Forecasting Ginger Harvest Yields: A Comparative Study of Double Exponential Smoothing and Long Short-Term Memory Models

Devie Rosa Anamisa[1*] , Fifin Ayu Mufarroha[1] , Achmad Jauhari[1] , Bain Khusnul Khotimah[1] , Mohammad Yanuar Hariyawan[2] , Ahmad Farisul Haq[1]

[1] Informatics Engineering Department, University of Trunojoyo Madura, Bangkalan 69162, Indonesia
[2] Electrical Engineering Department, Telkom University, Surabaya 60231, Indonesia

Corresponding Author Email: devros_gress@trunojoyo.ac.id

ABSTRACT

Ginger, a vital herbal commodity, experiences low yield rates, necessitating intensive cultivation and rigorous evaluation by farmers to ensure financial viability and alignment with market demands. This study was conducted to devise a harvest forecasting system that supports decision-making through minimal error rates by comparing double exponential smoothing (DES) and long short-term memory (LSTM) forecasting methods. The efficacy of these methods was assessed through a series of trials, analyzing data collected from 2015 to 2019, comprising 250 datasets. The evaluation focused on two primary metrics: the Mean Absolute Percentage Error (MAPE) and the Root MSE (RMSE), to determine the precision of forecast models. It was observed that the LSTM model outperformed the DES method, yielding a MAPE of 38.99% and an RMSE of 1244.85432, in contrast to the DES method which resulted in a MAPE of 43.49% and an RMSE of 12997.34261, at an alpha level of 0.4 and an optimal beta of 0.1. Given these findings, the LSTM model is recommended for the forecast of ginger yields due to its superior accuracy and lower standard error compared to the DES method. This comparative analysis underscores the importance of selecting appropriate forecasting models to enhance agricultural planning and productivity, particularly in crops with fluctuating yields such as ginger.

1. INTRODUCTION

The agricultural sector has considerable economic potential. However, the potential for agricultural products, especially ginger, is still relatively low. Therefore, farmers need to plan the yields obtained to avoid substantial losses and meet market demand. This leads to systems that help them in decision-making and future planning, such as a forecasting system using past track record data [1]. Forecasting is the prediction of future results to meet precise and appropriate targets [2]. This research on forecasting the yield of ginger plants helps farmers make plans and meet market demand greatly.

To overcome the above problems, this research requires a forecasting method to produce accurate predictions. Some previous research has applied forecasting methods. Chung and Kim [3] compared the Single Exponential Smoothing (SES) and Artificial Neural Network (ANN) methods, and found that the Mean Square Error (MSE) value of SES was lower compared to that of the ANN method. In addition, they developed the DES method for efficient jitter compensation, which showed that DES-based schemes ran about 100 times faster than Extended Kalman Filter (EKF)-based methods and 19 times faster than Kalman Filter (KF)-based methods. Siregar and Wibawa [4] compared the DES and ANN methods with the SES process on data input for foreign currency exchange, and found that the MAPE results were 53% with an

execution time of 561 seconds. Therefore, it can be concluded that the DES method is better than SES for improving ANN performance for forecasting foreign currency exchange rates. In 2020, Regina and Jodiawan [5] forecasted Fast Moving Consumer Goods using time series forecasting and compared it with the Autoregressive Integrated Moving Average Model (ARIMA), moving average (MA), DES, and linear regression (RL) methods. The results of this study indicate that the DES method has the smallest MAPE value among the three, with optimal alpha and gamma. In addition to statistical forecasting methods, machine learning methods have also been studied. The machine learning method is an algorithm that allows users to find and describe structural patterns in data so that the structural patterns can be used [6]. LSTM is one of the time-series forecasting methods of the deep learning group based on Recurrent Neural Network (RNN) with good accuracy. Several studies have used the LSTM method as a time series forecasting method. Shankar et al. [7] compared the LSTM method with seven different time series forecasting methods, such as ARIMA, simple exponential smoothing, Holt-Winter's, error-trend-seasonality, trigonometric regressors (TBATS), neural networks and hybrid ARIMA for container forecasting that Sonali Shankar developed. This study shows that the accuracy of the LSTM method outperforms the seven methods in long-term forecasting. In addition, Bathla [8] predicted stock prices and found that the LSTM method provided better

accuracy than Support Vector Regression (SVR). Agarwal and Tarar [9] also found that deep learning methods, such as LSTM, yielded higher accuracy compared to machine learning algorithms for crop prediction, and achieved clear information regarding the amount of soil materials required by removing them separately.

Some of these studies indicate that the accuracy of the LSTM and DES methods has a small error value in forecasting [10]. However, both methods have weaknesses, requiring a lot of data and regular maintenance. This takes quite a while to produce accurate predictions. Therefore, it is necessary to compare the two to find patterns with a small error rate. This study uses seasonal data to compare the LSTM and DES methods for predicting ginger yields in Madura. When changes occur repeatedly over a certain period, it creates seasonal data. With these two methods, a forecasting model can be selected as the best and most accurate method for predicting ginger yields. This study aims to find the best way to provide accurate forecasting results using the LSTM and DES methods via error measures, such as MAPE and RMSE values. After knowing the suitable model for predicting ginger yields, forecasting results can undoubtedly be obtained, which can reduce the accumulation of excess profits from erratic consumer demand and reduce losses experienced by ginger farmers in Madura.

2 PRELIMINARIES

2.1 Literature review

The crop yield forecasting literature is quite extensive. Therefore, it is possible to find forecasting methods for systematic and pragmatic prediction through previous relevant data. Forecasting methods can be used to recognize the elements of data that affect the amount of deviation due to unexpected factors. Many studies have compared several forecasting models to find a suitable model to predict crop yields, such as least squares [11, 12], SES [13], winter exponential smoothing [14], weight moving average [15], and ARIMA [16]. However, these models are not accurate at predicting variables. In 2022, Asrul [17] used the DES method to predict potato vegetable yields, which met the annual potato crop production as a material for consideration and the amount of potato production for subsequent market demand was recommended. Therefore, this method can improve forecasting by smoothing the average previous value of time series data in a decreasing (exponential) way. This can provide accurate short-term forecasting, and is easily adapted to

changes in data without requiring a lot of data [18].

In addition to the DES method, the LSTM method also has advantages in forecasting time series data, which can be used to overcome long-term dependencies on the input [19]. In addition, LSTM also has a memory block that determines which value is selected as the output, which is relevant to the given input [20]. Several studies have been conducted in the domain of agricultural forecasting, specifically ginger yield prediction. Elpawati et al. [21] analyzed the relationship between ginger-exporting countries, such as Indonesia, China, India, and the Netherlands, using the Value at Risk/Vector Error Correction Model (VaR/VECM) method. It was found that Indonesia's ginger exports increased by 92%, followed by the Netherlands (7%), China (0.2%), and India (0.8%). Apart from that, Das et al. [22] forecasted ginger production in Bangladesh, and tested the performance of eight trend models using the coefficient of determination (R^2) and adjusted R^2. This research shows that the compound growth rate of ginger production is 1.019 per year.

