Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/130118

Time Series Forecasting of New Cases for COVID-19 Pandemic in Jordan Using Enhanced Hybrid EMD-ARIMA

Ahmad M. Awajan[1,], Bilal Al-Hasanat[1], Elarbi Elkaroui[2], Ayed AL e'damat[1], Remal S. Al-Gounmeein[1], Abdullah S. Al-Jawarneh[3], Heba N. Ayyoub[4], and Eyad AlFarajat[1]*

[1]Department of Mathematics, Al Hussien bin Talal University, Ma'an, Jordan
[2]BCM/IEM Program, Faculty of Resilience, Rabdan Academy, Abu Dhabi, UAE
[3]Department of Mathematics, Faculty of Science, Jerash University, Jerash, Jordan
[4]Department of Mathematics, Faculty of Science and Technology, Irbid National University, Ibrid, Jordan

Received: 18 Feb. 2023, Revised: 27 May 2023, Accepted: 2 Jun. 2023.
Published online: 1 Jan. 2024.

Abstract: In this study, the enhanced hybrid empirical mode decomposition with autoregressive integrated moving average (EMD- ARIMA) method is proposed and applied to forecast daily new COVID-19 reported cases in Jordan. The EMD method is applied to decompose the COVID-19 data into a number of IMFs components as a simple time series. Then, the appropriate ARIMA(p,d,q) model is applied to evaluate the forecasting value for the low-frequency components. Then, the forecasting results are collected together. Data for this study are collected from the Jordanian Ministry of Health. Seven forecasting accuracy measures are employed to compare the forecasting results of the proposed technique with the results of seven forecasting methods. The comparison of forecasting results shows that the enhanced EMD-ARIMA method is better than selecting forecasting methodologies in COVID-19 data.

Keywords: COVID-19, ARIMA Models, EMD, Combined Forecasting Method.

1 Introduction

Day after day, COVID-19 is increasing the concern of doctors, governments of countries, and their people since its spread. Meanwhile, this pandemic has had an impact on government and accounting bodies and the capacity of state employees during the COVID-19 [1]. Since COVID-19 was declared an international pandemic by the world health organization, has become the most influential disease in the world [2]. Also, this disease has a great impact on those infected with it, even after recovery from COVID-19. Moreover, these patients require long-term medical and psychosocial support programs even after recovery from this disease [3]. So that it has become the most important issue that is discussed in meetings of the high policies of the government and international meetings [4].

As this epidemic has worked to change the internal and external policies of the countries of the world. One of these countries is the state of Jordan. There are some factors affecting the number of daily injuries in Jordan. One of these factors that are expected to reduce the number of daily infections is receiving the vaccine for this pandemic [5]. Moreover, the carefully studied closures by the government helped avoid reaching a dangerous stage for the health system in Jordan. Based on that concept, this study will study the behavior of historical data for this epidemic in Jordan. Forecasting the number of new cases would be a useful way to forecast costs and services in the future [6]. The EMD-ARIMA model will be developed, and then this hybrid model employed to improve the forecasting of future COVID-19 cases in Jordan. Based on the accurate prediction results of the proposed method, researchers will resort to reliable prediction methods to be used in predicting future epidemics. In order to contribute to the speed of response of decision makers. Accordingly, their actions will be more effective and reduce the social and economic impact.

As will be presented in Part 5, the used data in this study are non-stationary and nonlinear. Accordingly, the EMD method will be applied, because this method is one of the best methods to deal with this type of data [7]. Moreover, ARIMA technique is the most widely used time series forecasting method in literature, due to the accuracy of its results [8]. Therefore, these two methods were combined to overcome the non-stationary and non-linear characteristics of the data and obtain accuracy in the forecasting results.

