed
Holt-Winters	62.9878	55.1808	81.3114	74.7323
Prophet	58.0489	43.1535	76.1923	64.1957
Random Forest	56.2866	39.1696	72.9327	57.3008

Table 3 represents the error measurements for the out-sample performance of the direct and transformed forecasting methods. Compared to the in-sample results, the errors for all models have increased, indicating reduced accuracy on unseen data. While Prophet performed the worst in-sample, its out-sample performance improved slightly and outperformed Holt-Winters. Holt-Winters now shows the highest MAE and RMSE, which means that the model struggles to capture the underlying wind direction patterns. Although Random Forest still achieves the best performance overall, the gap between it and the other models has narrowed in the out-sample results.

The superior performance of Random Forest arises from its capacity to capture nonlinear relationships and interactions within the data, which are especially crucial for complex meteorological variables like wind direction. Conversely, Holt–Winters and Prophet, though effective for linear and seasonal data, lack the flexibility to model the nonlinear dependencies inherent in directional data. From a practical standpoint, the enhanced accuracy of transformation-based Random Forest forecasts is highly relevant for wind turbine alignment, aviation runway safety, and pollution monitoring, where small angular improvements can lead to significant operational advantages.

Across the three models examined, Random Forest consistently emerged as the most effective approach, achieving the lowest MAE and RMSE values. This finding suggests that Random Forest not only captures complex, nonlinear relationships within the training data but also exhibits strong generalization capacity when applied to out-of-sample forecasts. The Prophet model delivered a moderate level of accuracy, performing better than Holt–Winters but failing to match the robustness of Random Forest, likely due to its relative limitations in modeling nonlinear and high-dimensional dependencies. More importantly, the consistent superiority of the transformed data over the direct data highlights the critical role of appropriately accounting for the circular nature of wind direction. By leveraging sine and cosine transformations, the models were able to reduce angular error and improve predictive reliability, underscoring the methodological significance of transformation strategies in forecasting circular variables.

To further validate these findings, future research should utilize statistical significance tests like the Diebold-Mariano test to formally evaluate whether differences in forecast accuracy between models are significant. Other limitations of this study are the limitations of access to wind direction data, limitations in visualizing the comparison of the directional forecasts, and applications of time series modelling with directional data.

4. CONCLUSION

This study provides clear evidence on the effectiveness of trigonometric transformation in forecasting wind direction. The main conclusions are as follows:

- 1 Trigonometric transformation of wind direction significantly improves forecasting accuracy compared to direct modelling, effectively addressing the challenges of circular data.

- 2 Among the models tested, Random Forest with transformation achieved the lowest error values, outperforming Holt-Winters and Prophet in both in-sample and out-sample evaluations.
- 3 The study is among the first in Malaysia to apply transformation strategies for wind direction forecasting, offering methodological novelty and practical insights.
- 4 The findings have important applications in renewable energy optimization, aviation safety, and environmental monitoring, where accurate wind direction forecasts are critical.
- 5 Future research should explore hybrid or ensemble approaches, test across multiple locations and longer time periods, and integrate wind speed to build more comprehensive forecasting models.

Author Contributions

Nur Arina Bazilah Kamisan: Conceptualization, Data Curation, Methodology, Resources, Validation, Writing-Review and Editing. Pow Jing Huei: Data Curation, Draft Preparation, Formal Analysis. Muhammad Hisyam Lee: Validation, Writing-Review and Editing. All authors have read and approved the final manuscript.

Funding Statement

This work is supported by the Universiti Teknologi Malaysia under Grant Q.J130000.2754.04K49.

Acknowledgment

We would like to thank Jabatan Meteorologi Malaysia for the data used in this study.

Declarations

The authors declare no competing interests.

Declaration of Generative AI and AI-Assisted Technologies

Generative AI tools (e.g., ChatGPT) were used solely for language refinement, including grammar, spelling, and clarity. The scientific content, analysis, interpretation, and conclusions were developed entirely by the authors. All final text was reviewed and approved by the authors.