This study tries to compare the accuracy of the DES and LSTM methods in predicting ginger yields based on the nature of the composition, forecasting time, and data patterns. The MAPE and RMSE approaches are used to measure accuracy. The best forecasting result is based on the level of prediction error. The smaller the error rate, the more precise a method is in prediction [23].

2.2 Data collection

This study uses the data on ginger crop yields in the Madura Region in 2015-2019, which was obtained from the Department of Agriculture and Food Security in Pamekasan, Madura, after determining the quantity of the ginger harvest through observation. The ginger harvesting process in Madura is carried out once a week from January to December each year so that approximately 250 datasets are obtained, as shown in Table 1. The yield of the ginger commodity fluctuates from year to year, as shown in Figure 1. This original data consists of yields, years and names of the ginger-producing areas. Before data processing, the data was plotted first to determine the pattern of data flow, making forecasting easier. This data processing aims to ensure that the raw data obtained can be analyzed and conclusions can be drawn easier. Based on the data plot in Figure 1, which shows actual harvest data, it can be concluded that the data distribution is seasonal. In this research, raw data was obtained before analysis. Therefore, the data was preprocessed first, aiming to clean the data from distorted and missing values. This process is highly valid for providing accurate output for forecasting ginger harvest yields.
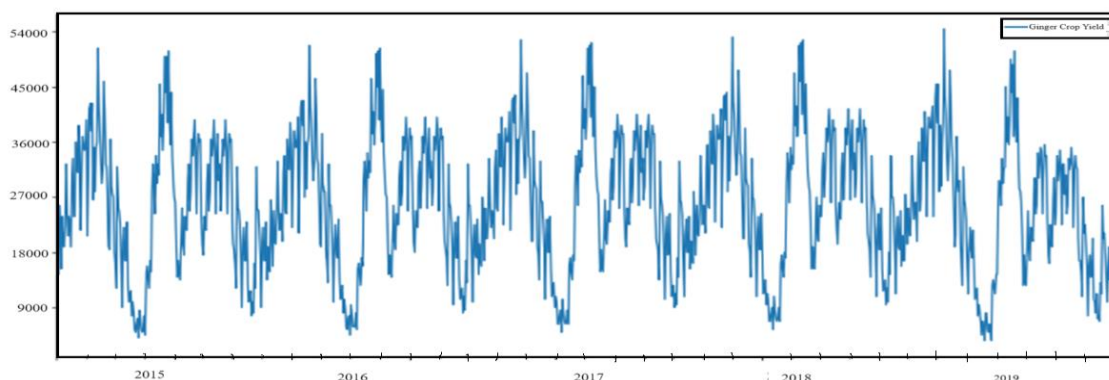

Figure 1. Data on ginger commodity yields

Table 1. Sample dataset of ginger plant harvesting

No.	Region	Year	Ginger Crop Yield (kg)
1	Arosbaya	2015	48347
2	Bangkalan	2016	49554
3	Blega	2017	19705
4	Burneh	2018	15960
5	Galis	2019	35468
6	Ambunten	2015	38602
7	Arjasa	2016	42956
8	Batang-batang	2017	45087
9	Batuan	2018	30961
10	Batuputih	2019	24931
….	…….	….	……..
225	Galis	2015	36303
226	Ambunten	2016	37905
227	Arjasa	2017	47622
228	Batang-batang	2018	25870
229	Batuan	2019	22023
230	Batuputih	2015	19519
231	Banyuates	2016	14634
232	Camplong	2017	32276
233	Jrengik	2018	19229
234	Lenteng	2019	12407
…	…	…	…

To predict ginger yields, a time-series forecasting method was used. This means that the data is presented based on the time of occurrence without indicating the factors that influence it. The time series method is a quantitative prediction, which is based on the analysis of the pattern of relationships between the variables to be searched (dependent) and the variables that affect them (independent), and changes from time to time, such as weeks, months, quarters, semesters, and years. According to the study of Estiningtyas et al. [24], ginger production is influenced by the abiotic environment, including all living things, such as pests, pathogens, and weeds, which interfere with ginger cultivation. Furthermore, the collected data was analyzed by comparing forecasting methods.

2.3 Forecasting method

Forecasting is an approach to predicting the possibilities for future situations by testing data that occurred in the past [25]. It predicts event history data or events in the business sector. The importance of forecasting is that it influences someone in making decisions and can also be used as a basis for long-term planning efforts in an organization [26]. Forecasting is usually divided into three types, such as short-term, medium-term, and long-term [27]. Short-term forecasting is usually used to predict events using periods of the next day, week, or month [28]. Medium-term forecasting is a forecasting approach that utilizes time data from one year to two years into the future [29]. Finally, long-term forecasting aims to find out events more than two years in the future. Usually, this forecasting uses a time series method which is based on historical data and predicts future event data as output [30]. The purpose of forecasting is to reduce the risk in conditions of uncertainty about something that might happen in the future, thereby minimizing this uncertainty. According to the study of Wang and Chaovalitwongse [31], the approach through forecasting or forecasting methods is divided into two parts, i.e., quantitative and qualitative methods. Qualitative methods are used when no sufficient previous data is available. In other words, this method can be used as a basis for consideration in making decisions for forecasting cases with previous data. However, if a lot of previous data is available and meets the

criteria, then forecasting with quantitative methods is more effective than qualitative ones [32]. Forecasting using quantitative methods has several requirements. For example, previous data is available, can be quantified, and has the same trend as future data [33]. To find patterns in the past series and extrapolate these patterns in the future by analyzing the data, this study compares the DES and LSTM methods as a reference for predicting future values.

2.4 DES (Holt)

DES is a linear model proposed by Afiyah et al. [18]. In the DES method, the smoothing process is carried out twice. In principle, Holt's linear and exponential smoothing methods use the multiple smoothing formulae directly [34]. Instead, Holt decides on a seasonal value with parameters different from the two used in the original series. The trend is a smoothed estimate of the average growth at the end of each period [35]. Forecasting Holt's linear and exponential smoothing has several steps [36]:

(i) The first smoothing value is determined.

$$S_t' = \alpha X_t + (1 - \alpha)S_t' \tag{1}$$

(ii) The second smoothing value is determined.

$$S_t'' = \alpha S_t' + (1 - \alpha)S_t'' \tag{2}$$

(iii) The constant value (α_t) is determined.

$$\alpha_t = S_t' + (S_t' - S_t'') = 2S_t' - S_t'' \tag{3}$$

(iv) The slope value (β_t) is determined.

$$\beta_t = \frac{\alpha}{1 - \alpha}(S_t' - S_t'') \tag{4}$$

(v) The forecasting value is determined.

$$F_{t+m} = S_t' + t_t m \tag{5}$$

where, x_t is the demand data in period t, S_t' is the SES value, t_t is the trend value in t, α and β are the intermediate smoothing parameters ranging between 0-1, F_{t+m} is the forecasting of m periods, and m is the number of future periods to be forecast.

