RESEARCH ARTICLE • OPEN ACCESS

Comparative Analysis of ARIMA and LSTM for Forecasting Maximum Wind Speed in Kupang City, East Nusa Tenggara

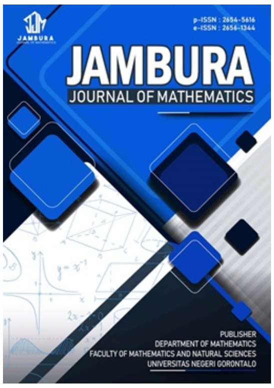
Indah Magfirrah *et al.*

Volume 6, Issue 2, Pages 169–175, August 2024

Received 4 June 2024, Revised 4 July 2024, Accepted 7 July 2024, Published 1 August 2024

To Cite this Article : I. Magfirrah, *et al.*,"Comparative Analysis of ARIMA and LSTM for Forecasting Maximum Wind Speed in Kupang City, East Nusa Tenggara", *Jambura J. Math*, vol. 6, no. 2, pp. 169–175, 2024, *https://doi.org/10.37905/jjom.v6i2.25834*

JOURNAL INFO • JAMBURA JOURNAL OF MATHEMATICS



	Homepage	:	http://ejurnal.ung.ac.id/index.php/jjom/index
	Journal Abbreviation	:	Jambura J. Math.
	Frequency	:	Biannual (February and August)
	Publication Language	:	English (preferable), Indonesia
	DOI	:	https://doi.org/10.37905/jjom
	Online ISSN	:	2656-1344
	Editor-in-Chief	:	Hasan S. Panigoro
	Publisher	:	Department of Mathematics, Universitas Negeri Gorontalo
	Country	:	Indonesia
	OAI Address	:	http://ejurnal.ung.ac.id/index.php/jjom/oai
	Google Scholar ID	:	iWLjgaUAAAAJ
	Email	:	info.jjom@ung.ac.id

Jambura Journal of Mathematics, Volume 6, Issue 2, Pages 169–175, August 2024 https://doi.org/10.37905/jjom.v6i2.25834

Research Article

Check for updates

Comparative Analysis of ARIMA and LSTM for Forecasting Maximum Wind Speed in Kupang City, East Nusa Tenggara

Indah Magfirrah[1,*], Meisyatul Ilma[1], Khairil Anwar Notodiputro[1], Yenni Angraini[1], and Laily Nissa Atul Mualifah[1]

[1] *Department of Statistics, IPB University, Bogor, Indonesia*

ARTICLE HISTORY
Received 4 June 2024
Revised 4 July 2024
Accepted 7 July 2024
Published 1 August 2024

KEYWORDS
Wind Speed
ARIMA
LSTM
Grid Search

ABSTRACT. *This study compares the Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) models for predicting maximum wind speed based on accuracy measured by Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). Based on the results of the research, the LSTM model is better than the ARIMA model in predicting maximum wind speed in Kupang City, East Nusa Tenggara Province. The best LSTM model has hyperparameters of 200 epochs; batch size of 32; learning rate of 0,001; and 8 neurons. Based on the evaluation results of predicted data against actual data, the MAPE value of the LSTM model is 19,40%. The benefit of this research is that it can contribute to the literature on the development of wind utilization as a basis for building power plants on small islands as a renewable resource, particularly in Kupang City, East Nusa Tenggara.*

1. Introduction

Forecasting is a method of estimating future events by examining information from the past [1]. The Autoregressive Integrated Moving Average (ARIMA) model is a classic method often used to forecast time series data. The ARIMA model is an ARMA model for non-stationary time series data. Time series data can be stationary by differencing, where the differencing value (d) is a non-negative integer. ARIMA models are usually better able to describe and anticipate geographic seasonal patterns in stable historical data but often face difficulties in forecasting unexpected changes, so deep learning can be used to overcome these weaknesses.

Deep learning uses end-to-end concepts [2], where learning layers in deep learning make accuracy and performance better than other algorithms [3]. One of the deep learning methods commonly used for forecasting time series data is Long Short-Term Memory (LSTM). Its advantage over the ARIMA model is that LSTM is superior and reliable in forecasting long-term periods [4]. LSTM has hyperparameters that can determine the reliability and performance of the model performance [5]. However, even though a method is said to be good, evaluation is still important to see the accuracy, such as with the use of time series cross-validation, which will divide the data into training data and testing data by paying attention to the time sequence to get the best model to be used for forecasting [6].

Research using the ARIMA and LSTM methods has been widely used, especially to forecast future events. Putri and Sadikin (2021), compared ARIMA and LSTM in predicting product sales to estimate raw material requirements. The research results prove that LSTM is the best method compared to ARIMA for this case [7]. Elsaraiti and Merabet [8] also compared the two methods in predicting wind speed. The results showed that LSTM was better in that case. Milniadi and Adiwijaya [9] compared the performance of LSTM and ARIMA in forecasting stock closing prices. The results showed that ARIMA was better than LSTM for that case. Umam and Ardiansyah [10] also compared the two methods to predict the number of library visitors. The results prove that ARIMA is better for this case.

