IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 3, June 2026, pp. 2633~2642
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i3.pp2633-2642

Forecasting world sugar contract futures using long short-term memory technique with multi-step ahead forecasting strategy

Khairil Anwar Notodiputro[1], Kayla Fakhriyya Jasmine[1], Indahwati[1], Wandee Wanishsakpong[2]
[1]School of Data Science, Mathematics, and Informatics, IPB University, Bogor, Indonesia
[2]Department of Statistics, Faculty of Sciences, Kasetsart University, Bangkok, Thailand

Article Info

Article history:

Received Jan 24, 2025
Revised Feb 9, 2026
Accepted Apr 20, 2026

Keywords:

Direct
Long short-term memory
Multi-input multi-output
Recursive
Sugar

ABSTRACT

Time series analysis using stochastic and dynamic models for data forecasting is a key in assisting planning and decision-making processes in various sectors. Long short-term memory (LSTM), with its advantage in understanding patterns and non-linearity in sequential data, is applied in a multi-step ahead forecasting strategy on world sugar futures prices. Fluctuations in sugar prices have a significant impact on the agriculture, trade, and food industry sectors. Forecasting sugar prices becomes a crucial tool for industries, investors, and traders to anticipate changes and make informed decisions. The objectives of this study are to identify the best strategy for forecasting the world sugar contract price and to perform forecasting using the best model. The research results indicate that hyperparameter tuning in LSTM models produces varied combinations and effects. Furthermore, the recursive strategy is suitable for long-term forecasting, while the direct strategy is appropriate for short-term forecasting. Forecasting values for long-term periods remains challenging in achieving high accuracy.

Corresponding Author:

Wandee Wanishsakpong
Department of Statistics, Faculty of Sciences, Kasetsart University
50 Ngamwongwan Rd, Lat Yao, Chatuchak, Bangkok, Thailand
Email: wandee.w@ku.th

1. INTRODUCTION

Time series data are formed by a series of observations arranged in accordance with order of events. Analysis of time series data focuses on developing stochastic and dynamic models for data dependence, applying them to various important areas [1]. One of the capabilities of time series analysis is forecasting, which allows us to predict future values based on past observations. Montgomery *et al.* [2] state that this forecasting is crucial as it provides a future perspective, aiding in the planning and decision-making process.

There are various methods that can be used for forecasting, ranging from classical techniques such as autoregressive (AR), moving average (MA), autoregressive integrated moving average (ARIMA) to deep learning models like long short-term memory (LSTM). LSTM, as a deep learning model, is effective in capturing and understanding patterns from sequential data. Furthermore, LSTM is a basic nonlinear model for time series analysis that can comprehend the non-linear nature of data [3]. A real-world application of LSTM is in forecasting global sugar futures prices using a multi-step ahead forecasting strategy. This strategy can predict values for several future periods with various types of strategies, namely recursive, direct, and multi-input multi-output (MIMO).

Historically, sugar has been one of the most volatile commodities [4]. Fluctuations in sugar prices impact various sectors, including agriculture, trade, and industry, particularly the food and beverage industry. According to Hanani *et al.* [5], an increase in sugar prices will prompt sugarcane farmers to expand the area of sugarcane plantations, causing a reduction in the production of other crops and agricultural products. Moreover, easier market access can lead to a decrease in domestic sugar production and an increase in imports, triggering high export levels from exporting countries such as Brazil and ultimately raising global sugar prices [6]. The rise in sugar prices will also increase production costs, causing the food and beverage industry to reduce production levels [5]. Futures contracts serve as crucial instruments used by traders and investors to manage and hedge against various risks in the financial markets [7]. By applying LSTM with a multi-step ahead forecasting strategy, forecasting sugar futures prices is expected to become an important tool for market participants to anticipate changes and make informed decisions.

Research related to sugar price forecasting has been conducted by Fauziah and Gunaryati [8], comparing the double exponential smoothing (DES) and artificial neural network (ANN) models in forecasting daily sugar prices in Depok. The results of this study indicate that ANN has better performance. Another study conducted by Yurtsever [9] compared deep learning models, namely LSTM, bidirectional long short-term memory (Bi-LSTM), and gate recurrent unit (GRU), finding that the LSTM model performed best with root mean square error (RMSE) values of 61.728, mean absolute error (MAE) of 48.85, and mean absolute percentage error (MAPE) of 3.48%. This study also showed that economic indicators influence gold prices. In addition to research comparing models, there is also research by Ferreira and Cunha [10] comparing recursive, direct, and MIMO strategies in forecasting daily reference evapotranspiration using deep learning. The results indicate that the MIMO strategy is the best strategy for obtaining the most accurate forecasts. The objectives of this research are to identify the hyperparameter combinations that yield the best modeling performance, identify the multi-step ahead forecasting strategy that provides the best forecasts for world sugar futures prices, and forecast world sugar futures prices using the best model.