The S_t' parameter is a basic and pragmatic technique used for forecasting with time series anticipation that estimates only the level components. Meanwhile, S_t'' is specifically for trends in univariant time series which can help change trends over time in different ways, either in a linear trend or an exponential trend. α_t is the level smoothing factor and β_t is the trend smoothing factor with the best trend estimate at time t.

2.5 Deep learning LSTM

Deep learning is a branch of machine learning that uses a deep neural network to solve problems [37]. Neural networks are inspired by how neurons work in the human brain. Each neuron in the human brain is interconnected and information flows from and between each of these neurons [38]. In deep learning, the network consists of several layers which are collections of nodes [39]. A node becomes the place where calculations occur. Compared to neural networks, deep learning has more hidden layers, such as more than three (including input and output), or even up to hundreds. LSTM is

often used to overcome deficiencies found in RNNs, which are the phenomenon of the magnitude of the gradient disappearing [40]. LSTMs use larger data sets and all data information as input to build deep networks. LSTM has three gates that control the use and updating of past text information, namely, the input gate, forget gate, and output gate [41]. There are several steps in LSTM among others [42]:

(i) Calculation of forget gates (f_t). The value of the forget gate is between 0 and 1, as shown in Eq. (7). Information that is not needed for the cases being managed is removed using the sigmoid function (σ);

(ii) Calculation of cell states (c_t). The tanh activation function is used, which forms a new context candidate, as shown in Eq. (8);

(iii) Calculation of the gate output (o_t). Sigmoid (σ) is used to generate output values and process cell state (c_t) on tanh activation, as shown in Eq. (9).

$$f_t = \sigma\left(W_f.[h_{t-1}, x_t] + b_t\right) \tag{7}$$

$$\begin{aligned} i_t &= \sigma(W_i.[h_{t-1}, x_i] + b_i \\ \hat{c_t} &= \tanh(W_c.[h_{t-1}, x_t] + b_c \end{aligned} \tag{8}$$

$$\begin{aligned} c_t &= f_t * c_{i-1} + i_t * \hat{c_t} \\ o_t &= \sigma(W_0.[h_{t-1}, x_t] + b_0 \\ h_t &= o_t \tanh(c_t) \end{aligned} \tag{9}$$

Before the model was trained, the parameters needed for the LSTM model were determined. The hidden state process in LSTM goes through four gates, namely, the forget gate, the input gate, the cell state and the output gate. The forget gate is the first gate that the input passes through using a sigmoid activation function or a sigmoid gate. The hidden state determines the gate input value. At the input gate, the value of the new candidate cell state is calculated and obtained using the tanh activation function. Meanwhile, the gate output uses the sigmoid activation function. The gate output value is used to generate a new hidden state value along with the cell state value.

2.6 Evaluation method

According to the study of Peñaloza et al. [43], the accuracy of the forecasting results can be seen from the large difference between the actual and estimated values of the forecasting. The accuracy value is the difference between the actual and forecast results. The residual value is obtained by measuring the accuracy of forecasting results, i.e., MAPE and RMSE. MAPE is the absolute average percentage error for evaluation calculations in measuring the precision or accuracy of a prediction [44]. MAPE is calculated using the absolute error for each period divided by the real observed value for that period. The general formula for MAPE can be seen in Eq. (10) [45]. The smaller the MAPE value, the more accurate the forecasting model [46]. RMSE is the root result of the average square of the difference between actual and predicted data, as shown in Eq. (11) [7].

$$MAPE = \sum_{t=1}^{n} \left| \frac{y_i - \hat{y}_i}{\hat{y}_i} \right| x100\% \tag{10}$$

where, $\hat{y}_1$ is the forecasting result, y_1 is the actual value, X_t is the actual value in t data, F_t is the forecast value on t data, and

n is the number of data periods.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i}^{n} (\dot{y}_t - y_i)^2} \tag{11}$$

where, n is the amount of data, y_i is the predicted data, and y_i is the actual data.

3 RESULTS

This chapter explains several test scenarios for ginger yield forecasting systems. The best α and β parameters influence the forecasting process of the DES method in producing the smallest MAPE and RMSE values, and the forecast results are close to the actual value. Meanwhile, the LSTM method is influenced by the input, forget and output gates for the most minor MAPE and RMSE results.
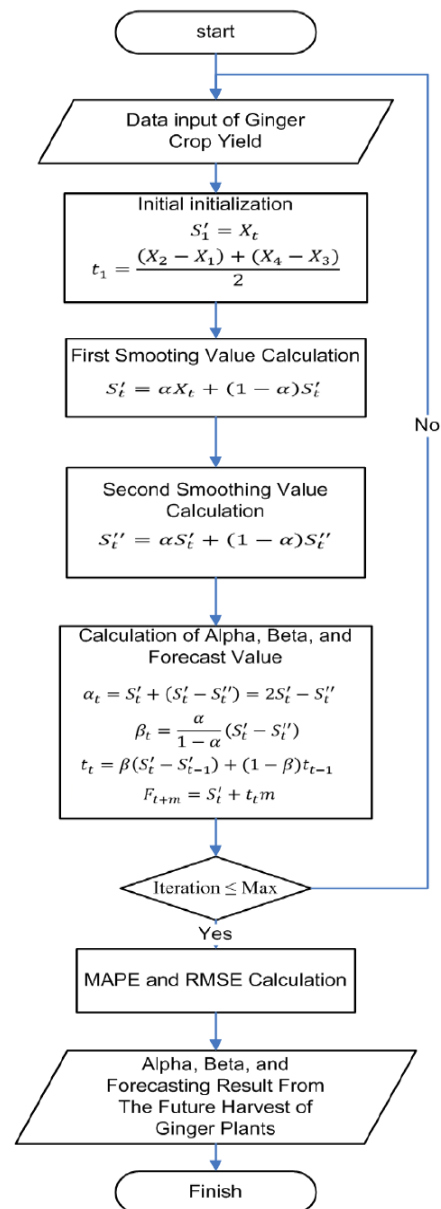
3.1 System flow and DES test scenario