Accurate forecasting results are needed for decision-making and future planning, including wind speed forecasting. Wind is a movement of air caused by the rotation of the earth and the difference in air pressure around it. Wind intensity is determined by speed, which fluctuates and changes over time [11]. Accurate forecasting methods can help manage and reduce the impact of weather phenomena and can be used as a reference in planning wind energy utilization. According to the National Research and Innovation Agency (BRIN), East Nusa Tenggara is one of Indonesia's provinces with the most significant wind power potential.

Therefore, this study will compare ARIMA, considered a benchmark in time series data analysis, and LSTM, which has shown great promise in handling complex patterns in sequential data. ARIMA is essential for its robustness and simplicity in forecasting stability, making it a reliable baseline method. On the other hand, LSTM's ability to capture long-term dependencies and nonlinear relationships in the data makes it a powerful tool for more accurate and dynamic forecasting.

By comparing these two methods, this research aims to determine which model performs better in forecasting wind speed, especially in Kupang City, East Nusa Tenggara Province. The

*Corresponding Author.

Email : *indahmagfirrah@apps.ipb.ac.id* (I. Magfirrah)
Homepage : http://ejurnal.ung.ac.id/index.php/jjom/index / E-ISSN : 2656-1344

findings could offer significant implications for improving wind energy planning and management, and reducing the impact of weather-related uncertainties. This comparative analysis could also contribute to the broader field of time series forecasting by highlighting the strengths and limitations of traditional and modern approaches.

2. Methods

2.1. Research Procedures

The research flow is exploring and preprocessing data, splitting data, modeling ARIMA and LSTM using training data, and comparing the best model between ARIMA and LSTM using training data and testing data so that the best model based on the smallest MAPE and RMSE values will be used to predict/forecast [12]. This research was conducted following the steps of the research flowchart in Figure 1.

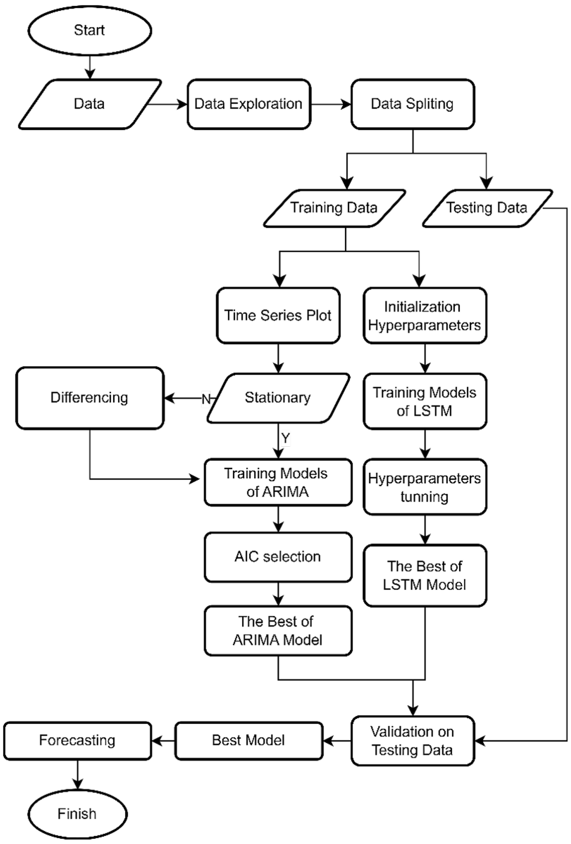


Figure 1. Research procedures

2.2. Data Source

This research uses secondary data obtained from https://power.larc.nasa.gov/data-access-viewer, namely Maximum Daily Wind Speed Data in Kupang City, East Nusa Tenggara, for the period January 2019-April 2024.

2.3. Data Splitting

Data splitting is done to find out how the model can work well. The data is divided into two parts, namely training and testing data. This study has training data for January 2019 - December 2023 and testing data for January 2024-April 2024. The division of training and testing data is presented in Table 1.

Table 1. Proportion of data splitting

Training Data	1.826
Testing Data	121
Total	1.947

2.4. ARIMA Model

The ARIMA model also called the Box-Jenkins model, is a combination of the Autoregressive (AR) model with order p, Moving Average with order q, and followed by the order d differencing process. The equation of the ARIMA model is expressed in Equation (1) [13].

$$\varphi_p(B)(1-B)^d y_t = \theta_q(B)\, e_t \tag{1}$$

where

y_t	:	observation value at time ke-*t*,
φ_p	:	coefficient at order p,
θ_q	:	MA coefficient at order q,
$(1-B)^d$	:	d-order difference series,
B	:	Backshift operator,
e_t	:	residuals at time-t.