2 Related Works

In literature, several researchers have applied different forecasting methodologies to forecast the COVID-19 data. Such as [9] proposed methods with a new technique, which is to use a data augmentation way to generate new data with keep the same attributes to the input data. In general, the proposed new methods (LSTM-Aug, LSTM-Aug, and GRU-Aug) performance which is applied to the new idea was significantly superior to the standard methods (LSTM, LSTM, and GRU)

[*]Corresponding author e-mail: ahmad.m.awajan@ahu.edu.jo

using COVID-19 data extracted from ten countries.

In [10], the ARIMA technique has been employed to forecast the COVID-19 epidemic behavior in Italian ten days ahead. This technique has been established based on the complete periodic data of COVID-19 epidemic behavior in Hubei, China. This technique is more advisable for short-term forecasting. In [11], the Error Trend Seasonal (ETS) model has been used to forecast the new cases and deaths in India. In [12], the ARIMA and SARIMA forecasting methods have been employed to forecast the COVID-19 cases (confirmed, recovered, and deaths) in 16 countries. Based on the AIC, BIC, MSE, MAE, RMSE, and MAP. A comparison has been achieved between these two methods. It is found, in some cases, these are more or less the same. In other cases, the SARIMA method is more accurate than the ARIMA method. In [13], the authors improved the Long short-term memory networks forecasting method to expect the COVID-19 cases in Canada. Also, they compared Canada's transmission rates with Italy and the USA.

3 Methodology

In this part of this study, the different methods for the implementation of the new proposed forecasting technique will be presented. Which are EMD, ARIMA, and Fourier transform. Also, we will present an overview of the forecasting methods used to compare the new proposed method.

3.1 Empirical mode decomposition (EMD)

The EMD method was introduced in [14]. The EMD effective method for processing and decomposing different types of non-stationary time series, the mode mixing problem is a limitation of the standard EMD algorithm [15]. One of our contributions to this study is that the EMD-ARIMA methodology is able to overcome these limitations. The EMD technique was employed in several areas as shown in [16]. The EMD method deals with the nonlinear and non-stationary time series [17], to decompose this time series into various simple time series [18]. Furthermore, the EMD decomposes the time series with a reserve for the time domain of the data. This attribute provides a powerful and adaptive process for decomposing a time series into a set of time series known as intrinsic mode functions (*IMF*) and residual [19]. As a result, the original time series can be built from the IMFs by using Equation (1).

$$x(t) = \sum_{i=1}^{n} IMF_i(t) + r(t) \tag{1}$$

where $x(t)$ is the original time series, IMF_i and $r(t)$ are i^{th} intrinsic mode function and residue, respectively. Which is the result of decomposition for the original time series by EMD.

To evaluate the *IMF*s, the sifting process algorithm is applied to the original data $x(t)$ [20]. This algorithm is presented in detail as follows:

1. Input the COVID-19 time series data $x(t)$ into the EMD algorithm, let the iteration value $i$is equal 1.
2. Determine the local extrema values for the $x(t)$.
3. Using the cubic spline line methods, the local maxima values are connected together, as a result, the local upper envelope function will be created, and it's denoted by $u(t)$.
4. Using the same last two processes, the local lower envelope function will be created from the local minimum values of $X(t)$, and it's denoted by $l(t)$.

Then the mean function will be determined from $u(t)$ and $l(t)$ using the following formula, and its denoted by $m(t)$;

$$m(t) = \frac{u(t) + l(t)}{2} \tag{2}$$

1. In this step, a new function $h(t)$ will be determined using $m(t)$ and the signal $x(t)$ by Equation (3)

$$h(t) = x(t) - m(t) \tag{3}$$

2. In this step, we check the following conditions on the function $h(t)$, which are called *IMF* conditions [21]; these are (a)

$$|N.Ex - N.C.Z| < 1, \tag{4}$$

And (b)

$$|m(t)| = \left|\frac{u(t) + l(t)}{2}\right| < \varepsilon, \tag{5}$$

where N.Ex means the number of local extreme points (local maxima and minima), while N.C.Z means the number of cross-zero points, and ε is a too-small non-negative value that approaches 0, sometimes equal to 0.