Figure 2. A block diagram of the DES method for forecasting ginger yields

Table 2. DES test result for ginger yields forecasting with α=0.4 and β=0.1

No.	Dataset	Level (S'_t)	Trend (t_t)	Prediction (F_{t+m})
1	0.69202			
2	0.39796	0.397955	-0.294067	
3	0.57513	0.173640	-0.290801	0.103888
4	0.64246	-0.004724	-0.285536	-0.117161
5	0.24749	-0.210664	-0.281808	-0.290259
6	0.08309	-0.407280	-0.277819	-0.492472
7	0.31699	-0.536773	-0.270873	-0.685099
8	0.63391	-0.594272	-0.260881	-0.807646
9	0.59737	-0.640156	-0.250814	-0.855154
10	0.51649	-0.682642	-0.241058	-0.890969
…	….	….	….	….
159	0.81000	0.479021	-0.000593	0.421520
160	0.11856	0.425160	-0.003088	0.478427
161	0.44047	0.424796	-0.002960	0.422073
162	0.00000	0.359397	-0.005884	0.421835
163	0.45080	0.367912	-0.005210	0.353513
164	0.46173	0.377360	-0.004523	0.362702
165	0.26501	0.356876	-0.005271	0.372836
…	…..	….	……	……..
166	0.59175	0.387150	-0.003606	0.351605
167	0.68821	0.428639	-0.001495	0.383543
168	0.33386	0.413336	-0.002141	0.427144
169	0.46529	0.419201	-0.001766	0.411195
170	0.09437	0.369616	-0.004006	0.417435
171	0.07645	0.322809	-0.006010	0.365610
172	0.19649	0.298992	-0.006844	0.316799
173	0.72694	0.356504	-0.003830	0.292148
174	0.84638	0.425750	-0.000408	0.352674
175	0.70816	0.467203	-0.001552	0.425342
…	….	……	……	…..
178	0.75077	0.468365	-0.001551	0.419302
179	0.29629	0.444216	-0,000348	0.469916
180	0.25824	0.416985	-0.000944	0.444564
181	0.29008	0.397396	-0.001817	0.416041
182	0.93113	0.474850	-0.001895	0.395580
183	0.31131	0.452258	-0.000748	0.476745
184	0.50689	0.460982	-0.001122	0.453006

Table 3. Test scenario of the DES method

No.	α	β	MAPE (%)	RMSE	Time (Second)
1	0.1	0.1	59.47	21397.09281	0.00986
2	0.1	0.2	50.65	17125.73846	0.00568
3	0.1	0.4	50.08	14891.05388	0.00580
4	0.1	0.6	50.92	14869.72711	0.00224
5	0.1	0.8	48.97	14667.31313	0.00984
6	0.2	0.1	45.07	14623.72117	0.00306
…	….	….	…..		…..
12	0.3	0.2	44.05	13859.94884	0.00106
13	0.3	0.4	44.52	13106.89026	0.00545
14	0.3	0.6	45.21	13555.78088	0.00005
15	0.3	0.8	46.21	14044.86861	0.00532
16	0.4	0.1	43.49	12997.34261	0.00004
17	0.4	0.2	43.76	13444.92185	0.00697
28	0,6	0.2	45.40	13680.54340	0.00895
29	0.6	0.4	48.53	14565.73533	0.00482
30	0.6	0.6	51.71	15581.54181	0.00015
…	….	….	…..		…..
43	0.9	0.2	51.04	15405.59069	0.00068
44	0.9	0.4	55.63	16818.54106	0.00459
45	0.9	0.6	60.81	18380.95135	0.00880

This section discusses the flow of the ginger yield forecasting system using the DES method, as shown in Figure 2. This figure outlines the system flow and the steps required

to obtain forecasting values using the DES method, as follows:

(i) The magnitude of the parameters α and β is determined between 0 and 1 to examine the accuracy of the prediction results obtained by calculating the MAPE value.

(ii) The first smoothing value is calculated using S'_t.

(iii) The second smoothing value is calculated using S''_t by paying attention to the first smoothing value.

(iv) The value of the constant α_t is determined by referring to the SES adjustment with the difference between SES and DES.

(v) The value of the trend coefficient β_t is determined, thereby determining the estimated trend from one time period to the next.

(vi) The forecasting results F_{t+m} are calculated after calculating the first and second smoothing values, and α_t and β_t values using the best α parameter.

This forecasting process began by inputting the normalized harvest results every week from 2015 to 2019. Then the number of harvests for the coming period was predicted by searching for the best α and β values to calculate the accuracy value, thereby examining the accuracy of the prediction results obtained by calculating the MAPE and RMSE values. In the prediction calculation process using the DES method, the dataset was separated, with 74% of the training dataset and 26% of the testing dataset at the training stage. The results can be seen in Table 2. Meanwhile, Table 3 shows the search for α and β and the effects of MAPE and RMSE measurements using the DES method. As shown in the table, when α=0.2 and β=0.1, a decreasing graphic pattern is produced for the MAPE and RMSE values, which is very significant because the α value used is 0.4 and β is 0.1, which is the lowest point. Therefore, the values of MAPE and RMSE are influenced by the value of the constant α in forecasting.
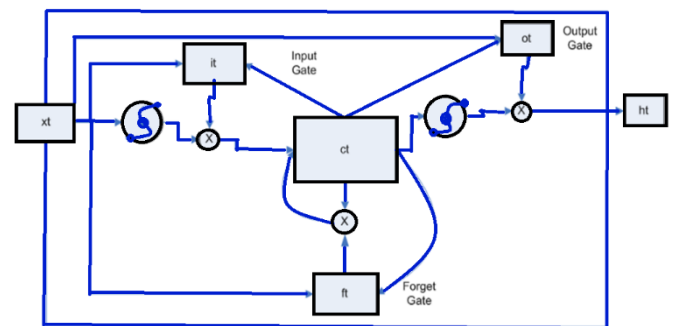

Figure 3. Model of LSTM memory cells

3.2 System flow and LSTM test scenarios

The previous section discusses the DES flow. This section addresses the description of the system flows for forecasting crop yields using the LSTM method, as shown in Figure 3. Model identification was carried out by creating a time-series plot. By plotting a time series, data patterns and trends in the observation series can be seen. The model identification process was carried out with stationarity in the variance and average. The step aims to examine whether the collected data is normally distributed or taken from a normal population. The stationarity of the data in the variance was investigated using the Box-Cox transformation so that the lambda (λ) value obtained in the Box-Cox plot was 1. If the lambda value is not 1, then a Box-Cox transformation must be carried out. The results of the Box-Cox test on ginger harvest data can be seen

