JURNAL INFOTEL
Informatics - Telecommunication - Electronics
Website: https://ejournal.ittelkom-pwt.ac.id/index.php/infotel
ISSN: 2085-3688; e-ISSN: 2460-0997

LSTM forecast of volatile national strategic food commodities

Herlina Jayadianti[1], Vynska Amalia Permadi[2,*], Partoyo[3]
[1,2]Informatics Department, Universitas Pembangunan Nasional Veteran Yogyakarta
[3]Soil Science Department, Universitas Pembangunan Nasional Veteran Yogyakarta
[1,2]J1. Babarsari, No. 2, Sleman 55281, Indonesia
[3]Jl. Padjajaran, No. 104, Sleman 55283, Indonesia
[*]Corresponding email: vynspermadi@upnyk.ac.id

Received 5 September 2023, Revised 8 Oktober 2023, Accepted 23 Oktober 2023

Abstract — Using the long short-term memory (LSTM) forecast, this research suggested a short-term projection model for national critical food pricing commodities. The model was trained using historical time-series data from each commodity price over the previous three years. The results demonstrated that the proposed LSTM architecture model was generalizable to all commodities and performed well in most cases. This result indicates that the model is resilient and can forecast commodity prices and offer accurate forecasts for most of the ten volatile national strategic foods, with an error value of less than 0.01 and an accuracy value of > 95 %. The model, however, needed to recognize the pricing pattern in cooking oil and beef commodities, both of which had increasing trend patterns. Hence, further research is needed to improve the model's performance for commodities with volatile prices. This could be done using a more extensive and diverse dataset and a more extended lookup date. Additionally, differentiating the LSTM architecture for commodities with different data distribution characteristics may be helpful. Finally, training the data with more epochs may also improve model performance. As the implications of this study, policymakers and stakeholders can use this predictive model to make better decisions about food prices and inflation, which can significantly impact food security and economic stability. For example, policymakers can use the model to set price floors and ceilings for essential food items, develop targeted subsidies for low-income households, and make better production and inventory planning decisions.

Keywords – forecasting, national strategic food commodities, price, volatile food

I. Introduction

Analyzing how agricultural commodity prices and volatility behave is vital since they considerably impact trade terms, inflation expectations, government budget balance, development prospects, and external debt [1]. Several of them, however, are national strategic food commodities whose prices are affected by various factors, including adverse weather events [2], market speculation [3], national policies [4], and the coronavirus pandemic [5]. Meanwhile, as institutional investors grow their stakes in commodity markets [6], the agricultural market is becoming more financialized, which could destabilize agricultural commodity prices.

As a result, agricultural commodity futures prices are exceedingly complicated, nonlinear, and erratic, making precision forecasting difficult. Consequently, agricultural commodity futures prices are highly complex, nonlinear, and confused, making precise forecasting more difficult. Before 2008, according to the United Nations Food and Agricultural Organization (FAO), food prices rarely experienced significant volatility [7]. Agricultural commodities, on the other hand, have undergone massive price changes during 2008-2018, resulting in both high and low volatility regimes [8]. The G20 requested a report from several international organizations (including the World Bank, IMF, UNCTAD, OECD, and FAO, among others) to "develop options for G20 consideration on how to better mitigate and manage the risks associated with food and other agricultural commodity price volatility, without distorting market behavior." [9].

Numerous forecasting techniques have been devised in previous research to improve prediction accuracy. Various techniques are utilized to forecast the prices of various commodities and stocks. The proposed methods are often classified into three categories. The first group includes classic statistical methods, followed by an artificial intelligence-based approach and hybrid methods [10]–[12]. Traditional econometric and statistical procedures, such as smoothing (ETS), benefit from their concise methodology [13], in addition to the autoregressive integrated moving average (ARIMA) [14], which are the most commonly utilized strategies for forecasting commodity prices. These approaches, however, are inadequate at capturing the nonlinear component of future price series since they are based on the assumption of approximation linearity. Machine learning (ML) approaches, classified as artificial intelligence (AI) approaches, have powerful data-driven features and adaptive learning capabilities that allow them to extract latent factors that traditional methods cannot capture successfully.