2. METHOD

2.1. Data

The data used in this research consists of daily prices from one of the world sugar futures contracts, specifically the ICE NY sugar #11 futures (SBc3). This contact serves as a primary benchmark for raw sugar derived from sugarcane and enables trading at a predetermined future price. The “c3” in SBc3 signifies that it is the third contract available in sequence for future delivery. Data was sourced from Investing.com at http://www.investing.com/commodities/us-sugar-no11-hidtorical-data?cid=1186965, covering the observation period from January 2, 2019, to December 29, 2023, with a total of 1,258 observations. The unit used in the prices of this future contract is cents per pound.

2.2. Long short-term memory

LSTM is a deep learning method developed from recurrent neural networks (RNN) capable of learning long-term dependencies and retaining information over extended periods [11]. One of the primary issues that LSTM addresses is the vanishing gradient problem in RNNs. This problem makes it difficult for RNNs to learn long-term relationships in data because the gradients of the error function tend to become very small, causing the learning process to halt [12]. LSTM has a structure that can selectively remember or forget information through its cell state and three gates [11]. The structure of an LSTM consists of four main components that interact at each time step:

i) Cell state (C_t): the cell state is the main component in an LSTM, functioning as long-term memory [13]. The process for the cell state resembles a conveyor belt, where parameter information moves linearly with interactions such as multiplication and addition. The status of the information depends on these interactions, and if there is no interaction, the information remains unchanged [14].

ii) Forget gate (f_t): the forget gate determines which information should be discarded or retained from the cell state. The sigmoid activation function (σ) is used at this stage, producing values between 0 and 1. A value of 0 means the information will be discarded, while a value of 1 means the information will be retained [14].

iii) Input gate (i_t): the input gate controls the information that will be added to the cell state from the latest input value (x_t) and protects memory from irrelevant input. There are two layers that determine the new memory stored in the cell state: the sigmoid layer and the tanh layer. The sigmoid layer decides the updated values, while the tanh layer forms a vector of new candidate values for the cell state ($\tilde{C}_t$) [14].

iv) Output gate (o_t): the output gate controls the information contained in the updated cell state (C_t) that will be sent to the output at each time step. This gate determines the value of the next hidden state based on the previous hidden state (h_{t-1}), the current input, and the newly updated cell state [14]. This optimization method combines techniques from other optimization methods such as adaptive gradient (AdaGrad), which works well with sparse gradients and root mean square propagation (RMSprop) [15].

2.3. Multi-step ahead time series forecasting strategy

Multi-step ahead time series forecasting aims to predict the values for the next H time periods $[y_n, \ldots, y_{n+H}]$ based on existing historical time series data [16]. The following are strategies applied to obtain forecasts for several future periods.

2.3.1. Recursive strategy

The recursive strategy is the simplest approach to predicting values for multiple future time periods [16]. The principle of this strategy is to use the forecasted result as an input to generate the prediction for the next day [16].

$$\hat{y}_{n+h} = \begin{cases} \hat{f}(y_n, \ldots, y_{n-d+1}), if\ h = 1 \\ \hat{f}(\hat{y}_{n+h-1}, \ldots, \hat{y}_{n+1}, \hat{y}_n, \ldots, y_{n-d+1}),\ if\ h \in \{2, \ldots, d\} \\ \hat{f}(\hat{y}_{n+h-1}, \ldots, \hat{y}_{n-d+1}), if\ h \in \{d+1, \ldots, H\} \end{cases} \tag{1}$$

Where n is index of the current time period, h is predicted time period, d is number of previous time periods used as input to make the prediction, H is number of predicted time periods, and $\hat{f}(y)$ is model used for forecasting.

2.3.2. Direct strategy

This strategy does not use previous forecast results as inputs. Instead, each forecast result has its own model [16]. As a result, the error in one forecast value does not accumulate into the forecast value for the subsequent period [17].

$$\hat{y}_{n+h} = \hat{f}_h(y_n, \ldots, y_{n-d+1}) \tag{2}$$

Where n is index of the current time period, h is predicted time period, d is number of previous time periods used as input to make the prediction, and $\widehat{f_h}(y)$ is the model used to forecast the h-th period.

2.3.3. Multi-input multi-output

This strategy generates forecast values for several future periods simultaneously by using a model that takes into account several inputs at once [18].

$$[\hat{y}_{t+H}, \ldots, \hat{y}_{t+1}] = \hat{F}(y_n, \ldots, y_{n-d+1}) \tag{3}$$

Where n is index of the current time period, d is number of previous time periods used as input to make the prediction, H is number of predicted time periods, and $\hat{F}(y)$ is model used for forecasting.