in Figure 4. It can be seen from the figure that the data is not stationary in terms of variance because the value $\lambda = -0.50$, or the data is stationary in variance because the value $\lambda \neq 1$ on the Box-Cox plot. This value is smaller than the significance level specified. The probability value is smaller than the 5% significance level, leading to the rejection of H_0. This means that there is no unit root problem or it can be said that the data is stationary. Therefore, parameter estimation is carried out by trial and error with the LSTM model. LSTM uses a form of RNN to avoid long-term dependency problems. This RNN model filters information through gate structures to maintain and update the state of memory cells. Then the model's performance is evaluated using MAPE. In this research, three main stages were carried out below.
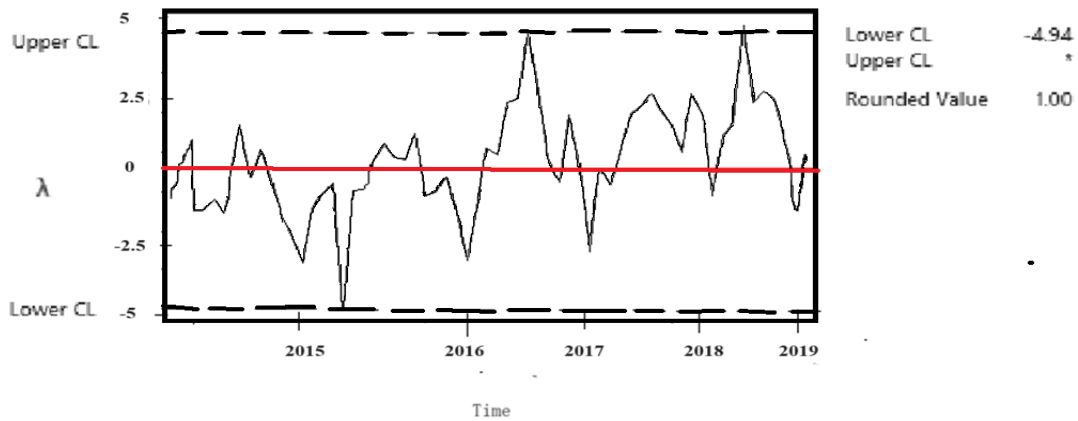

Figure 4. A time-series plot of ginger harvest data for Box-Cox stationary variance transformation

3.2.1 Preprocessing data

The use of data in this study is divided into two stages, i.e., the training stage for the input and the stage for testing the built input architecture. Before the data was processed by the LSTM model, the data was normalized first as part of the preprocessing with the Min-Max Scaler method. The preprocessing was divided into several phases as follows:

(i) The range of each attribute was paid attention to. The attribute range was converted to equal intervals. The minimum feature was made equal to zero and the maximum feature was equal to one.

(ii) The data was expanded to the desired scale. The ranges were provided in the form of tuples as minimum and maximum features.

(iii) If it was false, the in-place scaling was completed. If it was true, a copy was created instead of in-place scaling and the scaled data was truncated to the provided feature range.

The data preprocessing aims to examine the effect on the MAPE value generated by the LSTM method. The results can be seen in Table 4.

Table 4. Results of data preprocessing

No.	Ginger Crop Yield (kg)	No.	Ginger Crop Yield (kg)	No.	Ginger Crop Yield (kg)
…	…	…	…	235	0.59662
27	0.31699	210	0.81295	236	0.17268
28	0.63391	211	0.66261	237	0.72961
29	0.59737	212	0.23870	238	1.00000
30	0.51649	213	0.62090	239	0.50321
31	0.28596	214	0.08415	240	0.89732
32	0.39291	215	0.00568	241	0.73322
33	0.62201	216	0.62330	242	0.45716
34	0.82280	217	0.73731	243	0.59505
35	0.59071	218	0.37844	244	0.20876
36	0.17367	219	0.60043	245	0.91752
37	0.53625	220	0.70490	246	0.56453
38	0.77733	221	0.95403	247	0.07478
39	0.02002	222	0.15650	248	0.86617
40	0.58919	223	0.28061	249	0.84231
41	0.18197	224	0.15115	250	0.54368

3.2.2 LSTM model

The LSTM architecture consists of an input layer, an output layer, and a hidden layer. The hidden layer consists of memory cells, with three gates in one cell, namely, the input gate, forget gate, and output gate, as shown in Figure 3. The input gate functions to control how much information must be stored in the cell state. The forget gate controls how far the value remains in the memory cell. The output gate serves to decide how much content or value is in a memory cell, and is used to calculate the output. The forecasting values were obtained using the LSTM method in the following steps:

(i) The value of f_t was determined. The forget gate was the first gate passed, which recorded how much cell status C_{t-1} was from the previous time and returned to the cell status C_t from the current time. The gate produced a value between 0 and 1 based on h_{t-1} and x_t, with an output of 0. Then the information was considered no longer useful and deleted. Conversely, if the output was 1, then the information was stored for future use.

(ii) The number of network inputs at the current time (x_t) was determined. It returned to the cell state C_t which was input with two activation functions (sigmoid and tanh) to select the part to be updated.

(iii) The information was updated to the cell state of the new candidate $\hat{C}_t$ created via the tanh layer, aiming to control how much new information was added.

(iv) The old cell state C_t was updated to become a new cell state by multiplying the old state by f_t to delete the information determined on the f_t layer.

(v) What was produced by O_t, which passed through the tanh neuron layer, was determined. The result was h_t, which subsequently influenced the cell state.

(vi) After the value of O_t was obtained, the cell state was placed through the tanh and multiplied by the gate and sigmoid layer output to produce the value h_t.

The modeling of the LSTM network training process can be seen in Table 5. The forget gate, cell state, output gate, and sigmoid values can be seen from the table.

Table 5. Result of the training process with LSTM forecasting ginger crop yields

$h_{t-1}=0$	Week				f_t	i_t	$\hat{C}_t$	o_t	h_t
$c_{t-1}=0$	I	II	III	IV					
Data I	1	0.69	0.55	1	0.93	0.97	0.98	0.91	0.68
U_f	0.65	0.35	0.8	0.7					
Data II	0.15	0.39	0.10	0.16	0.67	0.86	0.71	0.66	0.56
U_i	0.34	0.63	0.74	0.95					
Data III	0.56	0.57	1	0.56	0.90	0.97	0.95	0.84	0.81
U_c	0.85	0.77	0.23	0.45					
Data IV	0.84	0.64	0	0.84	0.88	0.97	0.97	0.88	0.88
U_o	0.95	0.13	0.25	0.6					
Data V	0	0.25	0.97	0	0.75	0.90	0.64	0.65	0.64

3.2.3 Tests on data testing

There are several LSTM trial scenarios, as shown in Table 6. This learning rate test aims to obtain the optimal total learning rate for the MAPE and RMSE values from training and testing. The results of testing the learning rate parameter with a max epoch of 10 are shown in Table 7. Meanwhile, the number of max epochs was pushed, which aims to find out whether, before the max epoch, the MAPE has exceeded the limit or stopped at the specified max epoch. The max epoch numbers used are 10, 50, 100, 500 and 1000. The max epoch test results are shown in Table 8. When the max epoch number is 100, the optimal max epoch number with the smallest MAPE and RMSE is 38.99% and 1244.85432 with an estimated time of 8.726145 seconds. After obtaining an initial estimate of the LSTM from model identification, the parameters of the candidate LSTM model were estimated using the significant parameters generated at the model identification stage. A p-value is considered significant if it falls below the 5% significance level. The MAPE and RMSE values are the smallest compared with other models.