Back-propagation neural networks (BPNN) are neural networks that learn by doing [15]–[18], and popular machine learning algorithms for commodity price forecasting include extreme learning machines (ELM) [19], [20]. Among other shortcomings, ML approaches are sensitive to parameter values and prone to overfitting. No forecasting method is superior to all others; each technique has merits and weaknesses [21].

It is imperative to develop the frameworks necessary to accurately forecast agricultural commodity price volatility given the aforementioned market dynamics and recent agricultural volatility research so that policy institutions can design preventative policies or prepare for periods of high price volatility, as suggested by [22]. Policy institutions have recently demonstrated the need for precise agricultural price volatility forecasts. However, agricultural price volatility modeling techniques have been developed for over 15 years, and [23] were the first to attempt to produce actual out-of-sample forecasts.

The research [23] begins by focusing on the first of this research and uses GARCH-type models to create estimates for the price volatility of cocoa, coffee, and sugar. Some research [24]–[27] take advantage of the increased availability of ultra-high-frequency data to extend Corsi's [28] heterogeneous autoregressive (HAR) model to create short-run volatility forecasts (up to 20 days ahead). The research [25] estimated the realized volatility of five agricultural commodities sold on the Chinese market using two regime-switching Markov models: soybeans, soybean oil, white sugar, gluten wheat, and cotton. They find evidence that regime-switching dynamics outperform a simple AR(1) model and a Markov-switching AR(1) model in terms of predictive value.

In a similar manner to research [26], [27] fore-casted soybean, cotton, gluten wheat, and maize prices using intra-day data from the Chinese commodity futures markets (Zhengzhou commodity exchange and Dalian commodity exchange). The HAR model is extended in this study with potential predictors (such as day-of-week dummies, historical cumulative returns, and the leap component). According to their research, principal component and tagging-based HAR models can produce better forecasts than AR models alone.

The research [28] developed and fore-casted volatility measures using Chinese market futures prices for soybean, cotton, gluten wheat, maize, early Indica rice, and palm. They also employ additional calculated volatility measures (such as daily log-range volatility, real threshold multi-power variation, and real threshold bi-power variation) as possible predictors of observed volatility and the jump component. Predictors and coefficients in their prediction models can change over time. In terms of predictive ability, according to their findings, the dynamic model average and the Bayesian model average outperform the basic HAR model. Furthermore, they demonstrate that the HAR model with time-varying sparsity for all commodities evaluated generates the most accurate forecasts.

Given the lack of research on agricultural price volatility forecasting and the significance of such projections, this area of inquiry warrants additional attention. Furthermore, there hasn't been much research into forecasting agricultural commodity volatility using AI-based machine learning or deep learning models. The previous recent studies also did not consider Indonesia's national strategic food commodities. Hence, the study aims to implement one of the recurrent neural network-based forecasting techniques, Long Short-Term Memory (LSTM), to model and test the evaluation of forecasting results carried out for ten national food commodity products in Indonesia that are said to influence inflation projection, which are: rice, red pepper, cayenne pepper, onion, garlic, chicken, beef, cooking oil, and eggs.

This research will construct an LSTM model architecture for price forecasting and analysis of ten national food commodity products in the Yogyakarta area. Data from the previous three years is taken from open data sources on the Central Bank of Indonesia website to train the model. The models are then assessed based on their ability to create forecasts, indicating their trustworthiness. To present the findings of this study, this paper is organized as follows: The proposed research technique, the assumptions, needs, and limitations of this study are all addressed in Section II. The result section briefly reviews the LSTM model used to forecast national food commodity products. Section IV discusses and analyzes the Predicted vs. Actual Results of the LSTM model, and the last section summarizes these research findings.

II. Research Method

This section discusses the assumptions, needs, and limitations; businees objectives; and analysis approach.