2.4. Analysis procedure

Data analysis was conducted using Python software. Python is a popular software for statistics and data science research. Using the software and the theoretical framework the procedural steps for the analysis in this research are as follows:

i) Data exploration to determine the data characteristics, trends, seasonal patterns, trend change patterns over 5 years, and stationarity data.
ii) Splitting the data into training data (80%) and test data (20%).
iii) Normalizing the data using the min-max normalization method. Normalization is necessary even for univariate data because according to Li and Cao [19], LSTM using sigmoid and tanh as activation functions are sensitive to the input data range.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4}$$

iv) Adjusting the data dimension shape according to the approach used.
   - Single-output: the data dimension for the LSTM model in this approach will be arranged in the form of (number of samples, number of time steps, number of variables). The recursive and direct strategies will use models from LSTM training results with a single-output approach.
   - Multi-output: the data dimension for the LSTM model in this strategy will be arranged in the form of (number of samples, number of time steps, and number of input variables) for input and (number of samples, number of future time steps, and number of output variables) for output. The MIMO strategy will use models from LSTM training results with a multi-output approach.

v) Modeling with LSTM.
   – Creating k-fold time series cross-validation (CV) scenarios by dividing the training data into training and validation data for each fold while still considering the time order.
   – Applying the LSTM method to each fold for each combination of hyperparameters used as shown in Table 1.
   – Training the LSTM model with the initialized hyperparameter combinations on the 1st to the k-th fold.
   – Calculating the average RMSE value from the validation data for each combination of hyperparameter-trained models.
   – Repeating steps c and d with different hyperparameter combinations.
   – Selecting the combination of hyperparameters with the smallest average.
   – Training the LSTM model using all training data with the best hyperparameter combination.

vi) Repeating step 5 as many times as the length of data to be forecasted for the single-output approach.
vii) Denormalizing the data to return to the original data scale.
viii) Evaluating the model on the test data by applying the multi-step ahead forecasting strategy.
ix) Selecting the best strategy in forecasting sugar futures prices using RMSE, MAE, and MAPE metrics.
x) Selecting the strategy with the best model metric value.
xi) Forecasting world sugar futures contract prices with the best model and interpreting the results.

Table 1. Hyperparameter combination

Hyperparameter			
Hidden layer	Learning rate	Neuron	Time step
1, 2	0.001; 0.01	32, 64, 128	5, 22, 66, 125

According to Shcherbakov *et al.* [20], RMSE and MAE are error measures that can be used to evaluate the accuracy of forecasting models for time series data. Then, MAPE is also one of the commonly used error measures because it is not dependent on the data scale and easy to interpret [21].

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - f_t)^2} \tag{5}$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - f_t| \tag{6}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|y_t - f_t|}{yt} \times 100 \tag{7}$$

Where n is length of data, y_t is actual value, and f_t is predicted value.

3. RESULTS AND DISCUSSION

3.1. Data exploration

Over the course of five years, the price of sugar futures contracts rose by 66%, averaging 16.82 cents/pound. Sugar reached its lowest price, 9.67 cents/pound, on April 27, 2020, and its highest price, 26.26 cents/pound, on September 19, 2023. Figure 1 shows three significant timelines during this period that caused changes in the sugar price pattern. The year 2019 was before the COVID-19 pandemic, during which sugar prices fluctuated with small price changes and tended to be relatively stable at an average price of 13.01 cents/pound. Then, from 2020 to 2021, during the COVID-19 pandemic worldwide, the sugar price pattern changed significantly compared to the previous year. During the pandemic, almost the entire world implemented lockdowns, causing economic activities to halt temporarily. This led to a drop in commodity prices, including sugar.

Recovery from the COVID-19 pandemic occurred from 2022 to 2023, with sugar futures contract prices in 2022 showing a pattern quite similar to 2019. A different pattern emerged again in 2023 with significant price increases and decreases. High prices occurred because exporting countries like Thailand and India faced climate problems leading to reduced production. The Indian government also imposed export restrictions, further decreasing sugar export stocks, and there was a shift in sugarcane production to ethanol. Then, the decrease in sugar prices in December 2023 occurred because Brazil increased sugar production, and India also focused more on sugar production again.

According to Nurhambali *et al.* [22], data that are not spread around the mean indicate non-stationarity in the mean, while data that widens and narrows unevenly indicates non-stationarity in variance. The sugar futures price data shown in Figure 1 exhibit these characteristics. The results of the augmented Dickey-Fuller (ADF) test show a p-value of 0.346, which can be interpreted to mean that the sugar price data are non-stationary in the mean. Meanwhile, the lambda value is -0.101 with an upper bound of 0.02 and a lower bound of -0.263 within a 95% confidence interval, indicating that the sugar price data is also non-stationary in variance because the range does not include the value one. However, the data does not need to be treated with differencing or transformation because LSTM is a neural network method capable of handling non-stationary data.