Table 6. Trial scenario of the LSTM method

No.	Changes to Parameters	Description
1	Learning rate	Changes in learning rate (alpha) consist of 0.0001, 0.0002, 0.001, 0.002, 0.01 and 0.02.
2	Max epoch	Max epoch changes consist of 10, 50, 100, 500, and 1000.
3	LSTM units	Changes in LSTM units consist of 50, 75, and 100.

Table 7. Scenario results of LSTM based on learning rate

No.	LSTM Units	Learning Rate	Epoch	MAPE (%)	Time (Second)
1	50	0.002	10	45.91	1.761332
2	50	0.001	10	45.10	1.655466
3	75	0.02	10	43.46	1.724343
4	75	0.01	10	43.21	1.872740
5	100	0.0002	10	41.28	1.811447
6	100	0.0001	10	51.11	2.084307

Table 8. Scenario results of LSTM based on epoch and stationarity test results

No.	Learning Rate	Epoch	MAPE (%)	RMSE	p-Value	Time (Second)
1	0.0002	100	38.99	1244.85432	0.032	8.726145
2	0.0002	50	43.04	1872.63215	0.157	5.191170
3	0.0002	10	45.39	1873,85456	0.410	1.681532
4	0.0002	500	47.07	1251.56412	0.173	41.47692
5	0.0002	1000	51.47	1411.84513	0.130	90.93132

4. DISCUSSION

This study uses actual data from ginger yields from 2015 to 2019 to produce precise and accurate predictions for the future. A comparison between the DES and LSTM methods shows that both methods have the same advantage in forecasting, i.e., high accuracy. Before making predictions, the actual data was converted into a time series by preprocessing the data. Then ginger yields were predicted using both DES and LSTM. Several trials were carried out, including testing the DES model in two stages, i.e., testing the α value when the β value was fixed and vice versa to find the best α and β. This is shown in Table 3. The test results show that the MAPE value tends to be smaller when the α parameter value is large and the β parameter value is small. This shows that the DES method has the advantage on relatively little data. However, the weakness is that it should be maintained continuously, checked routinely and updated if there are bugs or new important features are added. Meanwhile, the tests carried out by the LSTM model in this study are shown in Table 6, where there are several parameters, including learning rate, epoch, and LSTM units. The initial weight at each LSTM gate was determined randomly and the initial bias was updated to achieve a loss function at a local minimum with the learning rate. Meanwhile, epoch and LSTM units were used to indicate the number of iterations and dataset size in LSTM training to produce a smaller MAPE. The analysis results using the LSTM model show that LSTM can be carried out for long-term time series data because it has memory to store information that will be reused in the calculation process for the next gate. In addition, the prediction results cannot be sufficiently compared with the original data. However, LSTM has a weakness, i.e., reliance on gradient values, where the weights cannot be updated for the next process if the gradient value is lost.

Table 9. Comparison between the results of the DES and LSTM methods and the actual data

No.	Date	DES Forecasting	LSTM Forecasting	Actual
1	2020-07-15	29563.80	27901.82	14886
2	2020-07-22	29684.35	27768.99	46048
3	2020-07-29	29804.90	28270.50	30132
4	2020-08-05	29925.45	27936.30	23869
5	2020-08-12	30046.11	27901.13	45119
6	2020-08-19	30166.55	28224.53	42309
7	2020-08-26	30287.10	28169.11	36354
8	2020-09-02	30407.65	28051.41	19563
9	2020-09-09	30528.20	27937.57	34702
10	2020-09-16	30648.75	28118.33	13441
11	2020-09-23	30769.30	27794.12	10333
12	2020-09-30	30889.85	27809.19	34797
13	2020-10-07	31010.41	28079.68	39313
14	2020-10-14	31130.96	28040.68	25098
15	2020-10-21	31251.51	27758.23	33891
16	2020-10-28	31372.06	27833.81	38029
17	2020-11-04	31492.61	27841.71	47897
18	2020-11-11	31613.16	28204.34	16307
19	2020-11-18	31733.71	27713.15	21223
20	2020-11-25	31854.26	27952.60	16095
21	2020-12-02	31974.81	27984.45	36303
22	2020-12-09	32095.36	28196.06	37905
23	2020-12-16	32215.91	27932.35	47622

According to the results of the trial scenarios, the yield prediction data produced by the ginger yield forecasting system in Table 9 is higher predicted with DES than predicted

with LSTM. The forecasting results can also be seen in Figure 5, with actual ginger yield data shown on the graph with a yellow line, DES prediction data with a blue line and LSTM prediction data with a red line. This shows that the LSTM method provides very good performance in this crop yield prediction model, even better than DES. However, LSTM in data processing requires quite a long time, as is evidenced in Table 8, where LSTM has a gate to regulate the entry and exit of processes that work with entry and exit gates. This shows that LSTM has a high complexity. After processing the data, the model was evaluated using MAPE and RMSE. The evaluation results of the LSTM model show that 38.99% takes 8.726145 seconds while the DES model produces an error rate of 43.49% with 0.00004 seconds, as shown in Table 3. Apart from that, it can be seen from the graph that the ginger harvest in July-December has increased, i.e., in the range of 31,000 tons. Meanwhile, the lowest demand occurs in January-June, with around 27,000 tons for both methods. Therefore, farmers are advised to plant ginger before June so that it can be harvested in July to meet market demand. According to the results of the trials, it can be concluded that this forecasting system provides more information for ginger farmers about the estimated future ginger harvest, thereby enabling farmers to handle the ginger planting in certain months more effectively.

According to the trial results in this study and literature review, it can be found that statistical and deep learning methods can be used to process time series data. Each has a relative effectiveness level, depending on the complexity of the data used. The difference between both methods is the use of time in processing data. In addition, deep learning methods tend to require more time than statistical methods. Therefore, further research could utilize several statistical methods to predict data with a short processing time. The LSTM and DES deep learning algorithms are two different prediction models and are suitable for performing time series analysis to improve performance and accuracy in forecasting.
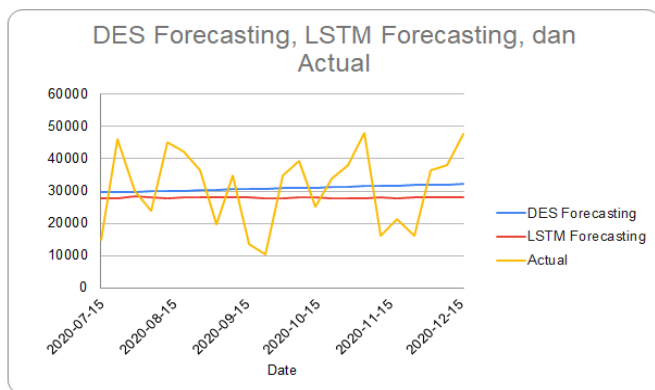

Figure 5. Graph of forecasting result comparison using the DES and LSTM methods with the actual data

5. CONCLUSIONS

Based on the research results and discussion of ginger crop yield prediction with a comparison of the LSTM and DES methods, conclusions can be drawn. Based on the prediction results using the DES method from ginger crop yield data for 2015-2019, the best value was obtained with MAPE 43.49% and RMSE 12997.34261 in parameters α=0.4 and β=0.1. The predicted results increased from July 2020 to December 2020

compared to the previous period. Based on the LSTM method forecasting results, with 74% training data, timestamp 10, and 100 hidden layer neurons, the MAPE value obtained is 38.99% at a learning rate of 0.0002 and RMSE of 1244.85432. After making predictions using these two methods, the LSTM method is the best one for this data compared to the DES method. This can be seen from the smaller error (MAPE) and RMSE values produced by LSTM compared to DES. Based on forecasting results using the DES and LSTM models with actual data, there is an upward trend. Therefore, it can be concluded that ginger production from 2019 to 2020 will increase by around 14% per year. Therefore, farmers can estimate future harvest plans by paying attention to the availability and price trends of ginger seeds in a certain month or at least three months beforehand. This affects the enthusiasm of farmers to continue to increase the productivity of ginger plants, especially in Madura.