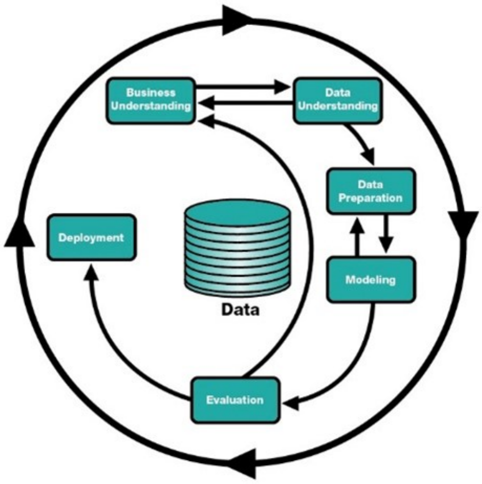

Fig. 1. The CRISP-DM methodology [29].

A. Assumptions, Needs, and Limitations

Price changes in uncontrolled critical commodities might result in losses for consumers and producers. Making accurate price estimates is one step toward addressing these challenges and implementing preventive actions to limit price volatility. In this study, the prices of ten essential national strategic food commodities—rice, purebred chicken meat, purebred chicken eggs, beef, onion, garlic, red chili, cayenne pepper, cooking oil, and sugar—will be predicted using long short-term memory (LSTM). Data is obtained from direct survey results from markets in each city/district that have been published by the Central Bank of Indonesia (https://www.bi.go.id/hargapangan).

The outcome of this research is the implementation of an LSTM model architecture for price forecasting and analysis of ten food commodities in the Yogyakarta area. The models are then assessed based on their propensity to produce forecasts, which summarizes how reliable each model is. The model is trained using historical commodity price data from the 26th August,2021 to the 26th August, 2023.

B. Business Objectives

This research is being carried out in to create a short-term market projection model for strategic food pricing commodities in the Yogyakarta region, using time-series forecasting of ten food commodities, which significantly contributes to the formation of inflation, particularly volatile food inflation. The expected outcome is the creation of a short-term price projection tool for staple food commodities that can be used regularly basis by the Regional Inflation Control Team or other relevant stakeholders in developing policy recommendations and/or short-term programs or market inspections to support market performance improvement, particularly for staple food commodities with the following details:

1) Rice: This commodity's data includes six varieties of rice quality based on price level: two types of regular/lower quality rice, two types of medium quality rice, and two types of premium quality rice. Rice kinds were chosen based on the type most often consumed by the community in the city/district of the sample location. Rason/rastra rice is not included in the price of ordinary / lesser quality rice. The price stated is the price per kilogram. The price utilized for subsequent analysis is the average of all available pricing data.
2) Onion: This commodity's data consists of only one quality of onion, which is local and of mediocre quality. The price stated is the price per kilogram.
3) Garlic: This commodity's data consists of only one quality garlic in medium-quality weevil. The price stated is the price per kilogram.
4) Red chili: This commodity's data consists of large redand fresh-curly red chili. The price stated is the price per kilogram. The price utilized for subsequent analysis is the average of all available pricing data. The price utilized for subsequent analysis is the average of all available pricing data.
5) Cayenne pepper: This commodity's data consists of two qualities: red cayenne pepper and green cayenne pepper in fresh condition. The price stated is the price per kilogram. The price utilized for subsequent analysis is the average of all available pricing data.
6) Beef: This commodity's data consists of two qualities: beef is fresh on the exterior and inside. The price stated is the price per kilogram. The price utilized for subsequent analysis is the average of all available pricing data.
7) Purebred chicken meat (chicken): This commodity's data only includes one quality, purebred chicken meat without offal in fresh condition. The price stated is the price per kilogram.
8) Purebred chicken eggs (eggs): This commodity's data only includes one quality, fresh chicken egg. The price stated is the price per kilogram.
9) Granulated sugar (sugar): This commodity's data is divided into local/bulk quality yellow color and premium quality. The price stated is the price per kilogram. The price for subsequent analysis is the average of all available pricing data.
10) Cooking oil: This commodity's data consists of three qualities: one local/bulk quality and two refill packaging qualities. The price stated is the price per liter. The price utilized for subsequent analysis is the average of all available pricing data.