REFERENCES

- [1] D. Alves, F. Mendonça, S. S. Mostafa, and F. Morgado-Dias, "THE POTENTIAL OF MACHINE LEARNING FOR WIND SPEED AND DIRECTION SHORT-TERM FORECASTING: A SYSTEMATIC REVIEW," *Computers*, vol. 12, no. 10, p. 206, Oct. 2023. doi: <https://doi.org/10.3390/computers12100206>
- [2] F. M. Noman, G. A. Alkaws, D. Abbas, A. A. Alkahtani, S. K. Tiong, and J. Ekanayake, "COMPREHENSIVE REVIEW OF WIND ENERGY IN MALAYSIA: PAST, PRESENT, AND FUTURE RESEARCH TRENDS," *IEEE Access*, vol. 8, pp. 124526–124543, 2020. doi: <https://doi.org/10.1109/ACCESS.2020.3006134>
- [3] A. Albani, M. Z. Ibrahim, S. S. Abdul Ghani, M. Z. Mat Rofi, and P. N. A. Taslin, "THE IMPACT STUDY OF EL NIÑO-SOUTHERN OSCILLATION TO THE WIND AND SOLAR DATA IN MALAYSIA USING THE WAVELET ANALYSIS," *Front Energy Res*, vol. 8, Jan. 2021. doi: <https://doi.org/10.3389/fenrg.2020.591469>
- [4] E. Yatiyana, S. Rajakaruna, and A. Ghosh, "WIND SPEED AND DIRECTION FORECASTING FOR WIND POWER GENERATION USING ARIMA MODEL," *2017 Australasian Universities Power Engineering Conference, AUPEC 2017*, vol. 2017-Novem, pp. 1–6, 2018. doi: <https://doi.org/10.1109/AUPEC.2017.8282494>
- [5] M. Santhosh, C. Venkaiah, and D. M. V. Kumar, "CURRENT ADVANCES AND APPROACHES IN WIND SPEED AND WIND POWER FORECASTING FOR IMPROVED RENEWABLE ENERGY INTEGRATION: A REVIEW," Jun. 01, 2020, *John Wiley and Sons Inc.* doi: <https://doi.org/10.1002/eng2.12178>
- [6] Wynnyk, Christopher M. "WIND ANALYSIS IN AVIATION APPLICATIONS." *2012 IEEE/AIAA 31st Digital Avionics Systems Conference (DASC)*. IEEE, 2012. <https://doi.org/10.1109/DASC.2012.6383045>
- [7] G. Zhang, J. Li, T. Chang, W. Zhang, and L. Song, "AUTONOMOUS NAVIGATION AND CONTROL FOR A SUSTAINABLE VESSEL: A WIND-ASSISTED STRATEGY," *Sustainable Horizons*, vol. 13, p. 100117, Mar. 2025. doi: <https://doi.org/10.1016/j.horiz.2024.100117>
- [8] Q. Xu, *et al.*, "A WIND VECTOR DETECTING SYSTEM BASED ON TRIBOELECTRIC AND PHOTOELECTRIC SENSORS FOR SIMULTANEOUSLY MONITORING WIND SPEED AND DIRECTION," *Nano Energy*, vol. 89, p. 106382, Nov. 2021. doi: <https://doi.org/10.1016/j.nanoen.2021.106382>