5. If the function $h(t)$ is satisfied the above conditions, then move to the next step. Otherwise, move to step 2 and we assume the function $h(t)$ is the original data, and apply steps 2 to 8 on the new original data.

6. As a result of the previous step, we have a new IMF; which is determined using Equation (6). And then, update the iteration value (by moving to the next integer) $i = i + 1$.

$$IMFi(t) = h(t). \tag{6}$$

7. In this step, we will define a new function, which is called the residue function $r_i(t)$; this function is evaluated by the Equation (7).

$$ri(t) = x(t) - IMFi(t) \tag{7}$$

8. Based on the result from the last step $r_{i+1}(t)$ we will tack one of the following decisions.

   (a) if $r_i(t)$ is a monotonic or constant function, that concludes this is the last process. Hence, we save the residue of the sifting process which is $r(t) = r_i(t)$, and we save all the IMFs that have been obtained.

   (b) If the residue component does not satisfy the monotonic (or constant) condition, go back to the second step.

The IMFs and the residual components of the original data are presented in Figure 1, while its boxplots are presented in Figure 2.

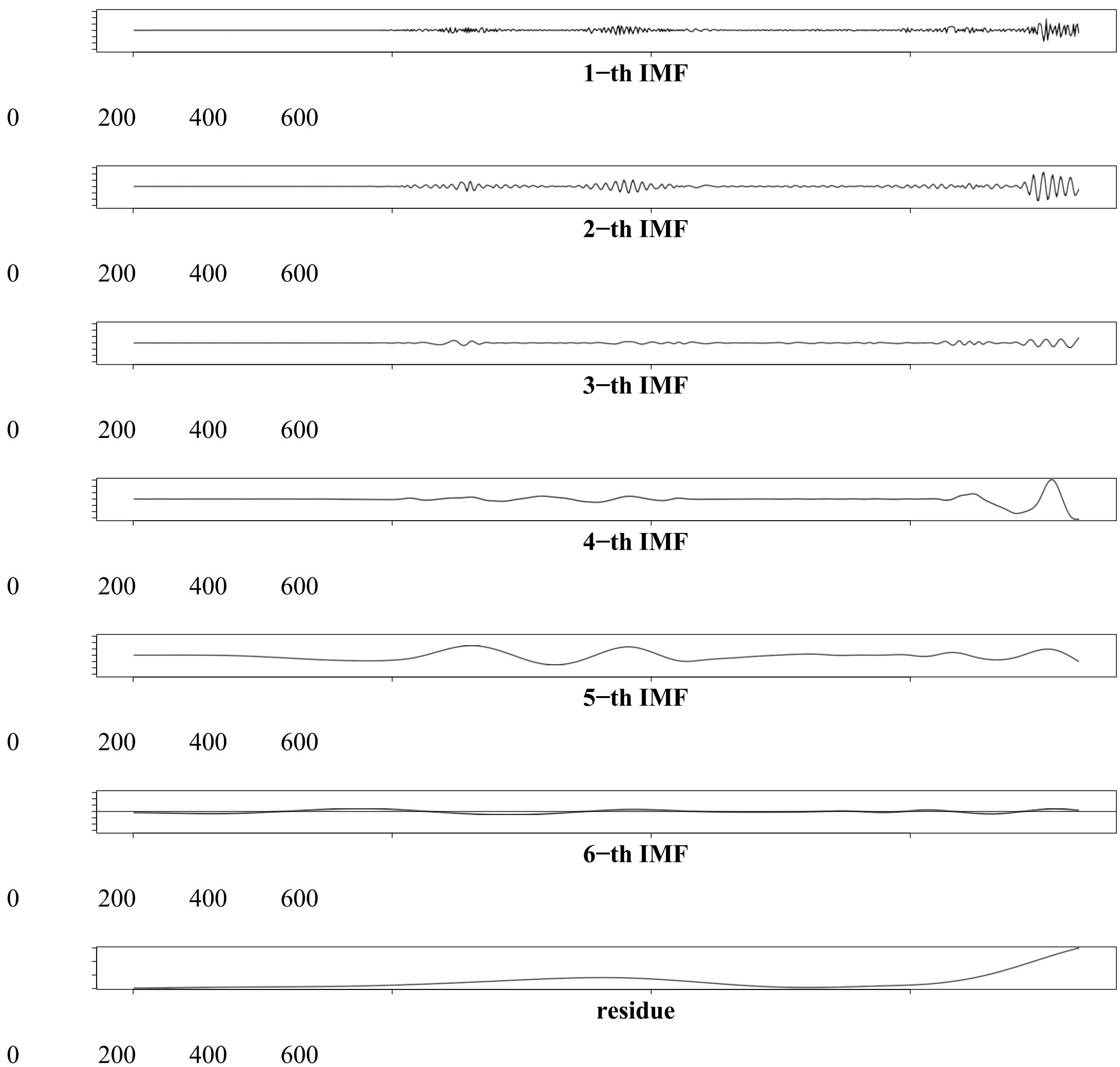


Fig.1: The IMF's and residual components of the original time series data.

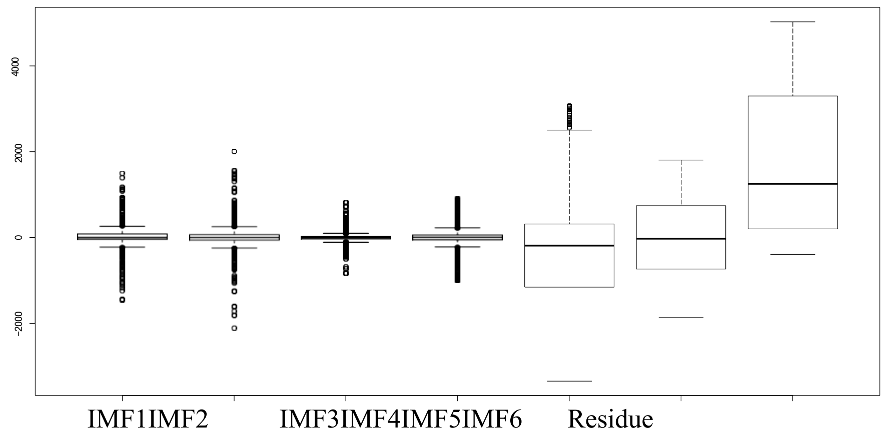


Fig.2: The box plot of IMFs and residue.

3.2 ARIMA methodology

In this study, ARIMA is used to forecast the daily COVID-19 new cases in Jordan. So, the ARIMA is presented in this section in detail. ARIMA or Auto-Regressive Integrated Moving Average model was introduced by [22]. The ARIMA method analyzes historical data to extract statistical information and then represents it as a linear aggregation of its own related historical values and historical errors. This linear formula is employed to forecast the future values of the time series data.

This method is represented as ARIMA (p, d, q) where $p, q \geq 1$ and $d \geq 0$. Such that, these symbols p and q are the degree for two statistical components techniques which are Auto Regressive and Moving Average, respectively. While d represents the degree of stationarity difference. If $d > 0$ this means the time series data is not stationary, which means, the primary time series has a trend and/or seasonality that must be removed [23]. The differencing degree d represents the number of different times that the time series needs to be stationary. To check if the time series is stationary or not, there are two useful tests that are employed in this study. These tests are the rolling statistics and the Augmented Dickey-Fuller test [24].

