Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 6 No. 4 (2022) 563 - 569 ISSN Media Electronic: 2580-0760

Time Series Temperature Forecasting by using ConvLSTM Approach, Case Study in Jakarta

Faishal Raihanur Rasyid[1], Didit Adytia[2]
[1,2]School of Computing, Telkom University
[1]faishalraihan@student.telkomuniversity.ac.id, [2]adytia@telkomuniversity.ac.id*

Abstract

Climate change has occurred in several countries, especially in tropical countries such as Indonesia. It causes extreme temperature changes in several Indonesian areas, especially Jakarta, one of the world's most populated cities. The population of Jakarta causes the activities carried out by residents to be disturbed by extreme temperature changes. In addition, drastic temperature changes also affect the energy consumption used by residents. Therefore, it is necessary to predict temperature to determine future temperature conditions so that residents can plan their activities. Temperature forecast can be done in several ways, one of which uses a machine learning approach. This research uses a deep learning model called the Convolutional Long Short-Term Memory (ConvLSTM). Moreover, we also compare the model with Multi-Layer Perceptron (MLP), and Long Short-Term Memory (LSTM). We use temperature data taken from the ERA-5 period years 2018 to 2020 located in Kemayoran, Jakarta, Indonesia. This research aims to investigate the accuracy of short-term temperature forecasting by using these three models. The model is built to predict short-term temperatures for 1, 3, and 7 days ahead. The performance of the three methods is measured by calculating the Root Mean Square Error (RMSE), Mean Square Error (MAE), and Coefficient Correlation (CC). The result shows that the LSTM performs better than the other methods to forecast 1, 3, and 7 days, i.e., with the lowest RMSE, MAE, and higher CC.

Keywords: *Temperature Forecasting, Machine Learning, MLP, LSTM, ConvLSTM*

1. Introduction

Temperature is a component of weather and an entity that can be used to understand global warming. The extreme climate change in Indonesia causes quite extreme temperature changes in several areas in Indonesia such as Jakarta and Bekasi[1]. Besides that, according to most climatologists, extreme temperature changes occur because of human activities that cause greenhouse gases increased[2]. Temperature is also one factor that affects energy consumption[3]. According to some studies, 40% of Energy generated worldwide is used by buildings half of this is used for heating and air conditioning[4]. The increasing use of air conditioning and heating due to being easily influenced by the air temperature. Therefore, studying the air temperature fluctuations is necessary to estimate the energy requirements needed[5]. Everyone can find information about future temperature conditions by doing temperature forecasts. This information helps in water resource management, irrigation, and health[6]. Besides, temperature forecast is important for society because it helps their work[7]. Temperature forecast has been done by applying machine learning because it can provide high accuracy results[8]. Using machine learning to make temperature forecasts can speed up the prediction process because it has faster computational capabilities to handle complex meteorological data and large data sizes [9].

Several machine learning methods can be used to predict Temperature, such as the Long Short-Term method Memory (LSTM). LSTM method was used by Park and his colleagues in 2019 in three locations in South Korea, namely Seoul, Gyeonggi, and Jeolla. They made temperature predictions for the next 6, 12, 24 hours, and 14 days using weather data from November 1981 until December 2017. They compare performance from the LSTM method with the Deep Neural Network (DNN) method in predicting Temperature. The result obtained from that research is that the LSTM model provides performance better than the DNN model in predicting temperature[10]. Another machine learning method that can be used is Multi-Layer Perceptron (MLP) which Salcedo-Sanz and colleagues used in 2015 to predict ten locations in Australia and New

Accepted: 21-06-2022 | Received in revised: 23-07-2022 | Published: 22-08-2022

Zealand. They made a prediction with weather data from the period 1910-2010. The research was conducted by comparing several methods other than MLP, Multi Linear Regression (MLR) and Support Vector Regression (SVR). Two cases will be used to train the SVR and MLP models; namely, the 30/70 case, where the training data contains a monthly average temperature data from January 1900 to December 1930 and the test data includes the monthly mean temperature data from the period 1931 to 2010. In the 70/30 case, the training data contains a monthly mean temperature data from 1900 to December 1970, and the test data includes a monthly mean temperature data from 1971 to 2010. The results obtained from this research are MLP method gives slightly better performance than MLR in the case of 70/30. While in case 30/70, the MLP method gives worse results than other methods[11].

