UNPAD

Vol. 09, No. 01 (2025) 75 – 88
JIIF (Jurnal Ilmu dan Inovasi Fisika)
https://jurnal.unpad.ac.id/jiif/

HYBRID SARIMA-LSTM MODEL FOR PREDICTING EROSION IN BUTTERFLY VALVE

AZHAR AIMAN DZULFIQAR[1], JULIAN EVAN CHRISNANTO[1,3], BUDI ADIPERDANA[2,*]

[1]*Program Studi Fisika, FMIPA, Universitas Padjadjaran*
[2]*Departemen Fisika, FMIPA, Universitas Padjadjaran*
Jl. Raya Bandung-Sumedang km.25, Tanjungsari 45362, Sumedang, Jawa Barat
[3]*PT Padjadjaran Information Techology (PLABS.ID)*
Jl. Setrasari Tengah No.10, Sukarasa, Kec. Sukasari, Kota Bandung, Jawa Barat 40512

**Corresponding author*
Email: b.adiperdana@unpad.ac.id

Diserahkan: 20/01/2025
Diterima: 05/02/2025
Dipublikasikan: 12/02/2025

Abstrak. Dalam industri minyak dan gas, katup seringkali mengalami erosi akibat diletakkan pada lokasi dengan tekanan dan suhu yang tinggi. Erosi pada katup dapat mengakibatkan kerugian besar pada industri minyak dan gas sehingga memerlukan antisipasi yang sedini mungkin. Untuk mengatasi permasalahan tersebut, dibuat model *hybrid* SARIMA-LSTM untuk melakukan prediksi massa erosi pada *butterfly valve* pada beberapa kondisi bukaan katup. Hasil pengujian menunjukkan model SARIMA-LSTM memiliki performa yang unggul dibandingkan model LSTM dan model SARIMA dengan nilai MSE pada bukaan katup 20° – 90° mencapai 1E-06; 1E-06; 6.2E-05; 2.34E-04; 1.35E-07; dan 1E-06 secara berurutan. Model *hybrid* SARIMA-LSTM berhasil mengidentifikasi karakteristik non-linear dari data dengan mengidentifikasi nilai residu yang dihasilkan dari selisih prediksi model SARIMA dengan data aktual. Pengujian ini juga menunjukkan hasil bahwa kombinasi antara model SARIMA dan LSTM secara signifikan mempengaruhi performa model LSTM. Penelitian ini juga berhasil menggunakan model SARIMA-LSTM untuk memprediksi nilai massa erosi untuk 30 *time-steps* selanjutnya. Melalui penelitian ini diketahui bahwa model *hybrid* SARIMA-LSTM memiliki kemungkinan untuk diterapkan pada industri minyak dan gas guna membantu proses pengamatan massa erosi pada katup.

Kata kunci: *butterfly valve, artificial intelligence,* SARIMA, LSTM, massa erosi, deret waktu

Abstract. *In the oil and gas industry, butterfly valves often undergo erosion due to being placed in fluid flow with high pressure and high temperature. Erosion on butterfly valves can result in huge losses so it requires early anticipation. To overcome these problems, a hybrid SARIMA-LSTM model was created to predict the mass erosion of butterfly valves under several opening conditions. The results show that the SARIMA-LSTM model has superior performance compared to the conventional LSTM and SARIMA model with MSE values at valves opening 20° – 90° reaching 1E-06; 1E-06; 6.2E-05; 2.34E-04; 1.35E-07; and 1E-06 respectively. The hybrid SARIMA-LSTM model successfully identifies the non-linear characteristics of the erosion data by identifying the residual value resulting from the difference between the SARIMA model prediction and the actual data. This test also show that the combination of SARIMA and LSTM models significantly affects the performance of the LSTM model. This study also successfully used the SARIMA-LSTM model to predict the erosion mass value for the next 30 time-steps. Through this study, it is known that the SARIMA-LSTM hybrid model has the possibility to be*

DOI: https://doi.org/10.24198/jiif.v9i1 *e-ISSN: 2549-7014*

applied to the oil and gas industry to help the process of observing the erosion mass on the butterfly valve.

Keywords: *butterfly valve, artificial intelligence,* SARIMA, LSTM, erosion mass, time-series

1. Introduction

Butterfly is a rotary valve with a seat element in the form of a disc that can be rotated to open and close the flow [1]. In the oil and gas industry, butterfly valves are used as an isolator and regulator of oil flow in the upstream segment. Butterfly valves are usually chosen because they have a simple shape, good durability, and are easy to operate [2]. Butterfly valves have a disc section located in the centre of the pipe and have a stem that connects the disc and actuator on the outside of the valve [3]. This type of valve is usually placed in a flow with high pressure and temperature [2]. The disc of this valve is always placed in the fluid flow, this makes the pressure drop always occur around the butterfly valve [3]. Due to its application in extreme fluid flow conditions, flow separation, secondary flow, and vortexes occur inside a butterfly valve [4]. This extreme condition makes the butterfly valves prone to leakage due to erosion [2].

In the oil and gas production process, the fluid flow from production wells often has contaminant particles in the form of solids such as salt, sulfur, and sand [5]. Particles in the fluid flow can hit the butterfly valve part, causing erosion which can ultimately result in valve passing (leakage) [2]. Erosion is a process of surface material loss resulting from recurrent collisions between the wall/valve surface and solid particle carried by the fluid and often occurs extensively in production, transmission, and processing equipment [6]. Researh show that 5-10% of valves in oil and gas industry are affected by internal leakage, which can result in economic loss, health and safety issues, or potetially to environment pollution [7]. This loss can be prevented by knowing the erosion rate of each valve so that it can be known then the valve needs maintenance or needs to be replaced.