Figure 1. Sugar contract futures price line chart

3.2. Long short-term memory modeling

In LSTM modeling, the model is trained by considering various combinations of hyperparameters to find the best model that can generate forecasts of world sugar futures contract prices with the smallest error [23]. According to Wang *et al.* [24], the initialization of hyperparameters has a significant impact on the complexity and accuracy of the model. The LSTM model is built with the initialization of hyperparameters consisting of hidden layers, learning rate, neuron units, and time steps. Selecting four hyperparameters is done to obtain the best model more effectively in terms of time.

The number of hidden layers and the number of neuron units in each hidden layer are considered as hyperparameters set during initialization because they can affect the complexity and depth of the network architecture. Learning rate is an important hyperparameter that can control the effective capacity of the model in a more complex way than other hyperparameters [25]. The time step can determine the length of the previous time period used as input to predict future sugar futures contract prices, thus affecting the model's complexity.

After determining the initialization of hyperparameters, CV is performed with k-fold specific to time series data [26]. CV is conducted to measure model performance, obtain the most optimal and consistent combination of hyperparameters, and detect overfitting. The value of k used is 6 with sequential time-based data splitting as shown in Figure 2. The choice of k =6 was made to ensure that each fold covers approximately 5-6 months of data for both training and validation. In the first CV iteration, the model is trained on about 5-6 months of data and validated on the following 5-6 months. In each subsequent iteration, the training data expands by an additional 5-6 months, while the validation data shifts forward by the same amount.

The average RMSE values are calculated for each fold in each combination of hyperparameters. The use of metrics provides a better understanding of how well the model can forecast sugar futures contract prices. Generally, each hyperparameter has a different impact on model performance. The optimal hyperparameters are determined based on the smallest average RMSE metric value on the validation data for each fold from the initialization combination results. The average RMSE values in Table 2 are numbers on a normalized scale. The results show that the single-output at H equal to one has the lowest average RMSE because the model is trained to obtain forecasts for the next period. This model is applied to a recursive strategy and direct strategy for the forecast of the first day. Furthermore, four single-output LSTM training results are displayed for the tuning results of forecast periods on the 5th, 22nd, 66th, and 125th days. The further the forecast period, the more difficult it is for the model to read the pattern, resulting in an increasing average RMSE value as the H value increases. Then, there is a multi-output approach with the highest average RMSE. This model is applied to the MIMO strategy. The high RMSE can occur because the model is trained to produce output for the entire desired period at once.

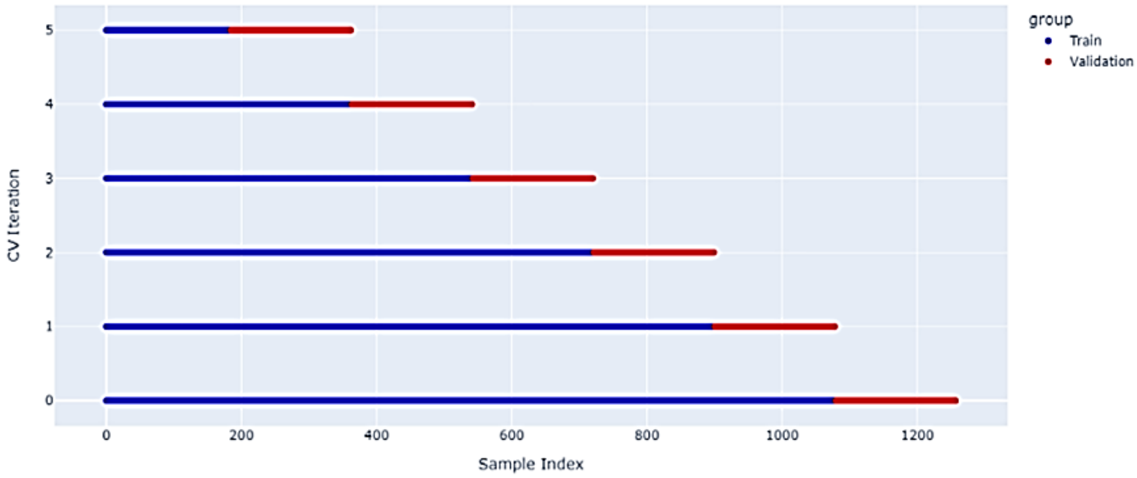

Figure 2. K-fold time series cross validation scheme

Table 2. Optimal hyperparameter combinations for each approach

Hyperparameter optimal	Approach					
	Single-output (H =1)	Single-output (H =5)	Single-output (H =22)	Single-output (H =66)	Single-output (H =125)	Multi-output
Hidden layer	2	2	2	2	2	2
Unit neuron	128	32	32	128	64	64
Learning rate	0.01	0.01	0.001	0.01	0.01	0.001
Time step	5	66	5	22	66	66
Rata-rata RMSE	0.039	0.059	0.108	0.176	0.195	0.21