REFERENCES

[1] Moiseev, G. (2021). Forecasting oil tanker shipping market in crisis periods: Exponential smoothing model application. The Asian Journal of Shipping and Logistics, 37(3): 239-244. https://doi.org/10.1016/j.ajsl.2021.06.002

[2] Anshory, M.I., Priyandari, Y., Yuniaristanto, Y. (2020). Peramalan penjualan sediaan farmasi menggunakan long short-term memory: Studi kasus pada apotik suganda. Performa: Media Ilmiah Teknik Industri, 19(2). https://doi.org/10.20961/performa.19.2.45962

[3] Chung, M.G., Kim, S.K. (2013). Efficient jitter compensation using double exponential smoothing. Information Sciences, 227: 83-89. https://doi.org/10.1016/j.ins.2012.12.008

[4] Siregar, S.A., Wibawa, A.P. (2018). Double exponential-smoothing neural network for foreign exchange rate forecasting. In 2018 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT), Makassar, Indonesia, pp. 118-122. https://doi.org/10.1109/EIConCIT.2018.8878591

[5] Regina, T., Jodiawan, P. (2021). Proposed improvement of forecasting using time series forecasting of fast moving consumer goods. Journal of Industrial Engineering and Management Systems, 14(1). http://doi.org/10.30813/jiems.v14i1.2418

[6] Al-Abri, E.S. (2016). Modelling atmospheric ozone concentration using machine learning algorithms. Doctoral dissertation, Loughborough University.

[7] Shankar, S., Ilavarasan, P.V., Punia, S., Singh, S.P. (2020). Forecasting container throughput with long short-term memory networks. Industrial Management & Data Systems, 120(3): 425-441. https://doi.org/10.1108/IMDS-07-2019-0370

[8] Bathla, G. (2020). Stock price prediction using LSTM and SVR. In 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC) Waknaghat, India, pp. 211-214.

https://doi.org/10.1109/PDGC50313.2020.9315800

[9] Agarwal, S., Tarar, S. (2021). A hybrid approach for crop yield prediction using machine learning and deep learning algorithms. In Journal of Physics: Conference Series. IOP Publishing, 1714(1): 012012. https://doi.org/10.1088/1742-6596/1714/1/012012

[10] Madhika, Y.R., Kusrini, K., Hidayat, T. (2023). Gold price prediction using the ARIMA and LSTM models. Sinkron: Jurnal Dan Penelitian Teknik Informatika, 7(3): 1255-1264. https://doi.org/10.33395/sinkron.v8i3.12461

[11] Bai, W., Fanghua, L., Yan, H., Yun, T. (2011). Application partial least squares regression in the analysis of maize regulated deficit irrigation. In 2011 International Conference on New Technology of Agricultural, Zibo, China, pp. 437-440. https://doi.org/10.1109/ICAE.2011.5943835

[12] Jin, X., Xu, X. (2012). Rmote sensing of leaf water content for winter wheat using grey relational analysis (GRA), stepwise regression method (SRM) and partial least squares (PLS). In 2012 First International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Shanghai, China, pp. 1-5. https://doi.org/10.1109/Agro-Geoinformatics.2012.6311706

[13] Altilar, D.T., Terliksiz, A.S. (2018). Comparison of statistical methods for predicting wheat yield trends in Turkey. In 2018 7th International Conference on Agro-geoinformatics (Agro-geoinformatics), Hangzhou, China, pp. 1-4. https://doi.org/10.1109/Agro-Geoinformatics.2018.8476125

[14] Aini, N.N., Iriany, A., Nugroho, W.H., Wibowo, F.L. (2022). Comparison of adaptive holt-winters exponential smoothing and recurrent neural network model for forecasting rainfall in Malang City. ComTech: Computer, Mathematics and Engineering Applications, 13(2): 87-96. https://doi.org/10.21512/comtech.v13i2.7570

[15] Meena, S.S., Peram, N.H., Sharma, A., Surliya, V. (2022). Price forecasting of tomato by using moving average forecasting model, simple exponential smoothing forecasting model and Arima model. Journal of Management & Entrepreneurship, 16(4(II)): 25-37.

[16] Bharati, R.C., Singh, A.K. (2019). Predicting rice production using autoregressive integrated moving average model. Journal of AgriSearch, 6(4): 205-210. https://doi.org/10.21921/jas.v6i04.16905

[17] Asrul, B.E.W. (2022). Implementasi metode double exponential smoothing untuk prediksi hasil panen sayuran kentang. Jurnal Fokus Elektroda: Energi Listrik, Telekomunikasi, Komputer, Elektronika dan Kendali), 7(3): 193-199. https://doi.org/10.33772/jfe.v7i3.9

[18] Afiyah, S.N., Kurniawan, F., Aqromi, N.L. (2021). Rice production forecasting system in East Java using double exponential smoothing method. Procedia of Engineering and Life Science, 1(2). https://doi.org/10.21070/pels.v1i2.988

[19] Bhimavarapu, U., Battineni, G., Chintalapudi, N. (2023). Improved optimization algorithm in LSTM to predict crop yield. Computers, 12(1): 10. https://doi.org/10.3390/computers12010010

[20] Tian, H., Wang, P., Tansey, K., Zhang, J., Zhang, S., Li, H. (2021). An LSTM neural network for improving wheat yield estimates by integrating remote sensing data and meteorological data in the Guanzhong Plain, PR China. Agricultural and Forest Meteorology, 310: 108629. https://doi.org/10.1016/j.agrformet.2021.108629

[21] Elpawati, E., Wirhanti, P.E., Aisyah, S.N. (2022). Forecasting Indonesia's ginger export with major competing countries in the international market. Anjoro: International Journal of Agriculture and Business, 3(2): 73-80. https://doi.org/10.31605/anjoro.v3i2.2061

[22] Das, K.R., Jahan, M., Noorunnahar, M. Performance of trend models of ginger production. Journal of Bioscience and Agriculture Research, 30(01): 2508-2512. https://doi.org/10.18801/jbar.300123.302