To date, there are several studies that discuss erosion on valves caused by solid particles in fluid flow. Liu et al. [8] conducted an erosion study due to solid particles on a butterfly valve. In the study, a CFD (Computational Fluid Dynamics) simulation model was created showing erosion due to two-phase flow by varying several parameters such as inlet velocity, particle mass fraction, and solid particle diameter. The result showed that the three parameters were directly proportional to the erosion rate of the butterfly valve. Lin et al. [9] conducted research about erosion on ball valves due to two-phase flow. In this study, a CFD-DEM (Discrete Element Model) model was made to determine the effect of valve opening and particle size on two-phase fluid flow and erosion characteristic in the ball valve. The result shows that under the condition of small valve opening degree, the erosion rate decreases as the particle diameter increases. Hanon and Ebrahimi [10] also conducted CFD simulation to investigate the erosion on butterfly valve. Erosion was observed by looking at the effect of several parameters, namely inlet pressure and valve opening angle. The result show that an increase in inlet pressure led to and increase in maximum velocity, turbulence intensity, shear stress, and particle erosion. However, turbulent intensity, particle erosion, and shear stress were reduced as the valve opening angle was decreased. At valve opening angle less than 50°, it was found that the parameter limitation decreased. Then, an increase in cavitation occurs when the valve opening is reduced and the inlet pressure is increased. It was found that the erosion at the front of the disk was due to particle erosion and the back due to cavitation erosion. Zhang et al. [11] conducted CFD-DEM simulations to observe liquid-solid two-phase flow and erosion in Y-type Slurry valve with different seat structures. This study conducted an observation on velocity and pressure distribution and surface erosion of the valve parts. The simulation results show that the variation of particle diameter and inner angle of the valve seat sealing ring greatly impacts the erosion of the back surface of the valve body, the inner surface of the valve seat, and the front surface of the valve disc.

Nowadays, the study of erosion on the valve is not only limited to the scope of CFD, but there is a development of the discussion to predict the mass erosion on oil and gas pipelines or valve using machine learning models. Othman et al. [12] conducted research about integration of CFD analysis and artificial neural networks for estimation of erosion rate in oil and gas pipelines. In this study, ANN (Artificial Neural Networks) models have been developed using data obtained from parametric studies of CFD simulations. The ANN model successfully predicted the maximum erosion rate on the bend pipe wall and show the similar pattern compared to CFD data with almost 95% comparable result data obtained with the error on RMSE (Root Mean Squared Error) is less than 10% . Wang et al. [13] studied solid particle erosion prediction in elbows based on machine learning and swarm intelligence algorithm. In this study, they used several machine learning models such as BPNN (Back Propagation Neural Network), SVR (Support Vector Regression), ELM (Extreme Learning Machine) and KELM (Kernel Extreme Learning Machine) to predict erosion on gas-solid flow simulation. The results shows that KELM model has a better fit to the actual erosion values and considered as the optimal model. Liu et al. [2] conducted research on the use of PSO-BP (Particle Swarm Optimizer-Back Propagation) model and PSO-LSTM (Particle Swarm Optimizer-Long Short Term Memory) to predict erosion on butterfly valves. In the study, the erosion rate was first examined through CFD simulation by varying valve opening and particle diameter. The result shows that erosion rate increased over time with larger valve openings. In addition, this study also showed that the PSO-LSTM model has a better ability than PSO-BP in predicting erosion rates.

In this research, the authors try to improve the accuracy in predicting the erosion rate of butterfly valves by introducing a hybrid SARIMA-LSTM (Seasonal Autoregressive Integrated Moving Average--Long-Short Term Memory) model to predict the erosion rate of butterfly valves based on CFD simulation data. In this model, SARIMA is a type of statistical model that is appropriate for handling time series data. Meanwhile, LSTM is a type of RNN with the additional feature of being able to remember data sequences, remember data trends up to a certain point in time through several gates and memory paths. LSTM has the ability to learn long-term dependencies in data and is widely used to solve problems in time series data. The use of the SARIMA-LSTM hybrid model is intended to overcome the shortcomings of SARIMA to solve problem on non-stationary time series and the shortcomings of LSTM which require a long training time and difficulty in determining the right and perfect combination of parameters [14]. The SARIMA-LSTM hybrid model is expected to improve accuracy in predicting erosion rates which are time-series data.

2. Research Methods

In this study, erosion mass data obtained from Liu et al. [2] research. This dataset involves erosion mass data with different valve openings and the particle size is set to 0.2

mm. The dataset used is the erosion mass data obtained from CFD simulations with a total flow time of 10 s which has a size of 10000 time-steps with a step size of 0.001.