3.3. Evaluation and determination of the best strategy

The previously obtained models from hyperparameter tuning are applied to the test data to assess how well the models forecast the test data. There are 252 observations in the test data with a period length of 1 year, from December 29, 2022, to December 29, 2023. The three forecasting strategies applied—recursive, direct, and MIMO—yield different forecasting patterns, as shown in Figure 3.

The forecasting pattern generated by the recursive strategy shows an increasing pattern up to the 8th month, followed by relatively flat forecast values. The flat condition in the later phase of forecasting indicates that the model has limitations in adapting to and responding to significant changes in the test data. This can occur because in the recursive strategy, errors from previous predictions accumulate and are used as input for subsequent predictions, potentially magnifying forecast errors over time.

The direct strategy produces predictions that are closer to the actual data, showing a strong ability to follow price fluctuations, especially in the short term. However, there are periods where this strategy fails to capture price changes accurately, especially during sharp price spikes or drops. This is due to the model's limitations in learning complex patterns of price changes that depend not only on the last value but also on historical patterns that may not be fully represented in the training process.

The MIMO strategy provides results farthest from the test data, with prediction lines relatively flat and not reflecting actual price volatility. The high complexity in the MIMO strategy causes the model to learn dependencies between various time steps simultaneously, making it difficult to capture constantly changing patterns. Furthermore, metrics are calculated using RMSE, MAE, and MAPE. These metrics provide information about the accuracy and average error of the forecasted values generated by the model.

The performance of each forecasting strategy tested in Table 3 shows that the best metric results are not consistently shown in one strategy over different time periods. In the short term, the RMSE, MAE, and MAPE values obtained from all three strategies can be considered quite good, with the direct strategy performing best in forecasting sugar prices up to 1 month. This indicates that the direct strategy has the best accuracy in dealing with short-term price fluctuations. Meanwhile, in long-term forecasting, from 3 months to 1 year, the recursive strategy becomes the best strategy due to its ability to have a stable pattern or slow-moving trend. Overall, the longer the forecast period applied to a strategy, the higher the resulting metric values. This phenomenon can occur because the model cannot accurately capture events that occur over long periods.

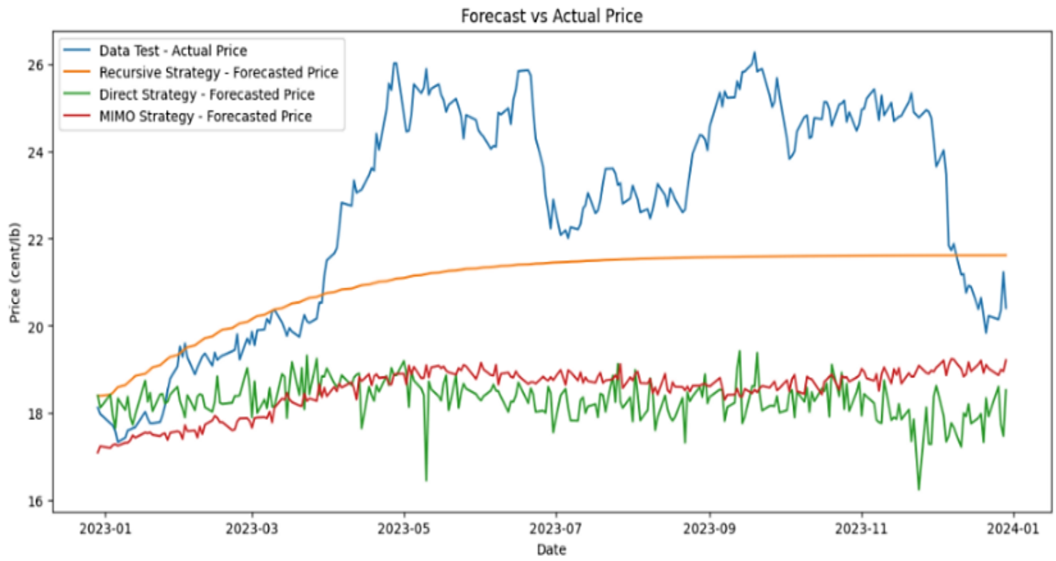

Figure 3. Comparison of test data and forecasting results for recursive, direct, and MIMO strategies