Mathematically, If we have a time series data X_t of real numbers and the integer t represents the index of X_t, then this time series X_t can be written ARIMA (p, q) model by the following formula:

$$X_t - \varphi_1 X_{t-1} - \cdots - \varphi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q},$$

or equivalently [25] by

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right) X_t = \left(1 - \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t \tag{8}$$

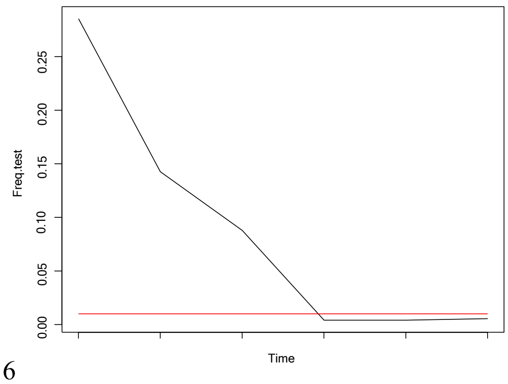
where L represents the lag operator, the φ_i represent the coefficients of the auto-regressive component of the method, the θ_i represent the coefficients of the moving average component and the ε_t represent the error expression. The error expression ε_t are usually possessing the white noise distribution (an independent, identically, and standard normal distribution) [26]. The maximum order of p is selected by using the partial auto-correlation function (*PACF*). Also, the maximum order of q is determined using the autocorrelation function (*ACF*) [27]. And the coefficients $\varphi_1, \ldots, \varphi_p$ and $\theta_1, \ldots, \theta_q$ are selected by using the Corrected Akaike's information criterion (*AICc*) by [28].

In literature, several studies employed the ARIMA models in forecasting. The result of these studies was accurate and useful. some of these studies are presented following. The authors in [29] seven-days forecast the dynamics of confirmed Covid-19 cases for several European countries using ARIMA (1,2,0). In [30], the ARIMA (3,2,6), ARIMA (4,2,6), ARIMA (0,2,1), and ARIMA (4,2,8) models are the dynamics of COVID-19 epidemic in four selected Indian states. Also, the ARIMA (4,2,7) model is applied to the dynamics of the COVID-19 epidemic in India. While, in [31] the ARIMA (2, 2, 2) model is applied to forecast the next 20 days of confirmed COVID-19 cases in India.

3.3Fourier transform

The Fourier transform (FT) is the conversion of the signal of data from the original domain to the frequency domain representation [32]. In this study, the fast Fourier transform (FFT) has been employed to estimate the spectral density of the IMFs time series components. The FFT is a mathematical technique for computing the finite discrete FT of signal data [33]. The discrete FT of finite length can be written as the following mathematical formula [34]. Given that $y_0, \ldots y_{N-1}$ be the N data points of a time series. Then, this time series can be written by discrete FT as the following equation.

$$Y_k = \sum_{n=0}^{N-1} y_n e^{-i2\pi k\frac{n}{N}} \; ; \; k = 0,\ldots,N-1 \tag{9}$$



12 3 4 5 6

Fig.3: The Frequency for the IMFs components with the threshold criteria equal to 0.01

Figure 3 presents the frequency for the IMFs components with the threshold criteria equal to 0.01 by the red horizontal line. Based on this line, the IMF_1, IMF_2, and IMF_3 are high frequencies. While the IMF_4, IMF_5, and IMF_6 are low- frequencies.

3.4 *Statistical techniques for consideration method*

In this study, the enhanced hybrid EMD-ARIMA method is compared with eight forecasting methods. These methods are ETS, HW, ARIMA, RW, STS, the taf, and EMD-HW. These forecasting methods were selected according to their degree of accuracy in forecasting time series data and their capacity to deal with the salient attribute of the data.

ETS is the exponential smoothing algorithm, the ETS method has been employed to forecast time series in several studies such as [35, 36, 37]. HW is the Holt-Winters forecasting method [38]. The HW was used in a number of forecasting studies, and the performance of the HW forecasting model was outstanding in the performance metrics used in [39, 40]. STS has been compared with other forecasting methods, and the performance of the STS method with the other methods was equally well [41]. The ARIMA models are used in order to validate the forecasting performance of the proposed forecasting model. RW is the random walk forecasting method, this method has been employed to improve forecasting results in several studies such as [42, 43, 44].