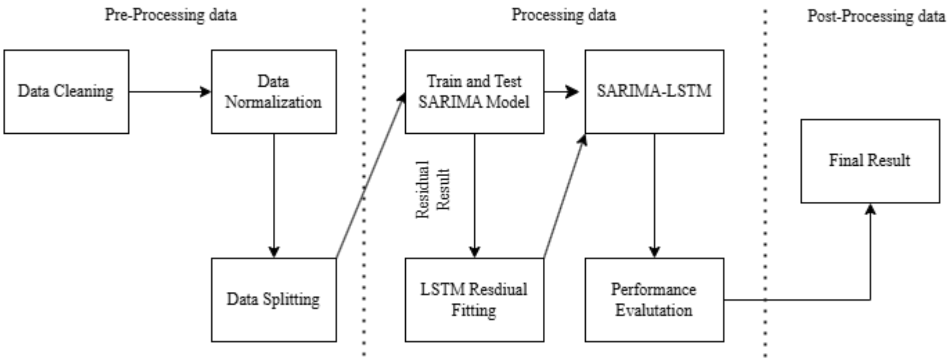

Figure 1. Flow Diagram of SARIMA-LSTM Hybrid Model

This study was conducted with three main stages such as preprocessing, processing, and post-processing (**Figure 1**).

2.1 Pre-Processing Data

In the pre-processing data stage, several processes are carried out to prepare data such as data cleaning, data normalization, and data splitting.

2.1.1 Data Cleaning

Data cleaning is a critical initial step to ensuring that the dataset is devoid of erroneous data. This process aims to make the data ready for analysis [15]. This process is done by fixing or removing duplicate or irrelevant data.

2.1.2 Data Normalization

Data Normalization is a stage to transform or scale data into certain range to make an equal contribution of each feature. Data Normalization is important for improving data quality and machine learning performance [16]. In this study, the Min-Max Normalization method is used to change the data value so that it has a value in the range of 0 - 1 by utilizing the Python Sklearn library. Normalization basically aims to make the data more structured and can be easily analyzed by the SARIMA-LSTM model. The mathematical equation of the Min-Max normalization method is as follows:

$$X' = \frac{X - X_{min}}{X_{max} X_{min}} \quad (1)$$

2.1.3 Data Splitting

Data splitting is a process to divide the data into several parts, namely train data and test data. This process aims to avoid overfitting, a condition where machine learning prevent us from perfectly generalizing the models to well fit observed data on training data, as well as unseen data on training dataset [17]. In this research, the data will be divided into training and testing dataset with a percentage of 80% and 20% respectively.

2.2 Processing Data

The data processing stage is a stage that includes the process of training and testing the SARIMA model, calculating the residual value, fitting the LSTM model to the residual data, until the final result of the SARIMA-LSTM hybrid model is obtained.

2.2.1 ARIMA

Auto Regressive Integrated Moving Average (ARIMA) is a statistical method for forecasting univariate time series dataset [18]. ARIMA Model is a combination of Autoregressive (AR) which is a regression model of a variable against itself and Moving Average (MA) model which is a model to see the movement of residuals from the past. ARIMA models are intended to describe autocorrelation in data.

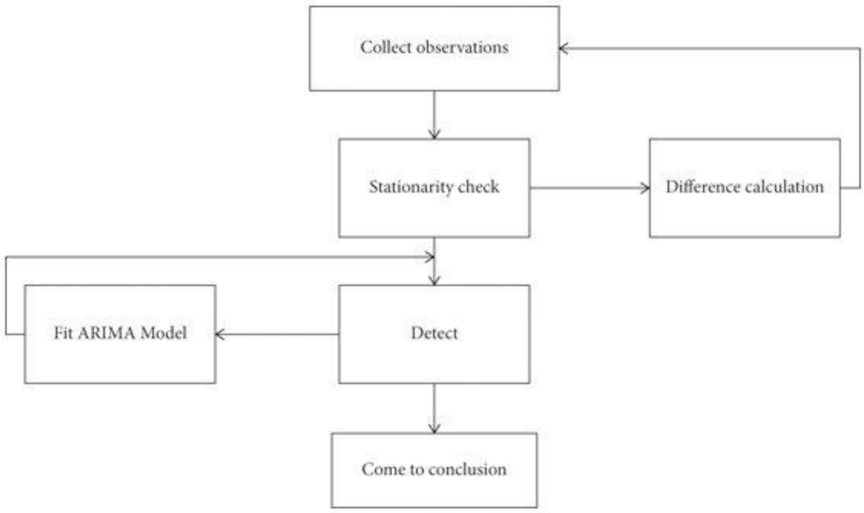

Figure 2. ARIMA Architecture [19]

The ARIMA model requires time series data with stationary characteristics. A stationary time series is a time series whose characteristics do not depend on the time of observation. A time series is classified as a non-stationary if the time series has a trend or seasonality or does not have drastic changes in the data. A stationary time series has constant fluctuations in value, and is independent of the time and variance of fluctuations in data value. Mathematically, the ARIMA model can be written as follows [20]:

$$\phi_p B(1-B)^d y_t = \theta_q B \varepsilon_t \tag{2}$$

Equation (2) can be changed into:

$$1 - B^d\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right) y_t = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right)\varepsilon_t \tag{3}$$

2.2.2 SARIMA

The SARIMA model, or Seasonal Autoregressive Integrated Moving Average is an extension of the ARIMA model that can be used to predict time series data with seasonal components. The SARIMA model has an order denoted as $(p, d, q)(P, D, Q)_S$, where p represents the non-seasonal component for autoregressive, d is the non-seasonal component for differencing, q is the non-seasonal component for moving average, P is the seasonal order for autoregressive, D is the seasonal order for differencing, and Q is the seasonal order for moving average [21]. Mathematically, the SARIMA model is expressed as follows:

$$(1 - \Phi B^p)(1 - \phi_1 B^s) y_t (1 - B)^d (1 - B^s) = (1 - B\theta)(1 - \theta B^s) e_t \tag{4}$$

The data processing stage will begin by finding the SARIMA order using the help of the autoarima() library. The obtained SARIMA model order will be used in the model fitting process with training data. After that, the trained SARIMA model will be tested to predict the data based on the test data. Through this process, the residual value will be obtained. which characterizes the difference value between the actual erosion mass data and the erosion data predicted by the SARIMA model.