- [9] Á. López-Oriona, Y. Sun, and R. M. Crujeiras, "FUZZY CLUSTERING OF CIRCULAR TIME SERIES WITH APPLICATIONS TO WIND DATA," *Environmetrics*, vol. 36, no. 2, Mar. 2025. doi: <https://doi.org/10.1002/env.2902>
- [10] A. Azhar and H. Hashim, "A REVIEW OF WIND CLUSTERING METHODS BASED ON THE WIND SPEED AND TREND IN MALAYSIA," *Energies (Basel)*, vol. 16, no. 8, p. 3388, Apr. 2023. doi: <https://doi.org/10.3390/en16083388>
- [11] M. S. Hanoon, et al., "WIND SPEED PREDICTION OVER MALAYSIA USING VARIOUS MACHINE LEARNING MODELS: POTENTIAL RENEWABLE ENERGY SOURCE," *Engineering Applications of Computational Fluid Mechanics*, vol. 16, no. 1, pp. 1673–1689, Dec. 2022. doi: <https://doi.org/10.1080/19942060.2022.2103588>
- [12] Y. Han, et al., "A SHORT-TERM WIND SPEED PREDICTION METHOD UTILIZING NOVEL HYBRID DEEP LEARNING ALGORITHMS TO CORRECT NUMERICAL WEATHER FORECASTING," *Appl Energy*, vol. 312, p. 118777, Apr. 2022. doi: <https://doi.org/10.1016/j.apenergy.2022.118777>
- [13] F. Sun and T. Jin, "A HYBRID APPROACH TO MULTI-STEP, SHORT-TERM WIND SPEED FORECASTING USING CORRELATED FEATURES," *Renew Energy*, vol. 186, pp. 742–754, Mar. 2022. doi: <https://doi.org/10.1016/j.renene.2022.01.041>
- [14] M. Yaghoubirad, N. Azizi, M. Farajollahi, and A. Ahmadi, "DEEP LEARNING-BASED MULTISTEP AHEAD WIND SPEED AND POWER GENERATION FORECASTING USING DIRECT METHOD," *Energy Convers Manag*, vol. 281, p. 116760, Apr. 2023. doi: <https://doi.org/10.1016/j.enconman.2023.116760>
- [15] X. Jing-Jing, H. Fei, X. Zi-Niu, and C. Xue-Ling, "BIAS CORRECTION IN WIND DIRECTION FORECASTING USING THE CIRCULAR-CIRCULAR REGRESSION METHOD," *Atmospheric and Oceanic Science Letters*, vol. 7, no. 2, pp. 87–91, Jan. 2014. doi: <https://doi.org/10.1080/16742834.2014.11447140>
- [16] E. Erdem and J. Shi, "ARMA BASED APPROACHES FOR FORECASTING THE TUPLE OF WIND SPEED AND DIRECTION," *Appl Energy*, vol. 88, no. 4, pp. 1405–1414, 2011. doi: <https://doi.org/10.1016/j.apenergy.2010.10.031>
- [17] N. A. Mokhtar, B. Badyalina, K. L. Chang, F. F. Ya'acob, A. F. Ghazali, and P. Shamala, "ERROR-IN-VARIABLES MODEL OF MALACCA WIND DIRECTION DATA WITH THE VON MISES DISTRIBUTION IN SOUTHWEST MONSOON," *Applied Mathematical Sciences*, vol. 15, no. 9, pp. 471–479, 2021. doi: <https://doi.org/10.12988/ams.2021.914521>
- [18] N. D. A. Halim, K. N. A. Maulud, K. C. H. Lun, W. S. W. M. Jaafar, F. A. Mohd, and F. Khan, "AIR POLLUTION PREDICTION BASED ON CHANGES IN MONSOON WIND DIRECTION BY USING TRAJECTORY-GEOSPATIAL APPROACH," *Jurnal Kejuruteraan*, vol. 35, no. 2, pp. 303–316, Mar. 2023. doi: [https://doi.org/10.17576/jkukm-2023-35\(2\)-04](https://doi.org/10.17576/jkukm-2023-35(2)-04)
- [19] N. H. Moslim, N. A. Mokhtar, Y. Z. Zubairi, and A. G. Hussin, "UNDERSTANDING THE BEHAVIOUR OF WIND DIRECTION IN MALAYSIA DURING MONSOON SEASONS USING REPLICATED FUNCTIONAL RELATIONSHIP IN VON MISES DISTRIBUTION," *Sains Malays*, vol. 50, no. 7, pp. 2035–2045, Jul. 2021. doi: <https://doi.org/10.17576/jsm-2021-5007-18>
- [20] N. H. Moslim, Y. Z. Zubairi, A. G. Hussin, S. F. Hassan, and N. A. Mokhtar, "UNDERSTANDING THE PATTERN OF WIND DIRECTION IN PENINSULAR MALAYSIA," 2023, p. 020007. doi: <https://doi.org/10.1063/5.0109944>
- [21] Blesgen, Thomas. "ON ROTATION DEFORMATION ZONES FOR FINITE-STRAIN COSSERAT PLASTICITY." *Acta Mechanica* 226.7 (2015): 2421-2434. <https://doi.org/10.1007/s00707-015-1326-2>
- [22] A. Ukil, V. H. Shah, and B. Deck, "FAST COMPUTATION OF ARCTANGENT FUNCTIONS FOR EMBEDDED APPLICATIONS: A COMPARATIVE ANALYSIS," in *Proceedings - ISIE 2011: 2011 IEEE International Symposium on Industrial Electronics*, 2011, pp. 1206–1211. doi: <https://doi.org/10.1109/ISIE.2011.5984330>
- [23] L. Breiman, "RANDOM FORESTS," 2001. <https://doi.org/10.1023/A:1010933404324>
- [24] M. J. Kane, N. Price, M. Scotch, and P. Rabinowitz, "COMPARISON OF ARIMA AND RANDOM FOREST TIME SERIES MODELS FOR PREDICTION OF AVIAN INFLUENZA H5N1 OUTBREAKS," *BMC Bioinformatics*, vol. 15, no. 1, p. 276, Dec. 2014. doi: <https://doi.org/10.1186/1471-2105-15-276>
- [25] R. P. Masini, M. C. Medeiros, and E. F. Mendes, "MACHINE LEARNING ADVANCES FOR TIME SERIES FORECASTING," Apr. 2021, [Online]. Available: <http://arxiv.org/abs/2012.12802>
- [26] Kalekar, Prajakta S. "TIME SERIES FORECASTING USING HOLT-WINTERS EXPONENTIAL SMOOTHING." *Kanwal Rekhi school of information Technology* 4329008.13 (2004): 1-13.
- [27] N. Nurhamidah, N. Nusyirwan, and A. Faisol, "FORECASTING SEASONAL TIME SERIES DATA USING THE HOLT-WINTERS EXPONENTIAL SMOOTHING METHOD OF ADDITIVE MODELS," *Jurnal Matematika Integratif*, vol. 16, no. 2, p. 151, Dec. 2020. doi: <https://doi.org/10.24198/jmi.v16.n2.29293.151-157>
- [28] T. O. Hodson, "ROOT-MEAN-SQUARE ERROR (RMSE) OR MEAN ABSOLUTE ERROR (MAE): WHEN TO USE THEM OR NOT," Jul. 19, 2022, *Copernicus GmbH*. doi: <https://doi.org/10.5194/gmd-15-5481-2022>
- [29] T. Chai and R. R. Draxler, "ROOT MEAN SQUARE ERROR (RMSE) OR MEAN ABSOLUTE ERROR (MAE)? - ARGUMENTS AGAINST AVOIDING RMSE IN THE LITERATURE," *Geosci Model Dev*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014. doi: <https://doi.org/10.5194/gmd-7-1247-2014>