The first step is to identify and select the appropriate model. If the data is not stationary in the mean, differencing can be done once or twice. However, logarithmic transformation can be performed if the data is not stationary in variance [14]. After the data is stationary, plots of the autocorrelation function (ACF) and partial autocorrelation function (PACF) are used to select the appropriate order of the AR and MA models. The second step is to estimate the model parameters using metrics such as the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) [15] to optimize the model. Finally, a diagnostic check of the model is performed by analyzing the residuals to determine the best final model [16].

2.5. LSTM Model

Long-short-term memory (LSTM) is one of the modifications of the Recurrent Neural Network (RNN) that was first introduced by Hochreiter and Schmidhuber in 1997 [17]. LSTM is present to overcome the weakness of RNN, which cannot predict data based on information stored over a long period. Specifically, the gate mechanism in LSTM is composed of three vector gate units: input gate, forget gate, and output gate [18]. Each vector gate unit has a different function. The input vector gate controls the number of input vectors affecting the memory. The forget vector gate controls the amount of old memory that will be erased. The output vector gate controls the amount of memory stored in the hidden state.

The first process in LSTM is to filter out the information that needs to be deleted. The forget gate determines the information that will be removed from the cell state. The variable equations of LSTM are x_t as the input vector to LSTM, F_t as the forget gate activation vector, I_t as the input gate activation vector, O_t as the output gate activation vector, H_t as the hidden state vector is known as the output vector of the LSTM unit, C_t as the cell state vector, W and U as the weight matrix, and b bias vector parameters that must be learned during training. The general equation form of LSTM can be seen in equations (2)-(6) [19].

$$F_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right), \tag{2}$$

$$I_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right), \quad (3)$$
$$O_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right), \quad (4)$$
$$C_t = F_t * C_{t-1} + I_t * \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right), \quad (5)$$
$$H_t = O_t * \tanh\left(C_t\right). \quad (6)$$

LSTM has two activation functions, each with its respective role. The sigmoid activation function has the role of converting an x value into a value that has a range of 0 to 1. The tanh activation function has the role of converting an x value into a value that has a range of -1 to 1. The sigmoid and tanh activation function formulas can be seen in equations (7) and (8) [20].

$$\sigma = \frac{1}{1+e^{-x}}, \quad (7)$$
$$\tanh = 2\sigma\left(2x\right) - 1, \quad (8)$$

where x is the input data and σ is the value of the sigmoid activation function.

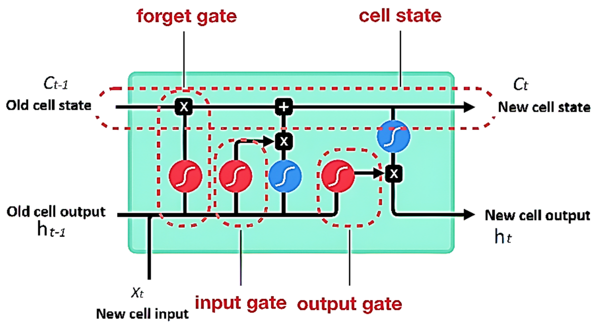


Figure 2. LSTM architecture

2.6. Hyperparameters Tuning

Hyperparameter tuning is the process of finding the optimal hyperparameter value combination that helps improve the performance of the LSTM model so that it provides the best performance [21]. One of the hyperparameter tuning methods is grid search. The main advantage of grid search is its ability always to try all possible hyperparameter combinations to find the best one to optimize the model based on accuracy and precision values [22].

3. Results and Discussion

3.1. Data Exploration

Descriptive analysis is used to see the maximum wind speed pattern in Kupang, East Nusa Tenggara, from 2019 to 2024 using daily data from January to December. The time series plot of the daily data is shown in Figure 3.

The minimum data of maximum wind speed in Kupang, East Nusa Tenggara is 2,96 knot, with an average maximum wind speed of 10,324 knot. The maximum data of maximum wind speed during 2019-2024 was in early April 2021, reaching 37,24 knot. In addition, generally, wind speed decreases from July to November and fluctuates from December to June, as seen in the time series plot.

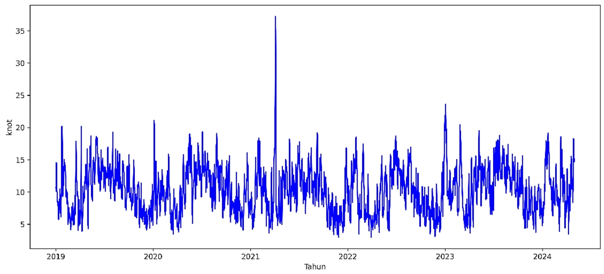


Figure 3. Time series plot

3.2. ARIMA Model

Before building an ARIMA model, the first step must be to test the stationarity of the data. Time series data can be said to be stationary if the data fluctuates around a relatively constant center value and variety across periods. Data stationarity can be checked with the ADF test (Augmented Dickey-Fuller) and KPSS (Kwiatkowski-Phillips-Schmidt-Shin) tests.