C. Analytic Approach

This research adopts a Data Science analysis model using the CRISP-DM [29]. As seen in Fig. 1, for the second stage, data understanding, data will be acquired from open data sources on the Central Bank of Indonesia website. The obtained data consists of only two attributes: the time-frame and price. Then, several sets of preprocessing steps will be carried out: adjusting the data format to the needs of analysis, filling in the null data (imputation), and carrying out the initial stages of descriptive statistics and visualization.

The LSTM Model is the solution or model that will be used to anticipate food prices based on the ten commodities chosen. Because LSTM models demand huge amounts of data, data availability in the last three years will be assessed as part of the evaluation process to achieve the best results. When food price prediction models yield the smallest error value, they are successful. This signifies that the model has a tiny difference between predicted and actual values. The evaluation matrix that will be used to assess the effectiveness of the built model will solely use mean absolute error (MAE) and accuracy by comparing predicted results to actual prices.

III. Result

The dataset that has been imported is a *CSV* file of 1097 rows of data from ten commodities over the last three years. Before the data understanding process is carried out, preprocessing of these files is carried out to combine price data of a type of item every day (according to the date of the *Excel* file). The ultimate goal of this preprocessing is a data frame that represents the dataset with index is the date with the data type of DatetimeIndex, and other columns labeled as item code of type $int64$ represented by commodities (*i.e.*, chicken, onion, garlic, rice, red pepper, cayenne pepper, cooking oil, beef, and egg), with subscripts from 1 to 9. The dataset is ready to be processed for data understanding, statistically or visually.

A. Data Understanding

Because the dataset only consists of two features, Date and Price, only feature understanding is carried out at this stage. Feature understanding includes understanding the feature's content and conducting simple statistical analysis by looking at data distribution through data visualization and checking data correlation. Through correlation analysis, some features have a strong correlation (close to 1, which is 0.876320); there is also a negative correlation between chicken and garlic with a correlation result is -0.162568.