The Convolutional LSTM (ConvLSTM) method, a variation of the LSTM method, can also be applied to prediction temperature as Kreuzer and colleagues did in 2020 in German. They predicted temperature for 24 hours using weather data from five meteorological stations in Germany for the period 2009-2018. The research used the ConvLSTM method with the SARIMA method, univariate LSTM, and multivariate LSTM. The result obtained from that research is that the SARIMA model provides the best performance for the first few hours, followed by a univariate LSTM model. But a longer time, the SARIMA and the LSTM model univariate performance decreased compared to a multivariate model. ConvLSTM model provides good performance to predict across locations. In contrast, seasonal naive predictions yield results poor at making predictions. Naive forecast seasonal only gives the same result as other models when predicting for 20-24 hours[12].

As Jakarta is one of the most populated cities in Indonesia, residents carry out many activities carried out by residents leading to increased use of air conditioning and heater use due to weather conditions. Therefore, temperature forecasting can help design an electricity forecasting system. Based on previous research, this research aimed to show the result of ConvLSTM to predict the temperature in Jakarta. Moreover, the result of ConvLSTM is compared with two other methods, MLP and LSTM. The accuracy of each model is measured by looking at the metric evaluation used in this research; each model is used to make temperature forecasts. This research uses meteorological data from ERA-5 as input.

2. Research Methods

In this research, several steps are carried out, as in Figure 1. Firstly, the step of preparing climate data will be used to make temperature predictions. The data comes from ERA-5, which contains information about weather from the Kemayoran weather satellite in Jakarta. After the data was obtained, The ERA-5 dataset will be divided into training and test data, with 80% for training data and 20% for test data. After splitting data, the next step is to determine the parameters used in the implemented model, such as the number of layers used, the number of hidden layers used, the epoch value, and the activation function used.

Input parameters affect the resulting performance of the implemented model because the built model's structure affects the model's performance in making predictions. Therefore, the next step is to determine the appropriate input parameters for the model used so that the built model can provide good performance in this study. The next step is to build MLP, LSTM, and ConvLSTM models with predefined parameters. The last step is to analyze the performance of the model that has been built to predict Temperature by measuring the value of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient Correlation (CC) of each model. After getting the value RMSE, MAE, and Coefficient Correlation of each model, these values of the three methods will be compared to find the best method for predicting Temperature using ERA-5 data.

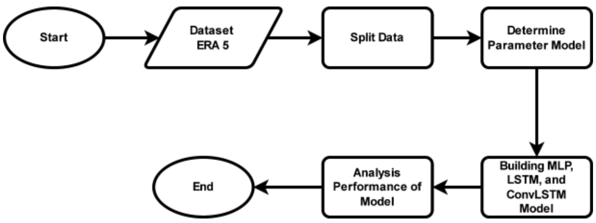

Figure 1. Methodology for temperature forecasting.

2.1 ERA-5 Dataset

The dataset used in this study is climate data from ERA-5 taken at a weather station located in Kemayoran Jakarta. The climate data is from 2018 to 2021. The dataset provides information on hourly climatic conditions in the Jakarta and Bekasi areas. There are two attributes of the ERA-5 dataset: skin temperature(skt) and air temperature at 2m(t2m). The machine learning model that is built in this research is a univariate machine learning model; therefore, the input used is only one attribute. The attribute used for input to three machine learning methods used in this research is temperature at 2m (t2m) because this attribute contains information about air temperature above 2 meters. As can be seen in Figure 2 is a visualization of the ERA-5 dataset based on temperature at 2m attributes. In this paper, we use three methods to forecast temperature with ERA-5 dataset, namely MLP, LSTM, ConvLSTM, which will be explained in the following subsection.

DOI: https://doi.org/10.29207/resti.v6i4.4162

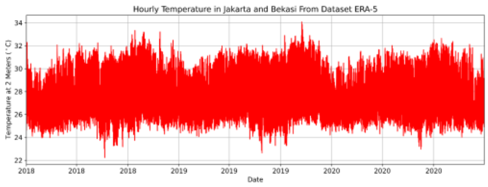

Figure 2. Temperature at 2m attribute from dataset ERA-5.