Table 2. ADF and KPSS tests on training data

Test	p-value	Description
ADF	0,0000	Stationary
KPSS	0,1	Stationary

Based on the ADF test results it shows that the p-value is less than α (0,05), meaning that the data is stationary, the KPSS test results also evidence this, the p-value is more than α (0,05), meaning that the data is stationary.

After proving the data is stationary, identify tentative models that will be built based on the ACF (Autocorrelation Function) plot and the PACF (Partial Autocorrelation Function) plot. The tentative model proposal is based on significant lags in the ACF and PACF plots in Figure 4.

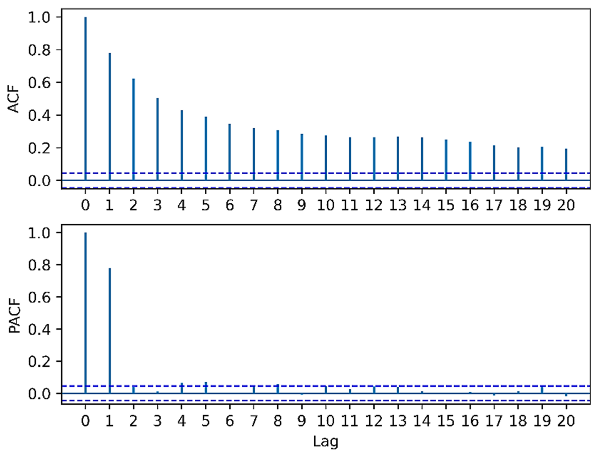


Figure 4. ACF and PACF plots on training data

Based on the ACF plot on the training data, it can be seen that the plot decreases slowly, meaning that the data does not have an MA (Moving Average) component. Then the PACF plot cuts off, meaning that the data contains an AR (Autoregressive) component so that the tentative models proposed are ARIMA (4,0,0), ARIMA (2,0,0), ARIMA (1,0,0) and ARIMA (2,0,1). The parameter estimation results of the four candidate models are presented in Table 3.

Table 3. Estimated values of ARIMA model parameters

Model	Parameter	coefficient	p-value	AIC
ARIMA (4,0,0)	AR(1)*	0,7475	0,000	8271,162
	AR(2)	0,0291	0,212	
	AR(3)	-0,0372	0,159	
	AR(4)*	0,0651	0,006	
ARIMA (2,0,0)	AR(1)*	0,7487	0,000	8275,154
	AR(2)*	0,0397	0,029	
ARIMA (1,0,0)	AR(1)*	0,7797	0,000	8276,039
ARIMA (2,0,1)	AR(1)*	1,6480	0,000	**8252,211**
	AR(2)*	-0,6575	0,000	
	MA(1)*	-0,9211	0,000	

*Parameters are significant at the 5% significance level

Based on Table 3, the ARIMA (2,0,1) model has significant parameter values at the 5% significance level with smaller AIC values. This indicates that the ARIMA (2,0,1) model is the best tentative model of the four tentative model candidates. After obtaining the best model, model diagnostic testing will be carried out to check the feasibility of the model. Examination of the randomness of the residuals will be carried out using the Ljung-Box test. The Ljung-Box test results for lags 5; 10; 15; 20 and 25 of the model are presented in Table 4.

Table 4. Ljung-Box test of the residuals

Model	*p-value*	Lag
ARIMA(2,0,1)	0,185953	5
	0,307689	10
	0,475262	15
	0,512566	20
	0,413771	25

Table 4 shows that the ARIMA (2,0,1) model has fulfilled the assumption of randomness of the samples because all lags have a p-value more than the 5% significance level. Next, we check the normality of the samples. Examination of the normality of the remainder will be carried out using the Shapiro test. The results of the Shapiro test show that the residuals do not spread generally because they have a p-value less than the 5% significance level. However, the Shapiro test results can be tolerated because the data used tends to be significant [23]. In this study, 1.947 data were used, so it can be said that the assumption of normal distribution has been fulfilled.

After obtaining the tentative ARIMA model with the smallest AIC value, overfitting will be carried out, which is presented in Table 5.