B. Data Preparation

The first data preparation process is normalization by imputing empty or null data. After data has been collected on ten national strategic food commodities: Chicken, Onion, Garlic, Rice, Red Chili, Cayenne Pepper, Sugar, Cooking Oil, Beef, and Eggs, on average, there is one to three blank data (missing value) in each commodity. To fill in null values, the Forward Fill method is selected. Applying this method across data frame indexes will fill in any missing values based on the corresponding value in the previous row.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm (LSTM)                  (None, 14, 100)           40800
_________________________________________________________________
dropout (Dropout)            (None, 14, 100)           0
_________________________________________________________________
lstm_1 (LSTM)                (None, 14, 100)           80400
_________________________________________________________________
dropout_1 (Dropout)          (None, 14, 100)           0
_________________________________________________________________
lstm_2 (LSTM)                (None, 14, 100)           80400
_________________________________________________________________
dropout_2 (Dropout)          (None, 14, 100)           0
_________________________________________________________________
lstm_3 (LSTM)                (None, 100)               80400
_________________________________________________________________
dropout_3 (Dropout)          (None, 100)               0
_________________________________________________________________
dense (Dense)                (None, 1)                 101
=================================================================
Total params: 282,101
Trainable params: 282,101
Non-trainable params: 0
_________________________________________________________________
```

Fig. 2. The LSTM model architecture.

The second normalization carried out is data normalization with the Min-Max Scaling technique. This normalization is needed because in the data understanding stage, through the visualization of the data carried out, there are several data points whose price differences are very far above the average.

C. Modeling

LSTM architectures for all commodities are given in Fig. 2. The LSTM model utilized for all commodities has a 4-layer input architecture (100 units each), dropout parameters *0.2 batch=12*, and *output=1*. Furthermore, the LSTM employed here employs an Adam optimizer and a 100-epoch iteration. After multiple tests with values of 10, 30, 50, and 100 and obtaining more small loss outcomes with units of 100, the number of units of 100 was chosen. The same tests are conducted for dropout values. When tested with a value of 0.5, the accuracy is lower than when tested with a value of 0.2. We had previously also attempted a batch size of 32, but better results were obtained when the batch size was 12. We also tried varying the epoch value from 10, 50, and 100, the lowest MSE result is obtained with an epoch of 100 due to the more repetitions.

D. Evaluation

As an evaluation, in addition to visualizing the results as given in Fig. 3. At each epoch, the mean-squared error (MSE) is calculated, which results are shown in the evaluation table in Table 1 below with epoch around 100 per batch.

IV. Discussion

The prediction given here requires a *look_up_date* or consideration date, which means if we want to predict prices on January 21, 2021, then the LSTM model built will consider the last 14 days of data (January

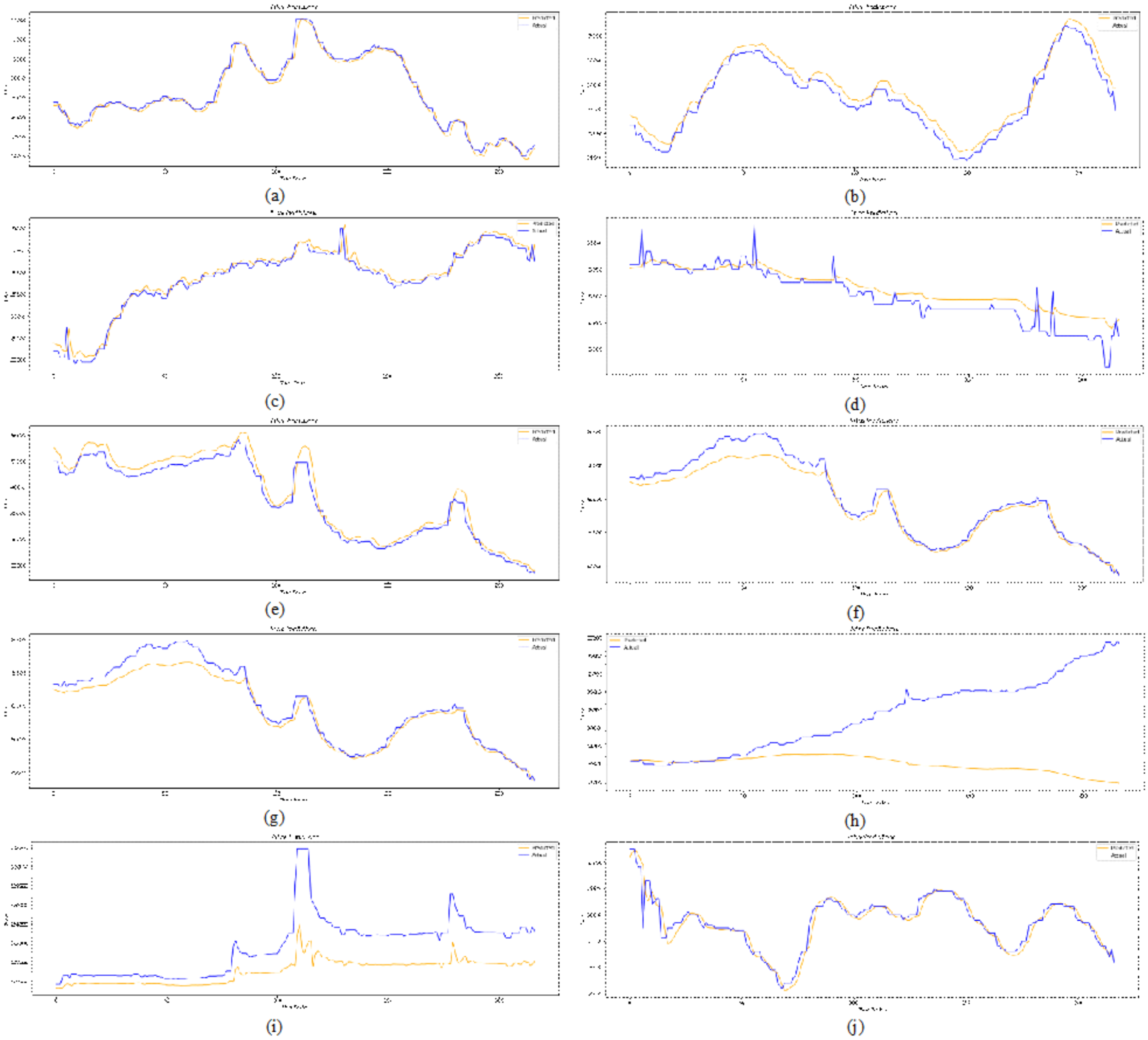

Fig. 3. Predicted vs. Actual evaluation results : (a) Chicken meat, (b) Onion, (c) Garlic, (d) Rice, (e) Red chili, (f) Cayenne pepper, (g) Sugar, (h) Cooking oil, (i) Beef, (j) Eggs.

Table 1. Average MSE Evaluation Results