2.2.3 LSTM

The residual value obtained from the SARIMA model will be given to the LSTM model for the fitting process. Long-Short Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) model that has the ability to learn long sequences of data and store long-term information that can be used to predict time series data [22]. LSTM has the ability to filter out redundant information. In simple terms, the architecture of LSTM is as follow:

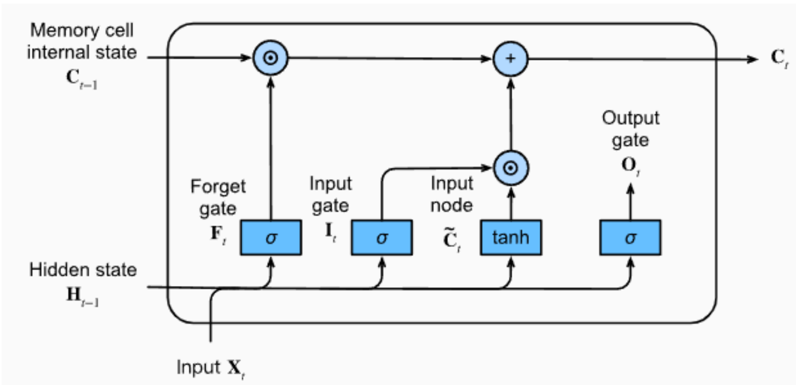

Figure 2. LSTM Architecture [23]

LSTM architecture has a memory cell that is controlled by three gates, namely input gate, forget gate, output gate. Basically, these three gates allow the LSTM to effectively determine what information to add, delete, and remove from the memory cell [14].

Input gate serve to control what information will be added to the memory cell

$$i_t = \sigma(w_i[h_{t-1}, X_t] + b_i) \tag{5}$$

$$\widehat{C_t} = \tanh(w_c[h_{t-1}, X_t] + b_c) \tag{6}$$

Forget gate serves to remove information from the memory cell

$$f_t = \sigma\left(w_f[h_{t-1}, X_t] + b_f\right) \tag{7}$$

The output gate serves as place for information to leave the memory cell

$$o_t = \sigma(w_o[h_{t-1}, X_t] + b_o) \tag{8}$$

$$h_t = o_t \odot \tanh(c_t) \tag{9}$$

This is what makes the LSTM network able to selectively maintain or delete information as long as data flows through the network and allows learning long-term dependencies. Furthermore, LSTM also has a hidden state that acts as short-term memory and will be updated based on input information, previous hidden state, and current memory state.

2.2.4 Hybrid SARIMA-LSTM

The final prediction results of the SARIMA-LSTM hybrid model are obtained by summing the predicted trend and seasonal components produced by the SARIMA model with the residual prediction values from the LSTM model. This process is based on the principle of additive decomposition. The additive model is a common approach that assumes the summation of seasonal, trend-cycle, and irregular components. Mathematically, it can be expressed as follows [24]:

$$Y_t = S_t + T_t + E_t \tag{10}$$

2.2.5 Model Evaluation

In this study, the performance evaluation of the machine learning models was conducted using Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). MSE measures the average of the squared differences between the predicted values and the actual data values. MSE serves as a useful parameter to determine the performance of the model. The smaller the MSE, the better the model’s performance. The MSE values is calculated using the following formula:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2 \tag{11}$$

Root Mean Squared Error (RMSE) is another machine learning loss metrics to test model performance of SARIMA-LSTM. RMSE measures the average differences between the predicted values of the machine learning model and the actual values. RMSE provides an estimate of how well a model can predict target values. The closer the RMSE value is to zero, the better the model’s performance in making predictions. The RMSE value can be calculated using the following equations:

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{\left(y_{prediksi,i} - y_{aktual,i}\right)^2}{n}} \tag{12}$$

2.3 Post-Processing Data

Finally, in the data post-processing stage, a comparison and analysis are conducted regarding the SARIMA-LSTM predicted result against actual data. Additionally, the hybrid SARIMA-LSTM model is used to forecast the erosion mass values for the next 30 time-steps.

3. Results and Discussion

Based on the experiments that have been conducted, we conduct an evaluation to the SARIMA-LSTM performance in predicting the erosion mass on 6 different datasets. Performance test done by calculating and comparing the MSE value between SARIMA-LSTM, SARIMA, and LSTM models. Performance of each model can be seen on **Table 1**.