Table 5. Estimated values of Overfitting model parameters

Model	Parameter	coefficient	p-value	AIC
ARIMA (3,0,1)	AR(1)	0,4671	0,988	8279,153
	AR(2)	0,2512	0,991	
	AR(3)	0,0105	0,993	
	MA(1)	0,2817	0,992	
ARIMA (2,0,2)	AR(1)	0,2965	0,990	8279,045
	AR(2)	0,4031	0,983	
	MA(1)	0,4509	0,985	
	MA(2)	-0,0258	0,984	

*Parameters that are significant at the 5% significance level

Based on Table 5, it can be seen that the parameter estimation results for both models show insignificant results at the 5% significance level. The ARIMA (2,0,1) model is the best model for the maximum wind speed data in Kupang, East Nusa Tenggara, because the parameters are significant and the AIC is the smallest. The ARIMA (2,0,1) formula is written as follows in Equation (9):

$$Y_t = 1.6480Y_{t-1} - 0.6575Y_{t-2} - 0.9211\varepsilon_{(Y)t-1} + \varepsilon_{(Y)t}. \quad (9)$$

3.3. LSTM Model

Modeling with LSTM does not require the assumption of data stationarity as in the ARIMA model. Because, before LSTM modeling the data is normalized using MinMaxScaler into the range (0,1). Data normalization helps improve model performance by adjusting the weights in the model [24]. The plot of maximum wind speed data after normalization is shown in Figure 5.

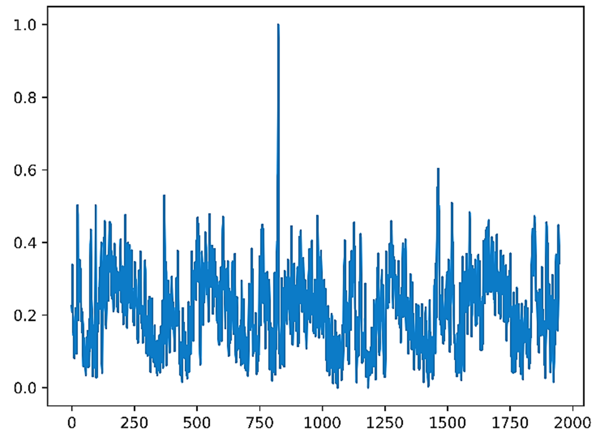


Figure 5. Data plot after normalization

Furthermore, the data is converted to supervised learning or labeled data. The data at time t is labeled as the prediction data $Y = x_t$, while the data at the previous time $(x_t - 1)$ is the input [25]. The supervised learning form helps the model learn the data pattern and the relationship between input and output to predict the following data set and avoid poor out-of-the-box performance. The data is then converted to a three-dimensional array by LSTM modeling and deep learning in general, where the typical input data is a 3-dimensional array, namely number of samples, time steps, number of features [26]. In this study, the training and testing data were transformed using the numpy.reshape() function to have the shape of (1824, 1, 1) and (119, 1, 1).

The LSTM model architecture used in this study consists of two LSTM layers and one Dense layer. The optimizer and loss function used to help model convergence are ADAM (Adaptive Moment Estimation) and MSE (Mean Squared Error). ADAM optimizer with MSE loss function is a commonly used algorithm in LSTM modeling. The ADAM optimizer combines the advantages of RMSProp, which uses the square of the gradient to adjust the learning rate, and the advantages of SGD (Stochastic Gradient Descent) with momentum, which utilizes momentum by using an exponentially weighted moving average gradient rather than the gradient itself so that ADAM works well as an optimizer [27]. LSTM and other deep learning methods use hyperparameters not directly learned by the algorithm to find the best value for the model. The hyperparameters used include neurons, batch size, epoch, and learning rate.

Neurons serve to remember the information needed to make predictions. Choosing neurons that are too small can make the LSTM unable to perform its function correctly, and if the neurons are too large, the LSTM will experience overfitting. It is written that epochs represent the total number of times a given data set is used for training purposes [28]. Epochs also serve to capture patterns in the data. One epoch indicates that the entire dataset is fed to the model only once. Choosing an epoch that is too small can cause the model not to capture the pattern correctly, while selecting an epoch that is too large can cause the model to overfit. Batch size is the number of data samples that pass simultaneously through the neural network [29]. According to Radiuk [30], the smaller the batch size value, the faster the training process. However, the larger the batch size value, the longer the training process will take because it requires more storage capacity. The learning rate, a configurable parameter, has a minimal positive value of 0 to 1. This parameter is used to determine the size of the step that determines whether the training results are getting better (loss value decreases) or worse (loss value increases) [31]. Some of the hyperparameters to be tuned are batch size (32; 64), epoch (50; 100; 150; 200) and learning rate (0,001; 0,005), neurons (4; 8; 16; 32; 64) [32, 33]. The four hyperparameters were chosen for tuning because they directly affect the ability and speed of the model to achieve optimal accuracy.

This research uses grid search in the scikit-learn package, namely GridSearchCV, with 10-fold cross-validation to perform hyperparameter tuning [34].They were delivering that the GridSearchCV model provides higher accuracy than models that do not use GridSearchCV. K-fold cross-validation generally uses 3; 5 or 10-fold. In this research, the k used is 10, where nine partitions in each fold are used for training and one partition for validation. Hyperparameter tuning shows the best hyperparameter combination results for modeling with LSTM Network: epochs 200; batch size 32; learning rate 0,001 and neurons 8. This model produces an average loss or error value of 0,0045.