Table 3. Determination of the best strategy

Forecast period	Strategy								
	Recursive			Direct			MIMO		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
1 day	0.280	0.280	1.517	0.280	0.280	1.517	1.027	1.027	5.663
1 week	0.634	0.610	3.430	0.324	0.244	1.899	0.664	0.614	3.434
1 month	0.742	0.670	5.455	0.545	0.448	2.979	0.753	0.577	3.209
3 months	0.695	0.603	6.012	1.300	1.086	6.116	1.632	1.440	7.819
6 months	2.428	1.888	12.145	4.324	3.430	14.846	4.124	3.438	15.382
1 year	2.537	2.129	11.942	5.086	4.489	18.895	4.704	4.190	18.055

Throughout the training to evaluation process, the strengths and weaknesses of each strategy can be identified. Recursive, as the best strategy for long-term price forecasting, only uses a single-output model, thus requiring less time in the model training and hyperparameter tuning process. With a simple scheme, errors in the forecasting results accumulate as the forecast period increases.

The weakness of recursive in terms of error accumulation is not experienced by direct and MIMO. This can occur in the direct strategy because the model is built for as many periods as desired for forecasting. However, the main drawback is the extensive time required in the model training and hyperparameter tuning process. Additionally, the direct strategy also has limitations in capturing complex temporal dependencies, especially if the patterns and relationships between variables in the data are dynamic or change over time.

Another strategy, MIMO, only uses one model that maintains temporal dependencies and does not require much time in the training and hyperparameter tuning process. However, MIMO also has disadvantages such as a lack of flexibility and high model complexity. This causes MIMO to potentially be ineffective in adapting to rapid and complex changes in data dynamics.

3.4. Forecasting results

The data used represents the sugar futures contract price for July 2024, meaning the contract price is forecasted until the last working day before July, which is June 28, 2024. There are two strategies, recursive and direct, applied to forecast sugar prices with different forecast time periods. The forecast results in Figure 4 show a consistent decrease in sugar prices from January to June 2024 according to the recursive strategy. Companies can utilize this information to plan sugar purchases during this period at lower prices. This will help them optimize spending and reduce production costs, thereby increasing profits. Meanwhile, the price forecast shown by the direct strategy has more fluctuating results with potential short-term market volatility. The forecast results generated by the direct strategy are more suitable for traders seeking opportunities from short-term price movements for speculation or hedging, taking advantage of market volatility for short-term gains by buying when prices are low and selling when prices rise.