Table 1. MSE Values

Valve Opening	MSE		
	LSTM	SARIMA	SARIMA-LSTM
20°	1.2E-05	1.226E-03	1E-06
30°	7.3E-05	3.72E-04	1E-06
45°	1.39E-04	0.481952	6.2E-05
60°	3.57E-04	0.720209	2.34E04
75°	4.36E-07	2.70E-03	1.35E-07
90°	2.2E-05	1.53E-02	1E-06

Based on **Table 1**, it is known that the LSTM model has better performance compared to the SARIMA model. With the characteristics of the erosion mass data used is data that has nonlinear tendency, this clearly shows that LSTM has better performance in identifying nonlinear data. As for the SARIMA model, although it is equipped with the ability to detect seasonal components in the data, it is insufficient to understand the erosion mass data with complex characteristics due to the erosion mass value being influenced by various physical parameters. However, the hybrid SARIMA-LSTM model shows superior performance compared to LSTM and successfully addresses the shortcomings of SARIMA in understanding non-stationary and non-linear data. The SARIMA-LSTM model is able to demonstrate its ability to identify non-linear characteristics by identifying residuals obtained from the SARIMA model and actual data.

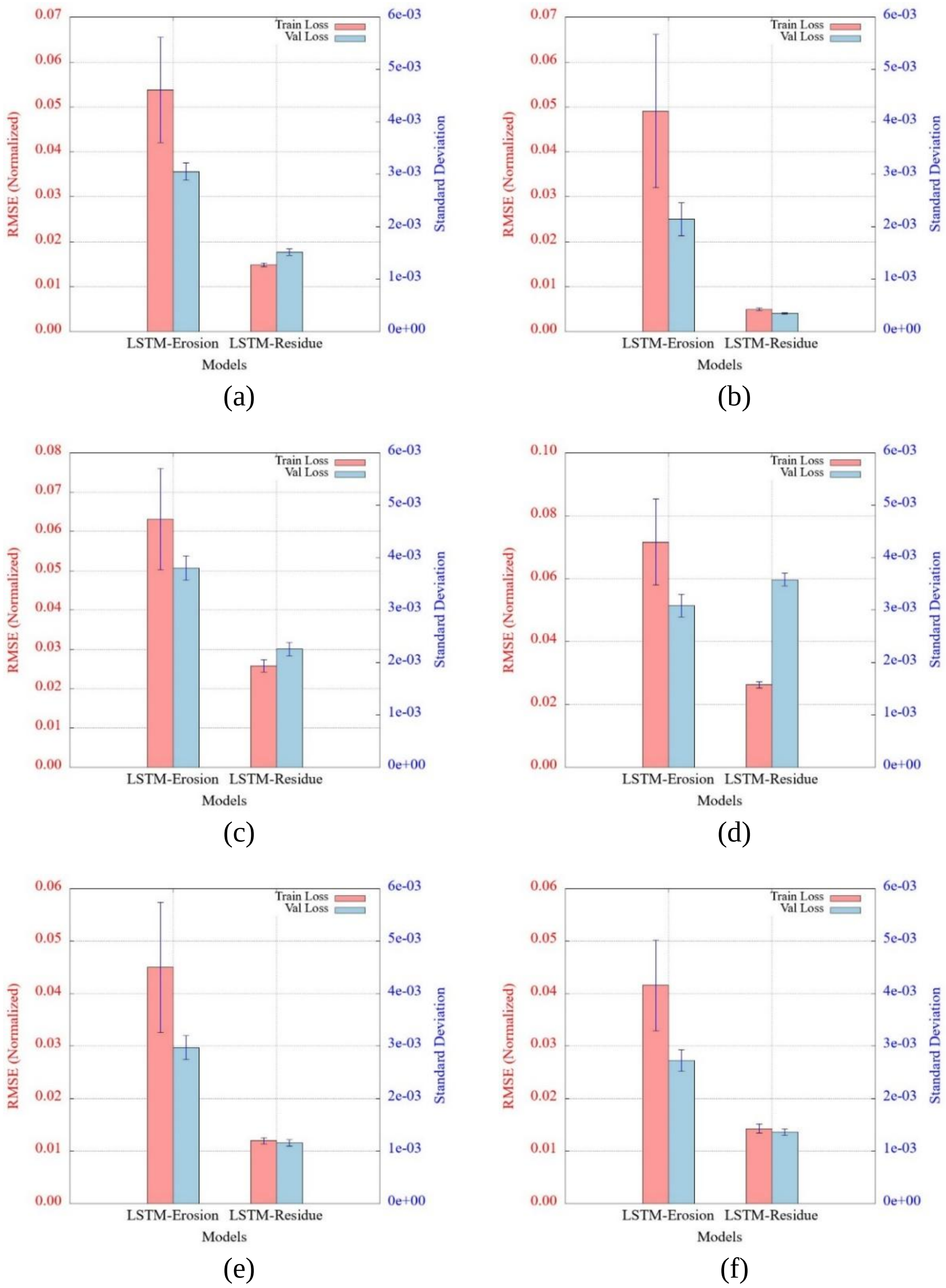

Figure 4. RMSE value of the LSTM Model on Valve Opening
(a) 20°, (b) 30°, (c) 45°, (d) 60°, (e) 75°, and (f) 90°.

In addition, we are also conducting the LSTM models performance test model using RMSE. The comparison between the LSTM model for erosion prediction and residue prediction can be seen on **Figure 4**. Using the same LSTM architecture, it can be seen that the RMSE form LSTM Residue model has a smaller value that the LSTM Erosion model. This shows that the combination of SARIMA and LSTM models to predict erosion can significantly improve the LSTM performance. Through residual data, LSTM learns complex patterns in erosion mass data that change over time. The learning process of LSTM aims to consider the interaction of other parameters such as particle size, particle density, valve material, valve structure, and fluid flow that essentially affect the erosion mass value at each time-step.