3.4. Validation of ARIMA and LSTM Models

At this stage, the prediction results of the ARIMA and LSTM models are evaluated against actual data using MAPE and RMSE. RMSE evaluates the model's goodness by calculating the difference between the predicted and actual values. RMSE gives higher error weights for prediction results with larger errors. RMSE does not have a maximum or minimum value, but in model comparison, it is known that the model gets better if it has a smaller RMSE value. MAPE is used to evaluate the model by showing the error rate or the difference between the predicted value and the actual value in percentage form to make it easier to understand.

Table 6. Model evaluation

Model	RMSE	RMSE-P	MAPE-P(%)
ARIMA	4,03	3,87	28,70
LSTM	**2,32**	**2,35**	**19,40**

*RMSE: model evaluation on training data and
*MAPE-P: model evaluation on testing data

Based on the evaluation results in Table 6, it is known that the best model is LSTM. The model produces a testing data RMSE value of 2,35 and a testing data MAPE of 19,40% which falls into the excellent category for forecasting. A comparison of the evaluation results of the ARIMA and LSTM models is also shown in Figure 6 and Figure 7.

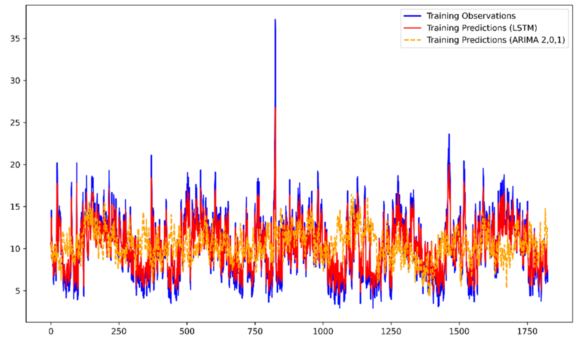


Figure 6. Comparison plot of training data with prediction data

Figure 6 compares the actual training data with the prediction results of the training data from the two models. Model training takes a long time, depending on the data and the complexity of the model. This can lead to overfitting events. One way to identify overfitting in the model is to predict and evaluate the model against the training data. If the prediction results are good, it is known that there is no overfitting in the model. Based on the comparison plot in Figure 6, it is known that the training data predicted by the two models is quite close to the actual training data. This is also supported by the RMSE value of the training data in Table 6, which is very small, so it is known that the two models work well to predict the training data or that there is no overfitting event.

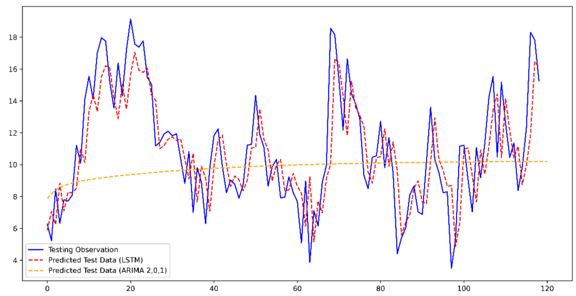


Figure 7. Comparison plot of testing data with prediction data

Figure 7 compares the actual testing data with the predicted testing data from the two models. The predicted data of the LSTM model is quite close to the actual testing data, so the LSTM model works very well. This is supported by the RMSE and MAPE values of the LSTM model testing data, which are the smallest according to the evaluation in Table 6, indicating that the LSTM model is better for forecasting the maximum wind speed in Kupang City, East Nusa Tenggara Province than the ARIMA model.

3.5. Forecasting

Forecasting the maximum wind speed for the May 2024 period was carried out using the LSTM model, as shown in Figure 8.

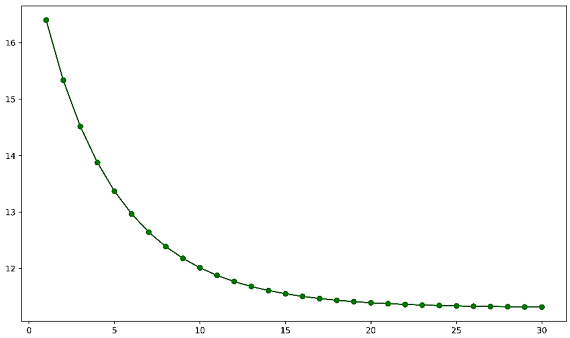


Figure 8. Maximum wind speed forecasting

The maximum wind speed in Kupang City, East Nusa Tenggara, has a downward trend, with the highest speed being 16,2 knot. Forecasting results for the next 30 days range from 10 to 16 knot, according to the BMKG statement that the East Nusa Tenggara region has begun the dry season, so the east monsoon winds have begun to be active. East monsoon winds are seasonal winds that blow from east to west in tropical and subtropical regions, with speeds of 10 to 25 knot [35].