2.2 Multi-Layer Perceptron

Multi-Layer Perceptron (MLP) is a class in ANN, namely a feed-forward class. MLP consists of several neurons whose relationships are adjusted by constructing input-output mapping in a learning process to predict the sample[11]. MLP trains the weight that had using supervised learning, which is back-propagation. MLP is usually used to solve classification problems and forecasting[13]. MLP architecture consists of the set of the input, hidden layers, and output layers. Layers in MLP are interconnected, making neurons between layers that are close together and connected. However, neurons in the same layer do not have a relationship [14].

The input contains the raw data $(x_1, \ldots, x_n)$ and its weights $(w_1, \ldots, w_n)$ that will feed into the hidden layer as follows:

$$I = \sum_{i=1}^{n} x_i w_i \quad (1)$$

The hidden layer is a processing unit for the learning process to occur. hidden layer changes the value received from the input layer using activation function. The activation functions that usually used is sigmoid as follows:

$$\sigma = \frac{1}{1 + e^{-x}} \quad (2)$$

The output generated by the hidden layer is as follows:

$$H = f(A(I)) = f\left(A\left(\sum_{i=1}^{n} (x_i w_i)\right)\right) \quad (3)$$

where A is the activation function. If sigmoid gives:

$$H = \frac{1}{\left(1 + e^{-\sum_{j=1}^{m}(x_i w_i)}\right)} \quad (4)$$

The output layer receives the output and associated weights from the neurons on the hidden layer as input. The output layer assuming the function sigmoid is given as seen in equation (5):

$$O = f\left(A \sum_{j=1}^{m} (h_j w_j)\right) \quad (5)$$

Where h_j and w_j are the outputs and weights of each existing neuron in hidden layers. In the next subsection will be explained about second method used in this paper which is LSTM

2.3. Long Short-Term Memory

Long Short-Term Memory (LSTM) is one of the classes in RNN developed by Hochreiter to study long-term dependencies[15]. LSTM has been widely used in various kinds of research such as natural language recognition and time series prediction[16]. LSTM is also used when there is a case to solve the vanishing gradient problem of RNN in long-term context memorization [16]. The difference between LSTM and RNN is that there is a unique component in the hidden layer of LSTM. The unique component is the memory cell. Memory cells are an essential component in the LSTM architecture because in the LSTM architecture, as can be seen in Figure 3 in each memory cell there are three gates, namely input gate (i_t), forget gate (f_t), and output gate (o_t). In LSTM framework, memory cell is important part because memory cell has the capacity to extract cell state information. LSTM cells carry hidden state of the previous step (S_{t-1}) and input (x_t). Afterwards hidden state (S_t) will be calculated as following [17].

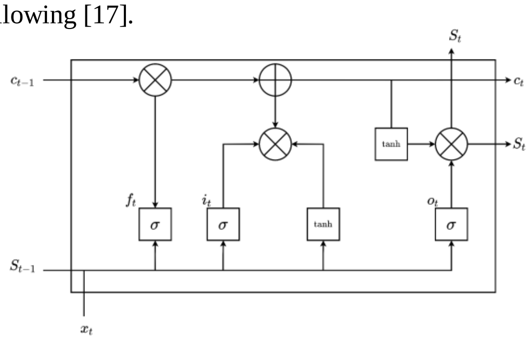

Figure 3. LSTM architecture.

The first stage carried out by LSTM is forget gate (f_t) decides what data to remove from cell state as seen in equation (6).

$$f_t = \sigma(W_{fx} x_t + W_{fs} S_{t-1} + b_f) \quad (6)$$

where is σ a sigmoid function. The next step is to decide what information to store in cell states. In determining this there are two steps. the first step is the input gate (i_t) selects the value to be updated.

$$i_t = \sigma(W_{ix} x_t + W_{is} S_{t-1} + b_i) \quad (7)$$

The second step is to use the tanh function to create a vector of the new value for applicant value $\tilde{c}_t$.

$$\tilde{c}_t = tanh(W_{\tilde{c}x} x_t + W_{\tilde{c}s} s_{t-1} + b_{\tilde{c}}) \quad (8)$$

After that update the cell state c_{t-1} with the new cell state c_t.