No	Commodities	Average MSE
1	Chicken Meat	0.0025
2	Onion	0.0015
3	Garlic	0.0009
4	Rice	0.0021
5	Red Chili	0.0021
6	Cayenne Pepper	0.0017
7	Sugar	0.0017
8	Cooking Oil	0.0014
9	Beef	0.0015
10	Eggs	0.0013

7, 2021—January 20, 2021) as a consideration s The average

MSE value from the modeling results is less than 0.01 The model built has provided predictions with an error value of $<$ 0.01. Using testing data for LSTM models built, we also analyze the accuracy of predictions by comparing prediction results with actual prices published on the same website. The resulting accuracy value is $> 95\%$, which means that the difference between the price difference between the actual price and the predicted price given by the LSTM model provides a difference that is not too far apart. However, in some cases, because several external factors sometimes affect price spikes that may be very high, prediction results will be obtained that become significantly less accurate because the input from this model requires the last 14 days of data as consideration. The value of 14 may be replaced, for example, the last 30/60/360, to produce better prediction performance.

The results of accuracy evaluation with testing data as much as 20% of the overall dataset (1097 data) are given in Table 2. Accuracy here serves as additional evidence that the model implemented can be general to all commodities and has good performance with an accuracy value of more than 95%. In addition, looking further from Fig. 3, one LSTM model can produce actual vs. predicted results that are pretty good in most commodities. Of the ten commodities evaluated, 7 have a difference in actual vs. predicted value that is not too far. However, the model needs to performs poorly in cooking oil and Beef commodities. In Cooking Oil, which has data with an increasing trend pattern, the LSTM model shows very different performance.

Table 2. Accuracy Evaluation Results

No	Commodities	Accuracy
1	Chicken Meat	99.56 %
2	Onion	99.01 %
3	Garlic	99.66 %
4	Rice	99.84 %
5	Red Chili	96.91 %
6	Cayenne Pepper	96.72 %
7	Sugar	97.16 %
8	Cooking Oil	96.48 %
9	Beef	97.94 %
10	Eggs	99.65 %

Meanwhile, in Beef commodities, patterns have been successfully identified by the LSTM model, but the resulting prediction results still need to be improved. Unlike the previous two commodities, pattern recognition by LSTM on Rice commodities is also not optimal. Fluctuations in data values that show multiple spikes also need to be adequately identified by this model.

V. Conclusion

The size and quality of the dataset limited the study. A more extensive and more diverse dataset would allow the model to learn more patterns and make more accurate predictions. Moreover, The study was also limited by the period used. The model could perform differently if trained and tested on data from a different period. Overall, the results of this study are promising. The LSTM model was able to provide accurate predictions for most commodities.

The LSTM model built in this study provided predictions with an error value of < 0.01 and an accuracy value of $> 95\%$. The model architecture could generalize to all commodities and performed well in most cases. However, it performed poorly in cooking oil and beef commodities, which have data with increasing trend patterns. However, of the ten commodities, these two do exhibit volatility that is challenging to predict, according to the findings of additional discussions with several experts. It is speculatively believed that this is caused by outside forces, such as price regulation by a particular group over a particular period. This incident might not be preventable, but in the future, variables could be added for more in-depth analysis, including taking into account news reports, particularly regarding these two commodities, to improve the accuracy of price forecasts.