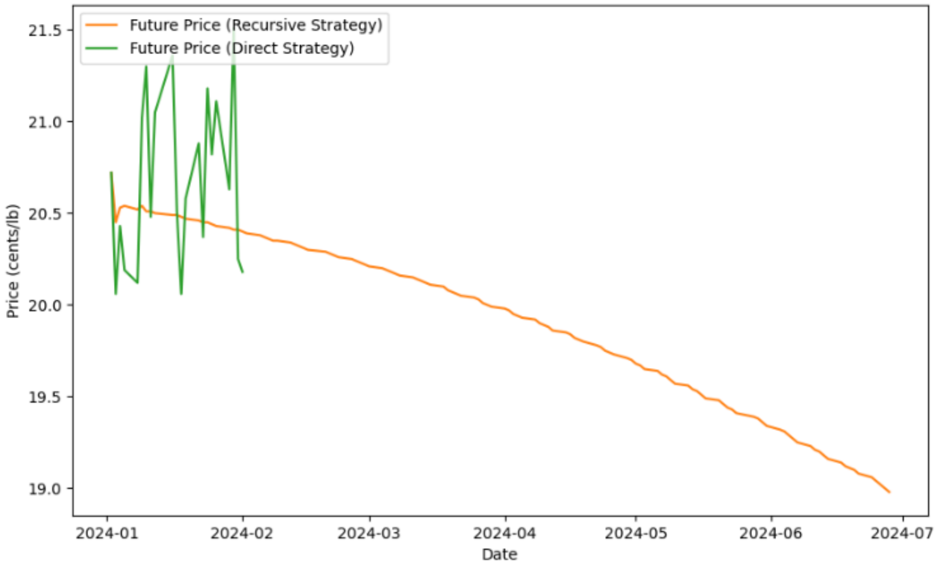

Figure 4. Sugar futures contract price forecast results

4. CONCLUSION

Based on the research conducted on sugar futures contract prices, the hyperparameter tuning process has various effects on each multi-step ahead forecasting strategy. The strategy that can be used to forecast sugar prices in the long term, ranging from 3 months to 1 year, is recursive. The forecasted prices have a continuous downward trend, so market participants are recommended to plan sugar purchases to optimize company profits. Meanwhile, short-term sugar price forecasts are best done by a direct strategy. Fluctuating prices can be considered by traders to take advantage of short-term trading opportunities that arise from price volatility, allowing them to make quick profits by capitalizing on rapid price changes within shorter forecasting horizons. Forecasting sugar futures contract prices for several periods ahead remains a challenge in achieving high accuracy. Therefore, the forecast values and patterns generated can still be considered with attention to market conditions and other external factors. This study focuses on the application of LSTM method using univariate data to forecast values several periods ahead by comparing the performance of three strategies. Future research can apply other modeling methods that can capture patterns undergoing significant changes. Additionally, adding other strategies such as direct-recursive (DirRec) and direct multi-output (DirMO) as efforts to obtain the most accurate forecasts. Moreover, multivariate data using external factors, such as political conditions, government policies, and climate change affecting sugar prices, can also be used, so that the forecast values are expected to yield better results.

ACKNOWLEDGMENTS

The authors would like to thank International SciKU Branding (ISB), Faculty of Science, Kasetsart University, Thailand for supporting this research.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Khairil Anwar Notodiputro	✓	✓	✓	✓	✓	✓		✓	✓	✓		✓		
Kayla Fakhriyya Jasmine		✓				✓		✓	✓		✓	✓		
Indahwati	✓		✓	✓			✓			✓	✓		✓	
Wandee Wanishsakpong	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available at https://www.investing.com/commodities/us-sugar-no11-contracts.

REFERENCES

[1] G. E. P. Box, G. C. Reinsel, G. M. Jenkins, and G. M. Ljung, *Time series analysis forecasting and control*. Hoboken, United States: Wiley, 2017, doi: 10.1002/wics.1391.

[2] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to time series analysis and forecasting*, 2nd ed. Hoboken, United States: Wiley, 2008, doi: 10.1080/00224065.2008.11917751.

[3] R. Sharma *et al.*, "A review of weather forecasting using LSTM model," in *AIP Conference Proceedings*, 2023, doi: 10.1063/5.0152493.

[4] L. N. Pop, M. Rovinaru, and F. Rovinaru, "The challenges of sugar market: an assessment from the price volatility perspective and its implications for Romania," *Procedia Economics and Finance*, vol. 5, pp. 605–614, 2013, doi: 10.1016/s2212-5671(13)00071-3.

[5] N. Hanani, R. Anindita, A. Q. Pudjiastuti, and D. Kaluge, "Effects of sugar price increase in Indonesia," *Studia Universitatis Babes Bolyai - Oeconomica*, no. 1, pp. 28–39, 2013.

[6] A. Elobeid, *How would a trade deal on sugar affect exporting and importing countries?*, Geneva, Switzerland: International Centre for Trade and Sustainable Development, 2009.

[7] M. A. Ehsan, "Futures contracts in Islamic finance: an analytical approach," *Global Review of Islamic Economics and Business*, vol. 1, no. 1, 2015, doi: 10.14421/grieb.2013.011-03.

[8] F. N. Fauziah and A. Gunaryati, "Comparison forecasting with double exponential smoothing and artificial neural network to predict the price of sugar," *International Journal of Simulation: Systems, Science and Technology*, vol. 18, no. 4, 2017, doi: 10.5013/IJSSST.a.18.04.13.

[9] M. Yurtsever, "Gold price forecasting using LSTM, Bi-LSTM and GRU," *European Journal of Science and Technology*, 2021, doi: 10.31590/ejosat.959405.

[10] L. B. Ferreira and F. F. D. Cunha, "New approach to estimate daily reference evapotranspiration based on hourly temperature and relative humidity using machine learning and deep learning," *Agricultural Water Management*, vol. 234, 2020, doi: 10.1016/j.agwat.2020.106113.

[11] X. H. Le, H. V. Ho, G. Lee, and S. Jung, "Application of long short-term memory (LSTM) neural network for flood forecasting," *Water*, vol. 11, no. 7, 2019, doi: 10.3390/w11071387.

[12] S. Basodi, C. Ji, H. Zhang, and Y. Pan, "Gradient amplification: an efficient way to train deep neural networks," *Big Data Mining and Analytics*, vol. 3, no. 3, pp. 196–207, 2020, doi: 10.26599/BDMA.2020.9020004.

[13] D. Peter, A. H. Alavi, B. Javadi, and S. L. Fernandes, *The cognitive approach in cloud computing and internet of things technologies for surveillance tracking systems*. Cambridge, United Kingdom: Academic Press, 2020, doi: 10.1016/C2018-0-00368-0.

[14] H. Okut, "Deep learning for subtyping and prediction of diseases: long-short term memory," in *Deep Learning Applications*, London, United Kingdom: IntechOpen, 2021, doi: 10.5772/intechopen.96180.