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \quad (9)$$

DOI: https://doi.org/10.29207/resti.v6i4.4162

Finally, the output gate (o_t)selects the components of the cell state that will be used as output.

$$o_t = \sigma(W_{ox}x_t + W_{os}S_{t-1} + b_o) \quad (10)$$

After that, the cell state passes through the tanh layer and multiplies as in the equation (11).

$$s_t = o_t \otimes tanh(c_t) \quad (11)$$

The next subsection will explain the last method used in this paper, Convolutional LSTM.

2.4 Convolutional LSTM

Convolutional LSTM(ConvLSTM) is a variation of LSTM that allows all existing gateways to use existing content in the previous memory cell[18]. ConvLSTM has been used in various kinds of research, such as video saliency detection, traffic accident prediction, and text recognition[19]. ConvLSTM has a complex structure in both input-to-state and state-to-state transitions because it can model spatial-temporal relationships quite well. Input $X_1, \ldots, X_t$, cell states $C_1, \ldots, C_t$, hidden states $H_1, \ldots, H_t$, input gate (i_t), forget gate (f_t), and output gate(o_t) of ConvLSTM is a 3D tensor that last two dimensions is a spatial dimension [20]. The following equations of ConvLSTM which Xingjian Shi has explained in his research about precipitation nowcasting using ConvLSTM [21].

$$i_t = \sigma(W_{xit} + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \quad (12)$$

$$f_t = \sigma\left(W_{xft} + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f\right) \quad (13)$$

$$o_t = \sigma(W_{xot} + W_{ho} * H_{t-1} + W_{co} \circ C_{t-1} + b_o) \quad (14)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ$$
$$\tanh\,(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \quad (15)$$

$$h_t = o_t \circ tanh(C_t) \quad (16)$$

Where ‘*’ is the convolution operator, and ‘∘’ is the Hadamard product. Moreover, σ is a sigma function and $W_{x\sim}$and $W_{h\sim}$are kernel 2-d convolution. This paper used three metric evaluations to help analyze the performance of some methods to predict temperature, which will be explained in the next subsection.

2.5 Metric Evaluation

In this research, the performance evaluation of the model was carried out to determine the best model of the three models that had been built, namely MLP, LSTM, and ConvLSTM. The performance of the model is measured by the value of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient Correlation (CC) which are methods used to measure the accuracy of the predictions made by the three models. Where in equation (17), (18), and (19) the value of *xi* is the result of the prediction by three methods that has been made, meanwhile the value of *yi* is the actual data from ERA-5 dataset, and the value of *n* is the size of data.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(xi - yi)^2}{n}} \quad (17)$$

$$MAE = \frac{\sum_{i=1}^{n}|xi - yi|}{n} \quad (18)$$

$$CC = \frac{\sum(xi - \bar{x})(yi - \bar{y})}{\sqrt{\sum(xi - \bar{y})^2 \sum(yi - \bar{y})^2}} \quad (19)$$

3. Results and Discussions

In this research, the ConvLSTM method will be implemented to show the performance to predict temperature. Here, we aim to perform short-term temperature prediction. Therefore, we propose three forecasting scenarios based on the prediction length, i.e., 1, 3, and 7 days ahead. Moreover, we also compare the performance of ConvLSTM with the other two machine learning methods. The two machine learning methods that will be used are MLP as one of the conventional machine learning methods and LSTM as the based model of ConvLSTM. The purpose of comparing those two models is to see how well ConvLSTM performs among those methods to predict temperature. The result will be analyzed based on the evaluation metrics used. The parameters used for the three models are shown in Table 1.

Table 1. Parameter used in ConvLSTM, MLP, and LSTM.

Parameter	ConvLSTM	MLP	LSTM
Epoch	100	100	100
Activation	ReLu	ReLu	ReLu
Optimizer	Adam	Adam	Adam
Batch Size	128	128	128
Units	200	100	400

As seen in Figure 4, these are the loss of methods for 100 iterations. Based on the visualization, LSTM perform well than other method when LSTM train the model with input from dataset ERA-5. The reason is LSTM convergence with less number of iteration than other method. Nevertheless, LSTM takes longer to train the model than other methods. It takes about 687.1833 seconds to train the model for LSTM as can be seen in Table 2. Meanwhile, MLP provides the fastest performance of all methods with 36.2713 seconds when training the model. Moreover, ConvLSTM takes shorter computational time than LSTM with 282.6990 seconds to train the model.