Furthermore, future research might use a larger, more varied dataset and a longer *look_up_date* for the model. Additionally, as cooking oil and beef have different data distribution characteristics from the other commodities, differentiating the LSTM architecture for these two commodities may be considered. Training the data with more epochs may also improve model performance.

Overall, the study's findings are promising and suggest that LSTM forecast can be used to develop a more accurate and reliable food price prediction model. Policymakers and stakeholders can use this predictive model to make better food prices and inflation, which can significantly impact food security and economic stability. Here are some specific examples of the use of the model:

1) Policymakers can use the model to set price floors and ceilings for these ten essential food items. This can help protect consumers from price spikes and ensure they have access to affordable food.
2) Policymakers can also use the model to develop targeted subsidies for low-income households. This can help to offset the impact of food price increases on the poorest members of society.
3) Stakeholders in the food supply chain, such as farmers and retailers, can use the model to make better production and inventory planning decisions. This can help reduce food waste and ensure a sufficient food supply to meet demand.

References

[1] L. Agnello, V. Castro, S. Hammoudeh, and R. M. Sousa, "Global factors, uncertainty, weather conditions and energy prices: On the drivers of the duration of commodity price cycle phases," *Energy Econ*, vol. 90, p. 104862, 2020, doi: 10.1016/j.eneco.2020.104862.

[2] A. Makkonen, D. Vallström, G. S. Uddin, M. L. Rahman, and M. F. C. Haddad, "The effect of temperature anomaly and macroeconomic fundamentals on agricultural commodity futures returns," *Energy Econ*, vol. 100, p. 105377, 2021, doi: 10.1016/j.eneco.2021.105377.

[3] Y. Zhang, S. Ding, and E. M. Scheffel, "A key determinant of commodity price Co-movement: The role of daily market liquidity," *Econ Model*, vol. 81, pp. 170–180, 2019, doi: 10.1016/j.econmod.2019.01.004.

[4] R. P. Roy and S. Sinha Roy, "Commodity futures prices pass-through and monetary policy in India: Does asymmetry matter?," *The Journal of Economic Asymmetries*, vol. 25, p. e00229, 2022, doi: 10.1016/j.jeca.2021.e00229.

[5] Z. Umar, M. Gubareva, and T. Teplova, "The impact of Covid-19 on commodity markets volatility: Analyzing time-frequency relations between commodity prices and coronavirus panic levels," *Resources Policy*, vol. 73, p. 102164, 2021, doi: 10.1016/j.resourpol.2021.102164.

[6] B. M. Ordu, A. Oran, and U. Soytas, "Is food financialized? Yes, but only when liquidity is abundant," *J Bank Financ*, vol. 95, pp. 82–96, 2018, doi: 10.1016/j.jbankfin.2017.06.001.

[7] FAO, "Price volatility in agricultural markets: Evidence, impact on food security and policy responses," 2010. [Online]. Available: https://www.fao.org/3/am053e/am053e00.pdf

[8] F. Greb and A. Prakash, "Has price volatility changed?," FAO Food Outlook, May 2015.

[9] FAO, "Price volatility in food and agricultural markets: Policy responses," 2011. [Online]. Available: http://www.fao.org/fileadmin/templates/est/Volatility/ Interagency_Report_to_the_G20_on_Food_Price_Volatility.pdf.

[10] H. Y. Kim and C. H. Won, "Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models," *Expert Syst Appl*, vol. 103, pp. 25–37, 2018, doi: 10.1016/j.eswa.2018.03.002.

[11] A. Vidal and W. Kristjanpoller, "Gold volatility prediction using a CNN-LSTM approach," *Expert Syst Appl*, vol. 157, p. 113481, 2020, doi: 10.1016/j.eswa.2020.113481.