4 Enhanced hybrid of Empirical Mode Decomposition with autoregressive integrated Moving Average Model

In this part of the study, we will present the methodology of the proposed method in detail. The methodology of the Enhanced hybrid of EMD-ARIMA consists of six steps. Figure 4 presents the flow chart Enhanced hybrid of EMD-ARIMA.

1. The EMD is applied on the COVID-19 Pandemic in Jordan, this data is denoted by x_t. As a result of this step, six IMFs components with one residue will be obtained. these components are presented in Figure 1.

2. The Fourier transform is employed to transform the IMFs components from the time domain into the frequency domain. The results of this step are presented in Figure 3.

3. Based on the IMFs frequency, these components are divided into two sets. These sets are high frequency and low frequency with 0.01 threshold criteria.

4. The ARIMA method is applied on each of the low-frequency components and residual to forecast 6 points ahead.

5. Then, the forecasting results of the high-frequency IMF and residual are aggregated to get the final forecasting output.

6. The result of the proposed method is compared with the results of eight forecasting methods using seven forecasting accuracy measurements.

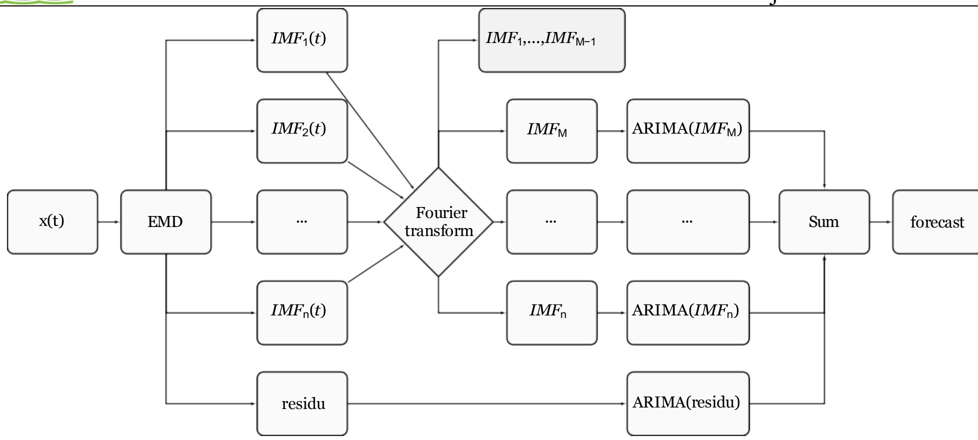


Fig.4: The Enhanced hybrid of the EMD-ARIMA flowchart.

5 Data

In this study, the COVID-19 time series data of Jordan are used. These data cover the period from 2/3/2020 (the date of the first case was shown in Jordan) to 26/3/2022 (the date for this study is done). Table 1 presents the basic statistics of the original data. This interval covers around two years (two time series periods). Figure 5 presents the original data with 728 observations. Figure 6 presents the box plot of the original data, which shows the density of this data is positively skewed. Moreover, the KPSS (Kwiatkowski-Phillips-Schmidt-Shin [45]), RESET (Ramsey Regression Equation

Table 1: The basic statistics of the original data.

Mean	Median	Min	Max	S.Deviation	Skewness	Kurtosis
2250.41	1021.5	0	22720	3475.5	3.11	11.89

Specification Error Test [46]), BP (Breusch-Pagan test [47]) has been applied on the original data. Based on the *p*-value of these tests, the time series is significantly non-linear and non-stationary, it has a high heteroscedasticity attribute also. Based on the mathematical pattern and statistical attributes of the new cases of the COVID-19 pandemic data in Jordan. This time series data significantly suitable for testing the performance of the proposed forecasting method in this study (the enhanced hybrid EMD-ARIMA). Moreover, no substantial work in the literature analyzed and forecast the new cases of the COVID-19 pandemic data in Jordan, while forecasting t the new cases of the COVID-19 pandemic data is one of the most important research topics at this time.