3.1 Result of 1 Day Prediction

The result of three methods to predict one day using dataset from ERA-5 can be seen in Figure 5, Figure 6, and Figure 7. Based on Table 3, the LSTM method prediction gives less error with the lowest with 0.3099 RMSE and 0.2443 MAE and Highest Coefficient Correlation with 0.9926 than other methods. Therefore,

DOI: https://doi.org/10.29207/resti.v6i4.4162

LSTM performs better than other methods to predict Temperature for 1-day scenario. Meanwhile, MLP in this scenario gives the worst result than other methods with the highest RMSE and MAE value and lowest coefficient correlation. In the other hand, ConvLSTM give a better performance than MLP but not quite well if the result of ConvLSTM compared to LSTM.

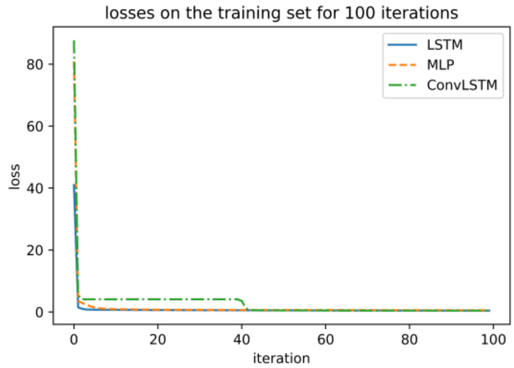

Figure 4. Losses for all of method for 100 iterations.

Table 1. CPU time for train model of all methods.

Method	CPU Time
MLP	**36.2713 (s)**
LSTM	687.1833 (s)
ConvLSTM	282.6990 (s)

Table 3. The result from MLP, LSTM, and ConvLSTM to predict 1 day ahead.

Method	Performance Evaluation		
	RMSE	MAE	CC
MLP	0.4822	0.3736	0.9797
LSTM	**0.3099**	**0.2443**	**0.9926**
ConvLSTM	0.3898	0.3376	0.9908

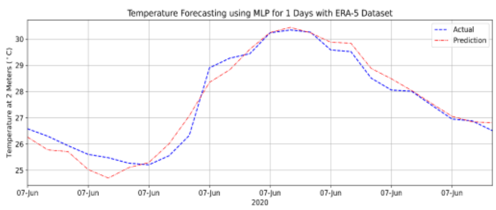

Figure 5. Result of prediction for 1 day ahead using MLP.

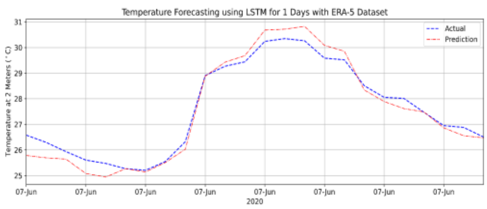

Figure 6. Result of prediction for 1 day ahead using LSTM.

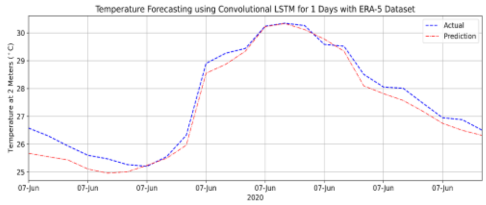

Figure 7. Result of prediction for 1 day ahead using ConvLSTM.

3.2 Result of 3 Days Prediction

The visualization of the result in 3 days scenario can be seen in Figure 8, Figure 9, and Figure 10. In this scenario, the prediction error of all methods increases from the previous scenario. However, based on the result shown in Table 4, the best method is still the same as the previous scenario because the LSTM prediction gives less error from the actual data with 0.3755 RMSE and 0.2734 MAE and Highest Coefficient Correlation value with 0.9865 than other methods. Moreover, ConvLSTM still provides better performance than MLP. The MLP gives the worst result from all three methods in this scenario but still not quite well than LSTM.