[12] M. Zolfaghari and S. Gholami, "A hybrid approach of adaptive wavelet transform, long short-term memory and ARIMA-GARCH family models for the stock index prediction," *Expert Syst Appl*, vol. 182, p. 115149, 2021, doi: 10.1016/j.eswa.2021.115149.

[13] L. Wu, S. Liu, and Y. Yang, "Grey double exponential smoothing model and its application on pig price forecasting in China," *Appl Soft Comput*, vol. 39, pp. 117–123, 2016, doi: 10.1016/j.asoc.2015.09.054.

[14] S. Torbat, M. Khashei, and M. Bijari, "A hybrid probabilistic fuzzy ARIMA model for consumption forecasting in commodity markets," *Econ Anal Policy*, vol. 58, pp. 22–31, 2018, doi: 10.1016/j.eap.2017.12.003.

[15] J. Manickavasagam, S. Visalakshmi, and N. Apergis, "A novel hybrid approach to forecast crude oil futures using intraday data," *Technol Forecast Soc Change*, vol. 158, p. 120126, 2020, doi: 10.1016/j.techfore.2020.120126.

[16] S. Zhou, K. K. Lai, and J. Yen, "A dynamic meta-learning rate-based model for gold market forecasting," *Expert Syst Appl*, vol. 39, no. 6, pp. 6168–6173, 2012, doi: 10.1016/j.eswa.2011.11.115.

[17] Z. Ji, D. Niu, M. Li, W. Li, L. Sun, and Y. Zhu, "A three-stage framework for vertical carbon price interval forecast based on decomposition–integration method," *Appl Soft Comput*, vol. 116, p. 108204, 2022, doi: 10.1016/j.asoc.2021.108204.

[18] J. Li, G. Li, M. Liu, X. Zhu, and L. Wei, "A novel text-based framework for forecasting agricultural futures using massive online news headlines," *Int J Forecast*, vol. 38, no. 1, pp. 35–50, 2022, doi: 10.1016/j.ijforecast.2020.02.002.

[19] P. Du, J. Guo, S. Sun, S. Wang, and J. Wu, "Multi-step metal prices forecasting based on a data preprocessing method and an optimized extreme learning machine by marine predators algorithm," *Resources Policy*, vol. 74, p. 102335, 2021, doi: 10.1016/j.resourpol.2021.102335.

[20] T. Xiong, C. Li, and Y. Bao, "Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China," *Neurocomputing*, vol. 275, pp. 2831–2844, 2018, doi: 10.1016/j.neucom.2017.11.053.

[21] R. Adhikari and R. K. Agrawal, "Performance evaluation of weights selection schemes for linear combination of multiple forecasts," *Artif Intell Rev*, vol. 42, no. 4, pp. 529–548, 2014, doi: 10.1007/s10462-012-9361-z.

[22] F. Greb and A. Prakash, "Assessing volatility patterns in food crops," 2017. [Online]. Available: http://www.fao.org/3/a-i7066e.pdf.

[23] P. Giot and S. Laurent, "Market risk in commodity markets: a VaR approach," *Energy Econ*, vol. 25, no. 5, pp. 435–457, 2003, doi: 10.1016/S0140-9883(03)00052-5.

[24] J. Luo, T. Klein, Q. Ji, and C. Hou, "Forecasting realized volatility of agricultural commodity futures with infinite Hidden Markov HAR models," *Int J Forecast*, vol. 38, no. 1, pp. 51–73, 2022, doi: 10.1016/j.ijforecast.2019.08.007.

[25] F. Tian, K. Yang, and L. Chen, "Realized volatility forecasting of agricultural commodity futures using long memory and regime switching," *J Forecast*, 2016, doi: 10.1002/for.2443.

[26] F. Tian, K. Yang, and L. Chen, "Realized volatility forecasting of agricultural commodity futures using the HAR model with time-varying sparsity," *Int J Forecast*, vol. 33, no. 1, pp. 132–152, 2017, doi: 10.1016/j.ijforecast.2016.08.002.

[27] K. Yang, F. Tian, L. Chen, and S. Li, "Realized volatility forecast of agricultural futures using the HAR models with bagging and combination approaches," *International Review of Economics & Finance*, vol. 49, pp. 276–291, 2017, doi: 10.1016/j.iref.2017.01.030.

[28] F. Corsi and R. Renò, "Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling," *Journal of Business & Economic Statistics*, vol. 30, no. 3, pp. 368–380, 2012, doi: 10.1080/07350015.2012.663261.

[29] C. Shearer, "The CRISP-DM model: The new blueprint for data mining," *Journal of Data Warehousing*, vol. 5, pp. 13–22, 2000.