[15] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *3rd International Conference for Learning Representations*, 2015.

[16] S. B. Taieb, G. Bontempi, A. F. Atiya, and A. Sorjamaa, "A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7067–7083, 2012, doi: 10.1016/j.eswa.2012.01.039.

[17] A. Sorjamaa, J. Hao, N. Reyhani, Y. Ji, and A. Lendasse, "Methodology for long-term prediction of time series," *Neurocomputing*, vol. 70, no. 16–18, pp. 2861–2869, 2007, doi: 10.1016/j.neucom.2006.06.015.

[18] S. B. Taieb, G. Bontempi, A. Sorjamaa, and A. Lendasse, "Long-term prediction of time series by combining direct and MIMO strategies," in *2009 International Joint Conference on Neural Networks*, 2009, pp. 3054–3061, doi: 10.1109/IJCNN.2009.5178802.

[19] Y. Li and H. Cao, "Prediction for tourism flow based on LSTM neural network," *Procedia Computer Science*, vol. 129, pp. 277–283, 2018, doi: 10.1016/j.procs.2018.03.076.
[20] M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, and V. A. E. Kamaev, "A survey of forecast error measures," *World Applied Sciences Journal*, vol. 24, no. 24, pp. 171–176, 2013, doi: 10.5829/idosi.wasj.2013.24.itmies.80032.
[21] S. Kim and H. Kim, "A new metric of absolute percentage error for intermittent demand forecasts," *International Journal of Forecasting*, vol. 32, no. 3, pp. 669–679, 2016, doi: 10.1016/j.ijforecast.2015.12.003.
[22] M. R. Nurhambali, Y. Angraini, and A. Fitrianto, "Implementation of long short-term memory for gold prices forecasting," *Malaysian Journal of Mathematical Sciences*, vol. 18, no. 2, pp. 399–422, 2024, doi: 10.47836/mjms.18.2.11.
[23] E. Bartz, T. B. Beielstein, M. Zaefferer, and O. Mersmann, *Hyperparameter tuning for machine and deep learning with R: a practical guide*. Singapore: Springer Nature, 2022, doi: 10.1007/978-981-19-5170-1.
[24] S. Wang, C. Ma, Y. Xu, J. Wang, and W. Wu, "A hyperparameter optimization algorithm for the LSTM temperature prediction model in data center," *Scientific Programming*, vol. 2022, no. 1, 2022, doi: 10.1155/2022/6519909.
[25] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. Cambridge, United Kingdom: MIT Press, 2016.
[26] D. Berrar, "Cross-validation," in *Encyclopedia of Bioinformatics and Computational Biology*, Amsterdam, Netherland: Elsevier, 2019, pp. 542–545, doi: 10.1016/B978-0-12-809633-8.20349-X.

BIOGRAPHIES OF AUTHORS

Khairil Anwar Notodiputro earned B.S. and master's degrees from IPB University and Ph.D. in Statistics from Macquarie University, Australia. He is a professor of Statistics and Data Science at IPB University. He is a member of the International Statistical Institute and a former chairman of the Indonesia Statistical Association. His research interests include mixed models, small area estimation, time series analysis, and statistical machine learning. He can be contacted at email: khairil@apps.ipb.ac.id.

Kayla Fakhriyya Jasmine holds a Bachelor of Statistics degree from IPB University, Indonesia, graduated in 2024. She is currently working as an internal auditor. During her undergraduate study, she developed a strong interest in data analysis, statistics, and machine learning, particularly in how data-driven models can be applied to support decision-making. She also gained experience in applying statistical methods and predictive techniques to real-world problems during her coursework and research projects. In her current professional role, she continues to explore how analytical approaches can contribute to process improvement and performance evaluation. She can be contacted at: kaylafjas@gmail.com or kjasmine@gmf-aeroasia.co.id.

Dr. Ir. Indahwati, M. Si. is a distinguished academic and researcher in the field of statistics. She earned her bachelor, master, and doctoral degree in Statistics from IPB University, Indonesia. As a researcher, she has extensively published in both national and international journals. She has served as a supervisor or member of the supervisory committee for undargraduate, master, and doctoral students. She can be contacted at email: indahwati@apps.ipb.ac.id.

Wandee Wanishsakpong is an assistance professor at Department of Statistics, Faculty of Sciences, Kasetsart University, Thailand. She completed her master's degree in Applied Statistics from National Institute of Development Administration, Thailand and Ph.D. in Research Methodology from Prince of Songkla University, Thailand. During the last ten years, she has taught and do research the following subjects: time series analysis, multivariate analysis, regression analysis sampling technique, statistical method, and data science. She can be contacted at email: wandee.w@ku.th.