Table 4. The result from MLP, LSTM, and ConvLSTM to predict 3 days ahead.

Method	Performance Evaluation		
	RMSE	MAE	CC
MLP	0.4572	0.3764	0.9755
LSTM	**0.3755**	**0.2734**	**0.9865**
ConvLSTM	0.3898	0.2993	0.9799

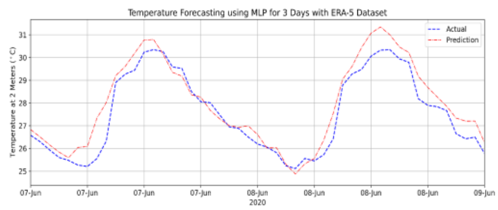

Figure 8. Result of prediction for 3 days ahead using MLP.

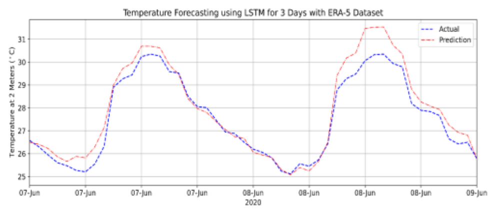

Figure 9. Result of prediction for 3 days ahead using LSTM.

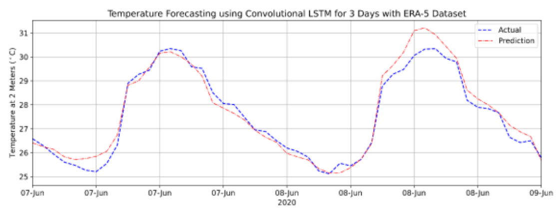

Figure 10. Result of prediction for 3 days ahead using ConvLSTM.

DOI: https://doi.org/10.29207/resti.v6i4.4162

3.3 Result of 7 Days Prediction

The result in this scenario is still the same from previous scenarios in terms of which give the best result. The difference is only on the increased prediction error from all methods, as seen in Figure 11, Figure 12 and Figure 13. Based on Table 5, the LSTM prediction performs well with less error with 0.4113 RMSE and 0.3277 MAE and Highest Coefficient Correlation than other methods in this scenario. Furthermore, ConvLSTM is slightly worse than LSTM, but if the performance is compared to MLP then ConvLSTM provides superior performance than MLP in this scenario.

Table 5. The result from MLP, LSTM, and ConvLSTM to predict 7 days ahead.

Method	Performance Evaluation		
	RMSE	MAE	CC
MLP	0.5381	0.4348	0.9617
LSTM	**0.4113**	**0.3277**	**0.9774**
ConvLSTM	0.4579	0.3598	0.9737

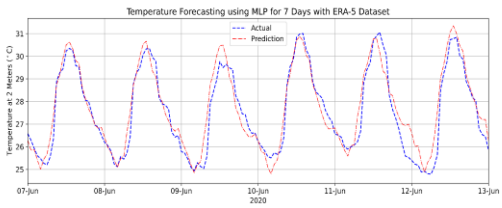

Figure 11. Result of prediction for 7 days ahead using MLP.

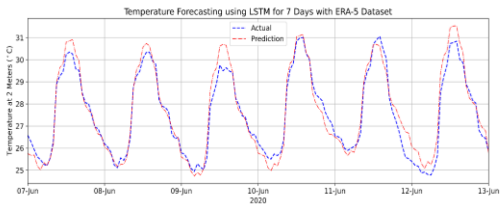

Figure 12. Result of prediction for 7 days ahead using LSTM.

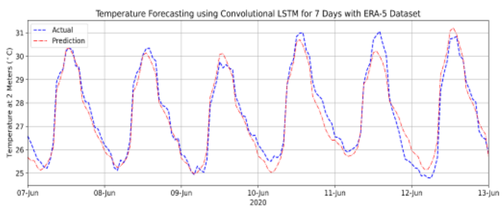

Figure 13. Result of prediction for 7 days ahead using ConvLSTM.

Table 6 shows the comparison result for 1, 3, and 7 days ahead using three methods used in this paper. As can be seen, all three methods perform better when they predict in a 1-day scenario than in other scenarios. That indicates that the data size affects the model's performance when predicting temperature.