[23] Rianti, A.T., Bafadal, A., Abdi, A. (2023). The forecasting analysis of rice production and sufficiency consumption of rice (Oriza sativa) in Konawe District. Jurnal Ilmiah Membangun Desa dan Pertanian, 8(3): 96-101. https://doi.org/10.37149/JIMDP.v8i3.131

[24] Estiningtyas, W., Surmaini, E., Suciantini, Susanti, E., Mulyani, A., Kartiwa, B., Sumaryanto, Perdinan, Apriyana, Y., Alifia, A.D. (2024). Analysing food farming vulnerability in Kalimantan, Indonesia: Determinant factors and adaptation measures. PLoS ONE ONE 19(1): e0296262. https://doi.org/10.1371/journal.pone.0296262

[25] Anggraeni, W., Aristiani, L. (2016). Using Google trend data in forecasting number of dengue fever cases with ARIMAX method case study: Surabaya, Indonesia. In 2016 International Conference on Information & Communication Technology and Systems (ICTS), Surabaya, Indonesia, pp. 114-118. https://doi.org/10.1109/ICTS.2016.7910283

[26] Pangestuti, D.C., Pasaribu, R.F. (2021). Analysis forecasting sales of tart products. Inovasi: Jurnal Ekonomi, Keuangan, dan Manajemen, 17(4): 792-801. https://doi.org/10.30872/jinv.v17i4.10180

[27] Daniawan, B. (2019). Evaluation of lecturer teaching performance using AHP and SAW methods. Journal Bit-Tech, 1(2): 74-83. https://doi.org/10.32877/bt.v1i2.41

[28] Naim, I., Mahara, T., Idrisi, A.R. (2018). Effective short-term forecasting for daily time series with complex seasonal patterns. Procedia Computer Science, 132: 1832-1841. https://doi.org/10.1016/j.procs.2018.05.136

[29] Mukhairez, H.H.A. (2019). Medium-Term forecasting for electricity end use. International Journal of Intelligent Computing Research (IJICR), 10(1): 967-970. https://doi.org/10.20533/ijicr.2042.4655.2019.0117

[30] Cruz-Nájera, M.A., Treviño-Berrones, M.G., Ponce-Flores, M.P., Terán-Villanueva, J.D., Castán-Rocha, J.A., Ibarra-Martínez, S., Santiago, A., Laria-Menchaca, J. (2022). Short time series forecasting: Recommended methods and techniques. Symmetry, 14(6): 1231. https://doi.org/10.3390/sym14061231

[31] Wang, S., Chaovalitwongse, W.A. (2011). Evaluating and comparing forecasting models. Wiley Encyclopedia of Operations Research and Management Science, eorms0307. https://doi.org/10.1002/9780470400531.eorms0307

[32] Zellner, M., Abbas, A.E., Budescu, D.V., Galstyan, A. (2021). A survey of human judgement and quantitative forecasting methods. Royal Society Open Science, 8(2): 201187. https://doi.org/10.1098/rsos.201187

[33] Ebrahim, S.A., Poshtan, J., Jamali, S.M., Ebrahim, N.A. (2020). Quantitative and qualitative analysis of time-series classification using deep learning. IEEE Access, 8: 90202-90215.

https://doi.org/10.1109/ACCESS.2020.2993538

[34] Saha, A., Sinha, K. (2020). Usage of Holt's linear trend exponential smoothing for time series forecasting in agricultural research. Not Available. https://www.researchgate.net/publication/345413376_U sage_of_Holt's_Linear_Trend_Exponential_Smoothing _for_Time_Series_Forecasting_in_Agricultural_Resear ch.

[35] Nurfaizah, Hariguna, T., Romadon, Y.I. (2019). The accuracy comparison of vector support machine and decision tree methods in sentiment analysis. In International Conference on Engineering, Technology and Innovative Researches, Purwokerto, Indonesia. https://doi.org/10.1088/1742-6596/1367/1/012025

[36] Febriyanti, S., Pradana, W.A., Saputra, M.J., Widodo, E. (2021). Forecasting the consumer price index in Yogyakarta by using the double exponential smoothing method. Parameter: Journal of Statistics, 2(1): 1-7. https://doi.org/10.22487/27765660.2021.v2.i1.15641

[37] Nifa, K., Boudhar, A., Ouatiki, H., Elyoussfi, H., Bargam, B., Chehbouni, A. (2023). Deep learning approach with LSTM for daily streamflow prediction in a semi-arid area: A case study of Oum Er-Rbia River basin, Morocco. Water, 15(2): 262. https://doi.org/10.3390/w15020262

[38] Sudriani, Y., Ridwansyah, I., A Rustini, H. (2019). Long short term memory (LSTM) recurrent neural network (RNN) for discharge level prediction and forecast in Cimandiri river, Indonesia. In IOP Conference Series: Earth and Environmental Science. IOP Publishing, 299: 012037. https://doi.org/10.1088/1755-1315/299/1/012037

[39] Sarker, I.H. (2021). Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. SN Computer Science, 2(6): 420. https://doi.org/10.1007/s42979-021-00815-1

[40] Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena, 404: 132306. https://doi.org/10.1016/j.physd.2019.132306

[41] Yu, Y., Si, X., Hu, C., Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. Neural Computation, 31(7): 1235-1270. https://doi.org/10.1162/neco_a_01199

[42] Chang, Y.S., Chiao, H.T., Abimannan, S., Huang, Y.P., Tsai, Y.T., Lin, K.M. (2020). An LSTM-based aggregated model for air pollution forecasting. Atmospheric Pollution Research, 11(8): 1451-1463. https://doi.org/10.1016/j.apr.2020.05.015

[43] Peñaloza, A.A., Leborgne, R.C., Balbinot, A. (2022). Comparative analysis of residential load forecasting with different levels of aggregation. Engineering Proceedings, 18(1): 29. https://doi.org/10.3390/engproc2022018029

[44] Urolagin, S., Sharma, N., Datta, T.K. (2021). A combined architecture of multivariate LSTM with Mahalanobis and Z-Score transformations for oil price forecasting. Energy, 231: 120963. https://doi.org/10.1016/j.energy.2021.120963

[45] Ariqoh, A.S., Nisfullaili, J., Salsabila, N.P., Prianjani, D. (2023). Selection of the best newspaper forecasting method using holt-winters and long short term memory method. In 12th Annual International Conference on Industrial Engineering and Operations Management, pp. 2836-2845. https://doi.org/10.46254/an12.20220525

[46] Khairina, D.M., Daniel, Y., Widagdo, P.P. (2021). Comparison of double exponential smoothing and triple exponential smoothing methods in predicting income of local water company. In Journal of Physics: Conference Series. IOP Publishing, 1943(1): 012102. https://doi.org/10.1088/1742-6596/1943/1/012102