4. Conclusion

The LSTM model is better than the ARIMA model in forecasting the maximum wind speed in Kupang City, East Nusa Tenggara Province. The best LSTM model has hyperparameter epochs 200, batch size 32, learning rate 0,001, and neurons 8. This method can learn trends and data patterns well. Based on the results of the evaluation of predicted data against actual data, the MAPE value of the LSTM model is 19,40% or falls into an outstanding category for forecasting maximum wind speed data in Kupang City, East Nusa Tenggara Province.

Author Contributions. **Indah Magfirrah:** Software, data curation, visualization, writing—review and editing. **Meisyatul Ilma:** Writing—original draft preparation, data curation. **Khairil Anwar Notodiputro:** Conceptualization, supervision. **Yenni Angraini:** Conceptualization, methodology, supervision. **Laily Nissa Atul Mualifah:** Conceptualization, validation, supervision. All authors have read and agreed to the published version of the manuscript.

Acknowledgement. The authors express their gratitude to the editor and reviewers for their meticulous reading, insightful critiques, and practical recommendations, all of which have greatly enhanced the quality of this work.

Funding. This research received no external funding.

Conflict of interest. The authors declare no conflict of interest related to this article.

References

[1] R. J. Hyndman and G. Athanasopoulos, "Forecasting: Principles and Practice (3rd ed)," Monash University: OTexts. Accessed: Jul. 02, 2024. [Online]. Available: https://otexts.com/fpp3/

[2] J. Patterson and A. Gibson, *"Deep Learning: A Practitioner's Approach."* Sebastopol, CA: O'Reilly Media, Inc., 2017.

[3] O. Al Qasem, M. Akour, and M. Alenezi, "The Influence of Deep Learning Algorithms Factors in Software Fault Prediction," *IEEE Access*, vol. 8, pp. 63945–63960, 2020, doi: 10.1109/ACCESS.2020.2985290.

[4] H. Abbasimehr, M. Shabani, and M. Yousefi, "An optimized model using LSTM network for demand forecasting," *Comput Ind Eng*, vol. 143, p. 106435, May 2020, doi: 10.1016/j.cie.2020.106435.

[5] J. N. Van Rijn and F. Hutter, "Hyperparameter importance across datasets," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Jul. 2018, pp. 2367–2376. doi: 10.1145/3219819.3220058.

[6] Z. Liu and X. Yang, "Cross validation for uncertain autoregressive model," *Commun Stat Simul Comput*, vol. 51, no. 8, pp. 4715–4726, Aug. 2022, doi: 10.1080/03610918.2020.1747077.

[7] E. S. Putri and M. Sadikin, "Prediksi Penjualan Produk Untuk Mengestimasi Kebutuhan Bahan Baku Menggunakan Perbandingan Algoritma LSTM dan ARIMA," *Format Jurnal Ilmiah Teknik Informatika*, vol. 10, no. 2, p. 162, 2021, doi: 10.22441/format.2021.v10.i2.007.

[8] M. Elsaraiti and A. Merabet, "A comparative analysis of the arima and lstm predictive models and their effectiveness for predicting wind speed," *Energies (Basel)*, vol. 14, no. 20, 2021, doi: 10.3390/en14206782.

[9] A. D. Milniadi and N. O. Adiwijaya, "Analisis Perbandingan Model Arima Dan Lstm Dalam Peramalan Harga Penutupan Saham (Studi Kasus : 6 Kriteria Kategori Saham Menurut Peter Lynch)," *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, vol. 2, no. 6, pp. 1683–1692, 2023, doi: 10.54443/sibatik.v2i6.798.

[10] K. Umam and Ardiansyah, "Perbandingan Metode ARIMA dan LSTM pada Prediksi Jumlah Pengunjung Perpustakaan," *MIND (Multimedia Artificial Intelligent Networking Database) Journal*, vol. 8, no. 2, pp. 119–129, 2023.

[11] F. K. Lutgens and E. J. Tarbuck, *The Atmosphere: An Introduction to Meteorology (13th Ed.)*. London: Pearson, 2018.

[12] T. Prasetyo *et al.*, "Perbandingan Kinerja Metode Arima, Multi-Layer Perceptron, dan Random Forest dalam Peramalan Harga Logam Mulia Berjangka yang Mengandung Pencilan," vol. 11, no. 2, pp. 265–274, 2024, doi: 10.25126/jtiik.2024117392.

[13] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *"Introduction to Time Series Analysis and Forecasting"*. New Jesrey: John Wiley & Sons, 2015.

[14] M. Yohansa, K. A. Notodiputro, and E. Erfiani, "Dynamic Time Warping Techniques for Time Series Clustering of Covid-19 Cases in DKI Jakarta," *ComTech: Computer, Mathematics and Engineering Applications*, vol. 13, no. 2, pp. 63–73, Nov. 2022, doi: 10.21512/comtech.v13i2.7413.