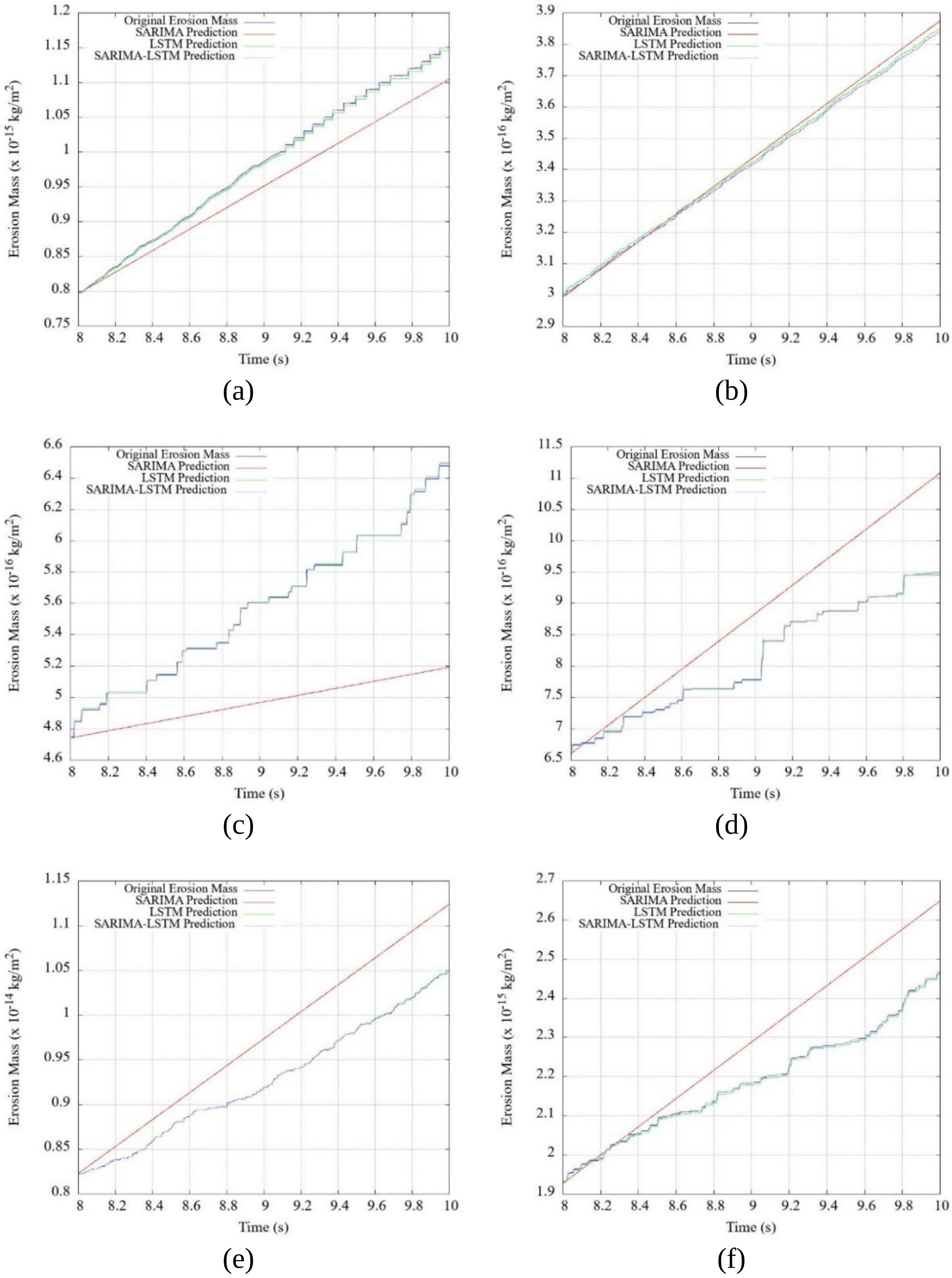

Figure 5. SARIMA-LSTM fitting results on erosion mass data with Valve Opening (a) 20°, (b) 30°, (c) 45°, (d) 60°, (e) 75°, and (f) 90°.

In this study, we also observe the SARIMA-LSTM capabilities in forecasting the erosion mass value for the next 30 time-steps. The results of the erosion mass forecast for each valve opening can be seen in **Figure 6.**