Table 6. The result from MLP, LSTM, and ConvLSTM to predict 1, 3, and 7 days ahead.

Prediction	Method	Performance Evaluation		
		RMSE	MAE	CC
1 Day	**MLP**	**0.4822**	**0.3736**	**0.9797**
	LSTM	**0.3099**	**0.2443**	**0.9926**
	ConvLSTM	**0.3898**	**0.3376**	**0.9908**
3 Days	MLP	0.4752	0.3764	0.9755
	LSTM	0.3755	0.2734	0.9865
	ConvLSTM	0.3898	0.2993	0.9799
7 Days	MLP	0.5381	0.4348	0.9671
	LSTM	0.4113	0.3277	0.9774
	ConvLSTM	0.4579	0.3598	0.9737

4. Conclusion

This research developed the performance of ConvLSTM to predict the temperature in Jakarta, Indonesia. The performance of ConvLSTM was analyzed based on comparison results with the performance of two other methods, namely LSTM and MLP. Three scenarios were tested to see their performances to predict Temperature with ERA-5 dataset accurately. Based on the result, the best method to predict in all scenarios is the same i.e., LSTM. It performs well in all scenarios with 0.3099 of RMSE, 0.2443 of MAE and 0.9926 of CC in 1 day scenario; 0.3755 RMSE, 0.2734 MAE, and 0.9865 CC in 3 days scenario, and 0.4113 RMSE, 0.3277 MAE, and 0.9774 CC in 7 days scenario. Nevertheless, the performance of ConvLSTM to predict temperature in all of the scenarios is relatively slightly similar to the performance of LSTM with 0.3898 RMSE, 0.3376 MAE, and 0.9908 CC in 1 day scenario, 0.3898 RMSE, 0.2993, and 0.9799 CC in 3 days scenario, and 0.4579 RMSE, 0.3598 MAE, and 0.9737 CC in 7 days scenario. Although LSTM performs slightly better than ConvLSTM; besides LSTM takes a lot of time than ConvLSTM to train the model. Meanwhile, MLP performs worse than other methods in every scenario that has been tested despite giving a faster computational time when MLP trains the model than other methods.

Reference

[1] C. Johnstone and E. D. Sulungu, “Application of neural network in prediction of temperature: a review,” *Neural Computing and Applications*, vol. 33, no. 18. Springer Science and Business Media Deutschland GmbH, pp. 11487–11498, Sep. 01, 2021. doi: 10.1007/s00521-020-05582-3.

[2] E. Fahimi Nezhad, G. Fallah Ghalhari, and F. Bayatani, “Forecasting Maximum Seasonal Temperature Using Artificial Neural Networks ‘Tehran Case Study,’” *Asia-Pacific Journal of Atmospheric Sciences*, vol. 55, no. 2, pp. 145–153, May 2019, doi: 10.1007/s13143-018-0051-x.

[3] M. A. Lokoshchenko and N. A. Nikolaeva, “Dependence of energy consumption on air temperature in Moscow,” *Russian Meteorology and Hydrology*, vol. 42, no. 12, pp. 783–791, Dec. 2017, doi: 10.3103/S1068373917120068.

[4] B. Spencer, O. Alfandi, and F. Al-Obeidat, “A Refinement of Lasso Regression Applied to Temperature Forecasting,” in *Procedia Computer Science*, 2018, vol. 130, pp. 728–735. doi: 10.1016/j.procs.2018.04.127.

[5] Y. Hirano, K. Gomi, S. Nakamura, Y. Yoshida, D. Narumi, and T. Fujita, “Analysis of the impact of regional temperature pattern on the energy consumption in the commercial sector in Japan,” *Energy and Buildings*, vol. 149, pp. 160–170, Aug. 2017, doi: 10.1016/j.enbuild.2017.05.054.

[6] B. B. Karthika S, “MODELING OF AIR TEMPERATURE USING HYBRID WAVELET TRANSFORM-ANFIS-SUPPORT VECTOR MACHINE COMPUTING TECHNIQUES,” 2016.