[15] M. Abror Gustiansyah *et al.*, "Aplikasi Model ARIMA dalam Peramalan Data Harga Emas Dunia Tahun 2010-2022," *Jurnal Statistika dan Aplikasinya*, vol. 7, no. 1, 2023, doi: 10.21009/JSA.07108.

[16] S. Novel Lase, M. Turnip, and E. Indra, "Application of Data Mining to Predicate Stock Price using Long Short Term Memory Method," *Jurnal Infokum*, vol. 10, no. 2, 2022, [Online]. Available: http://infor.seaninstitute.org/index.php/infokum/index.

[17] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," 1997, Accessed: Jul. 03, 2024. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.51.3117.

[18] N. Kalchbrenner, I. Danihelka, and A. Graves, "Grid Long Short-Term Memory," Jul. 2015, [Online]. Available: http://arxiv.org/abs/1507.01526.

[19] K. Zhang, W.-L. Chao, F. Sha, and K. Grauman, "Video Summarization with Long Short-term Memory," May 2016, [Online]. Available: http://arxiv.org/abs/1605.08110.

[20] B. Karlik and A. V. Olgac, "Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks." Accessed: Jul. 03, 2024. [Online]. Available: https://www.cscjournals.org/manuscript/Journals/IJAE/Volume1/Issue4/IJAE-26.pdf

[21] S. K. Sachin Kumar, A. D. Sachin Kumar, A. K. Agam Damaraju, S. K. Aditya Kumar, and C.-M. C. Saru Kumari, "LSTM Network for Transportation Mode Detection," *Journal of Internet Technology*, vol. 22, no. 4, pp. 891–902, Jul. 2021, doi: 10.53106/160792642021072204016.

[22] E. F. de S. Soares, C. A. V. Campos, and S. C. de Lucena, "Online travel mode detection method using automated machine learning and feature engineering," *Future Generation Computer Systems*, vol. 101, pp. 1201–1212, Dec. 2019, doi: 10.1016/j.future.2019.07.056.

[23] A. L. Schaffer, T. A. Dobbins, and S. A. Pearson, "Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions," *BMC Med Res Methodol*, vol. 21, no. 1, Dec. 2021, doi: 10.1186/s12874-021-01235-8.

[24] J. Brownlee, *"Long short-term memory networks with python: develop sequence prediction models with deep learning."* Machine Learning Mastery, 2017.

[25] J. D. Williams and G. Zweig, "End-to-end LSTM-based dialog control optimized with supervised and reinforcement learning," Jun. 2016, [Online]. Available: http://arxiv.org/abs/1606.01269.

[26] T. Phaladisailoed and T. Numnonda, "Machine Learning Models Comparison for Bitcoin Price Prediction," in *2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE)*, IEEE, Jul. 2018, pp. 506–511. doi: 10.1109/ICITEED.2018.8534911.

[27] M. Naeem and S. A. Maged, "Linear Time Invariant State Space System Identification Using Adam Optimization," in *2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE)*, IEEE, Feb. 2020, pp. 196–204. doi: 10.1109/ITCE48509.2020.9047808.

[28] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Trans Neural Netw Learn Syst*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.

[29] N. Bakhashwain and A. Sagheer, "Online Tuning of Hyperparameters in Deep LSTM for Time Series Applications," *International Journal of Intelligent Engineering and Systems*, vol. 14, no. 1, pp. 212–220, Feb. 2021, doi: 10.22266/ijies2021.0228.21.

[30] P. M. Radiuk, "Impact of Training Set Batch Size on the Performance of Convolutional Neural Networks for Diverse Datasets," *Information Technology and Management Science*, vol. 20, no. 1, Jan. 2017, doi: 10.1515/itms-2017-0003.

[31] N. Rochmawati *et al.*, "Analisa Learning rate dan Batch size Pada Klasifikasi Covid Menggunakan Deep learning dengan Optimizer Adam," 2021. doi: 10.26740/jieet.v5n2.p44-48.

[32] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, IEEE, Dec. 2018, pp. 1394–1401. doi: 10.1109/ICMLA.2018.00227.

[33] D. Fan, H. Sun, J. Yao, K. Zhang, X. Yan, and Z. Sun, "Well production forecasting based on ARIMA-LSTM model considering manual operations," *Energy*, vol. 220, p. 119708, Apr. 2021, doi: 10.1016/j.energy.2020.119708.

[34] G. N. Ahmad, H. Fatima, S. Ullah, A. Salah Saidi, and Imdadullah, "Efficient Medical Diagnosis of Human Heart Diseases Using Machine Learning Techniques With and Without GridSearchCV," *IEEE Access*, vol. 10, pp. 80151–80173, 2022, doi: 10.1109/ACCESS.2022.3165792.

[35] B. Wang, *The Asian Monsoon*. Berlin: Springer, 2006.