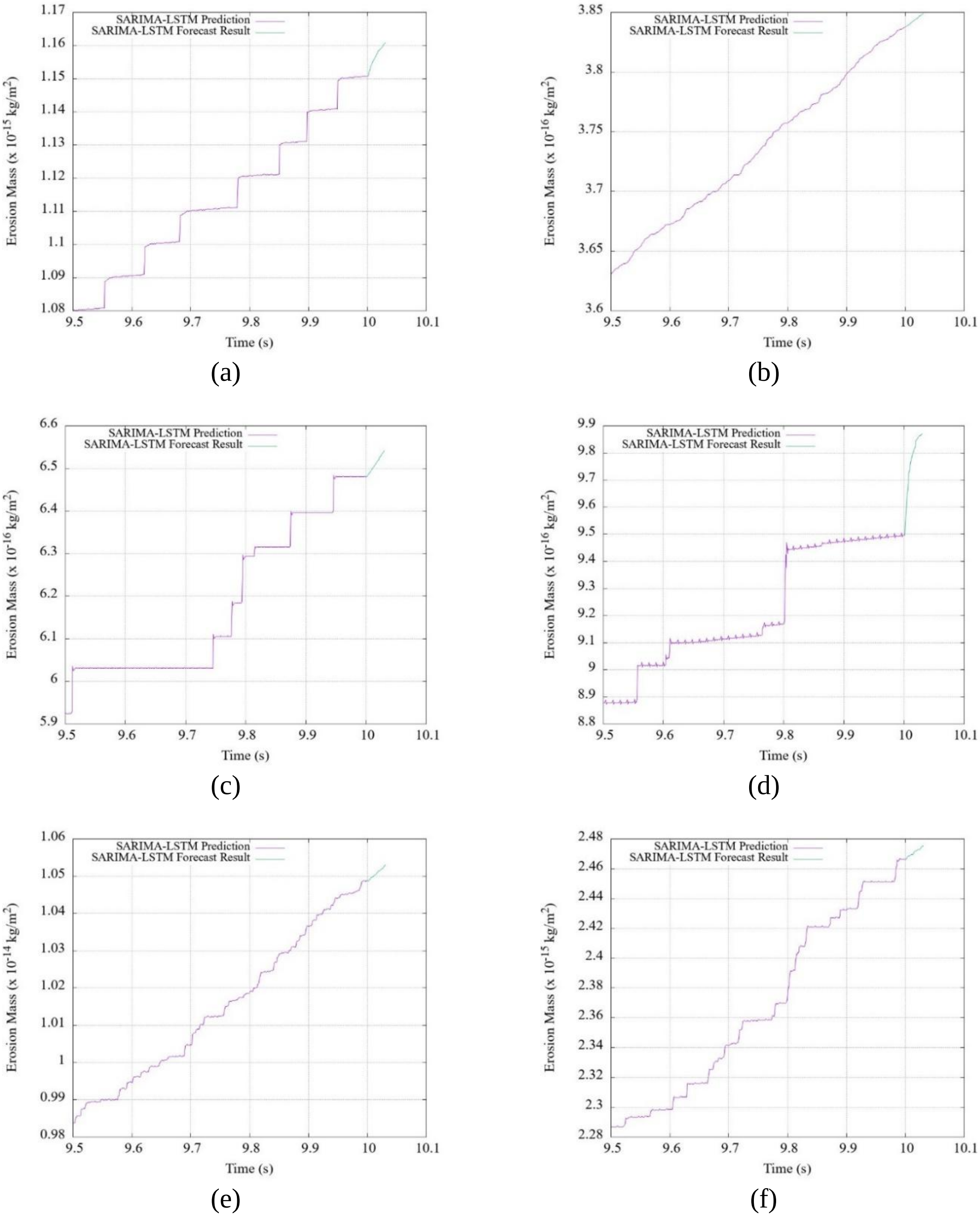

Figure 6. Hybrid SARIMA-LSTM erosion mass forecast results on valve opening (a) 20°, (b) 30°, (c) 45°, (d) 60°, (e) 75°, and (f) 90°.

Figure 6 shows the good ability of the SARIMA-LSTM model in forecasting erosion mass value. This is indicated by the forecast results which provide an erosion mass value that increase over time. In theory, the erosion mass will increase over time due to the effect of solid particle collisions that continuously hit the valve disc and are also influenced by various parameters related to fluid flow. The value of the forecasted erosion mass in each graph looks to have a pattern which is obtained from the ability of LSTM as an artificial intelligence model that is able to learn long-term interdependencies on data.

4. Conclusion

This study successfully implemented the hybrid SARIMA-LSTM model for predicting erosion mass on butterfly valves. The Hybrid SARIMA-LSTM model shows its advantages in predicting erosion mass compared to the SARIMA model and the LSTM model on each erosion mass data. The LSTM model combined with the SARIMA model proved to have a significant increase in performance in terms of RMSE value. The hybrid SARIMA-LSTM model also successfully forecast the erosion mass value up to 30 time-steps. These results provide an open possibility to apply the hybrid SARIMA-LSTM model in the oil and gas industry directly to observe erosion mass value over time. In future study, researchers can still investigate the performance of this model using real experimental data and conduct a more thorough validation of the forecast data generated by the hybrid SARIMA-LSTM model.

References

[1] M. Stewart, *Surface Production Operations*, 3rd ed. Virginia: Elsevier, 2016. doi: 10.1016/C2009-0-20127-9.

[2] Q. Liu, C. Xie, and B. Cheng, “Butterfly valve erosion prediction based on LSTM network,” *Flow Meas. Instrum.*, vol. 98, no. July, p. 102652, 2024, doi: 10.1016/j.flowmeasinst.2024.102652.

[3] S. Y. Jeon, J. Y. Yoon, and M. S. Shin, “Flow characteristics and performance evaluation of butterfly valves using numerical analysis,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 12, p. 012099, 2010, doi: 10.1088/1755-1315/12/1/012099.

[4] G. Zhang, R. Hu, D. Yin, D. Chen, H. Zhou, and Z. Lin, “Numerical Study of Cavitation Characteristics through Butterfly Valve under Different Regulation Conditions,” *Processes*, vol. 12, no. 4, 2024, doi: 10.3390/pr12040686.

[5] H. Devold, *Oil and gas production handbook: An introduction to oil and gas production, transport, refining and petrochemical industry*. Oslo: ABB Oil and Gas, 2013.