[7] Z. Zhang, Y. Dong, and Y. Yuan, “Temperature Forecasting via Convolutional Recurrent Neural Networks Based on Time-Series Data,” *Complexity*, vol. 2020, 2020, doi: 10.1155/2020/3536572.

[8] J. Cifuentes, G. Marulanda, A. Bello, and J. Reneses, “Air temperature forecasting using machine learning techniques: A review,” *Energies*, vol. 13, no. 6. MDPI AG, Aug. 01, 2020. doi: 10.3390/en13164215.

[9] E. Abrahamsen, O. M. Brastein, and B. Lie, “Machine Learning in Python for Weather Forecast based on Freely Available Weather Data,” in *Proceedings of The 59th Conference on imulation and Modelling (SIMS 59), 26-28 September 2018, Oslo Metropolitan University, Norway*, Nov. 2018, vol. 153, pp. 169–176. doi: 10.3384/ecp18153169.

[10] I. Park, H. S. Kim, J. Lee, J. H. Kim, C. H. Song, and H. K. Kim, “Temperature prediction using the missing data refinement model based on a long short-term memory neural network,” *Atmosphere (Basel)*, vol. 10, no. 11, Nov. 2019, doi: 10.3390/atmos10110718.

[11] S. Salcedo-Sanz, R. C. Deo, L. Carro-Calvo, and B. Saavedra-Moreno, “Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms,” *Theoretical and Applied Climatology*, vol. 125, no. 1–2, pp. 13–25, Jul. 2016, doi: 10.1007/s00704-015-1480-4.

[12] D. Kreuzer, M. Munz, and S. Schlüter, “Short-term temperature forecasts using a convolutional neural network — An application to different weather stations in Germany,” *Machine Learning with Applications*, vol. 2, p. 100007, Dec. 2020, doi: 10.1016/j.mlwa.2020.100007.

[13] A. N. Ahmed, T. van Lam, N. D. Hung, N. van Thieu, O. Kisi, and A. El-Shafie, “A comprehensive comparison of recent developed meta-heuristic algorithms for streamflow time series forecasting problem,” *Applied Soft Computing*, vol. 105, Jul. 2021, doi: 10.1016/j.asoc.2021.107282.

[14] A. Olawoyin and Y. Chen, “Predicting the future with artificial neural network,” in *Procedia Computer Science*, 2018, vol. 140, pp. 383–392. doi: 10.1016/j.procs.2018.10.300.

[15] T. T. K. Tran, S. M. Bateni, S. J. Ki, and H. Vosoughifar, “A review of neural networks for air temperature forecasting,” *Water (Switzerland)*, vol. 13, no. 9. MDPI AG, May 01, 2021. doi: 10.3390/w13091294.

[16] A. Sagheer and M. Kotb, “Time series forecasting of petroleum production using deep LSTM recurrent networks,” *Neurocomputing*, vol. 323, pp. 203–213, Jan. 2019, doi: 10.1016/j.neucom.2018.09.082.

[17] T. T. Kieu Tran, T. Lee, J. Y. Shin, J. S. Kim, and M. Kamruzzaman, “Deep learning-based maximum temperature forecasting assisted with meta-learning for hyperparameter optimization,” *Atmosphere (Basel)*, vol. 11, no. 5, May 2020, doi: 10.3390/ATMOS11050487.

[18] M. Rahman, D. Islam, R. J. Mukti, and I. Saha, “A deep learning approach based on convolutional LSTM for detecting diabetes,” *Computational Biology and Chemistry*, vol. 88, Oct. 2020, doi: 10.1016/j.compbiolchem.2020.107329.

[19] S. W. Lee and H. Y. Kim, “Stock market forecasting with super-high dimensional time-series data using ConvLSTM, trend sampling, and specialized data augmentation,” *Expert Systems with Applications*, vol. 161, Dec. 2020, doi: 10.1016/j.eswa.2020.113704.

[20] G. Zhu, L. Zhang, P. Shen, and J. Song, “Multimodal Gesture Recognition Using 3-D Convolution and Convolutional LSTM,” *IEEE Access*, vol. 5, pp. 4517–4524, 2017, doi: 10.1109/ACCESS.2017.2684186.

[21] X. Shi *et al.*, “Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting.”

DOI: https://doi.org/10.29207/resti.v6i4.4162