[6] D. Peng, S. Dong, Z. Wang, D. Wang, Y. Chen, and L. Zhang, “Characterization of the Solid Particle Erosion of the Sealing Surface Materials of a Ball Valve,” *Metals (Basel).*, vol. 11, no. 2, p. 263, Feb. 2021, doi: 10.3390/met11020263.

[7] M. Hettegger and D. Berezovsky, “Valve Leakage - How Acoustic Emission and Artificial Intelligence Can Reduce Losses and Increase Safety and Efficiency in Oil and Gas Plants,” in *Day 3 Wed, October 04, 2023*, SPE, Oct. 2023. doi: 10.2118/216596-MS.

[8] B. Liu, J. Zhao, and J. Qian, “Numerical study of solid particle erosion in butterfly valve,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 220, no. 1, pp. 8–13, 2017, doi: 10.1088/1757-899X/220/1/012018.

[9] Z. Lin, H. Yu, T. Yu, and Z. Zhu, “Numerical study of solid–liquid two-phase flow

and erosion in ball valves with different openings," *Adv. Powder Technol.*, vol. 33, no. 5, p. 103542, 2022, doi: 10.1016/j.apt.2022.103542.

[10] K. H. Hanon and E. Ebrahimi, "Numerical Study of the Fluid Flow and Erosion-Corrosion in an Industrial Valve," *Iran. J. Chem. Eng.*, vol. 20, no. 1, pp. 3–21, 2023, doi: 10.22034/ijche.2023.375055.1466.

[11] G. Zhang, H. Y. Zhang, J. L. Xin, D. S. Chen, and Z. Lin, "Numerical Study of Solid–liquid Two-phase Flows and Erosion in a Y-type Slurry Valve with Different Seat Structures," *J. Appl. Fluid Mech.*, vol. 17, no. 8, pp. 1657–1676, 2024, doi: 10.47176/jafm.17.8.2451.

[12] N. T. A. Othman, Z. A. Putra, S. A. Z. Abidin, and F. I. M. Ali, "Integration of CFD analysis and artificial neural networks for estimation of erosion rate in oil and gas pipelines," *Results Eng.*, vol. 24, p. 103288, Dec. 2024, doi: 10.1016/j.rineng.2024.103288.

[13] Z. Wang, H. Chen, M. Wang, X. Zhang, and Y. Dou, "Solid particle erosion prediction in elbows based on machine learning and swarm intelligence algorithm," *J. Pet. Sci. Eng.*, vol. 218, p. 111042, Nov. 2022, doi: 10.1016/j.petrol.2022.111042.

[14] M. M. Ridwan, "Perbandingan Hybrid ARIMA-LSTM dan ARIMA-GRU Pada Prediksi Harga Bitcoin," Universitas Islam Negeri Syarif Hidayatullah, 2023.

[15] G. Y. Lee, L. Alzamil, B. Doskenov, and A. Termehchy, "A Survey on Data Cleaning Methods for Improved Machine Learning Model Performance," Sep. 2021, [Online]. Available: http://arxiv.org/abs/2109.07127

[16] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," *Appl. Soft Comput.*, vol. 97, p. 105524, Dec. 2020, doi: 10.1016/j.asoc.2019.105524.

[17] X. Ying, "An Overview of Overfitting and its Solutions," *J. Phys. Conf. Ser.*, vol. 1168, p. 022022, Feb. 2019, doi: 10.1088/1742-6596/1168/2/022022.

[18] W. Hariadi and S. Sulantari, "Aplication of ARIMA Model for Forecasting Additional Positive Cases of Covid-19 in Jember Regency," *Enthusiastic Int. J. Appl. Stat. Data Sci.*, vol. 1, no. 01, pp. 20–27, Apr. 2021, doi: 10.20885/enthusiastic.vol1.iss1.art4.

[19] L. Xu and G. Dong, "Prediction Model of International Trade Risk Based on Stochastic Time-Series Neural Network," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–9, Jun. 2022, doi: 10.1155/2022/3119535.

[20] W. S. Harianto, "PENGGUNAAN METODE ARIMA (AUTOREGRESSIVE INTEGRATED MOVING AVERAGE) UNTUK PRAKIRAAN JUMLAH PERMINTAAN GULA RAFINASI (Studi Kasus: PT. Makassar Tene)," Universitas Islam Negeri Alauddin, 2018.

[21] A. A. Aminullah, M. Idhom, and W. S. J. Saputra, "FORECASTING SALES USING SARIMA MODELS AT THE SINAR PAGI BUILDING MATERIALS STORE," *JIKO (Jurnal Inform. dan Komputer)*, vol. 7, no. 2, pp. 71–78, Jul. 2024, doi: 10.33387/jiko.v7i2.8266.

[22] J. E. Chrisnanto *et al.*, "Bi-LSTM Approach for Seagrass Transplantation Site

Detection," in *2024 14th International Conference on System Engineering and Technology (ICSET)*, IEEE, Oct. 2024, pp. 38–43. doi: 10.1109/ICSET63729.2024.10774941.

[23] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive into Deep Learning*. Cambridge University Press, 2021. [Online]. Available: https://d2l.ai

[24] R. J. Hyndman and G. Anthanasopoulos, *Forecasting: Principles and Practice*, 2nd ed. Melbourne: OTexts, 2018. [Online]. Available: otexts.com/fpp2