Daily COVID−19NewCasesinJordan

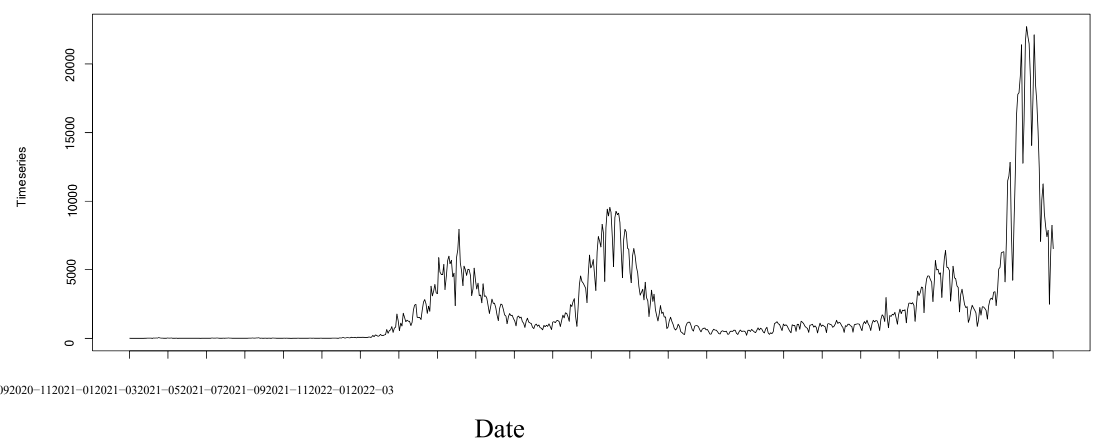


Fig.5: The plot of the original data

To evaluate the proposed forecasting method in this study, the time series is partitioned into two sets. The observations from day-1 into day$-m$ will be the first set. This data set is employed to select the suitable forecasting method. The second set is the observations from day$-(m+1)$ into day-N (the last observation). This data set is reserved for out-of-sample evaluation, which is employed to make a comparison of the accuracy with the selected forecasting methods. In this study, the observation number is $N = 728$, the first set is $m = 722$. The second set is $h = 6$.

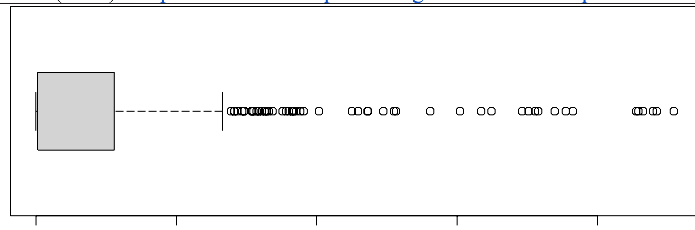


Fig. 6: The box plot of the original data

6 Result and Discussion

In this study, the number of new daily cases of COVID-19 in Jordan is employed to present the forecasting performance for the Enhanced hybrid of the END-ARIMA method. Seven forecasting methodologies are applied in order to confirm the forecasting performance accuracy of the newly proposed forecasting method (Enhanced hybrid of EMD-ARIMA). The first part of the proposed methodology is to input the original data into the EMD. As a result of this part are six IMF's components with one residual, these components and residual are presented in Figure 1. The second part of the proposed method is applying the FFT six IMF's components to estimate the maximum frequency value of each one. Then, we cluster IMF's components into high-frequency and low-frequency. The frequency of IMF's components is presented in Figure 3. As a result of this part, the IMFs 4, 5, and 6 are low-frequency IMFs components. The third part aims to apply the ARIMA technique to residual and low-frequency IMFs, to forecast the COVID-19 daily new cases in Jordan with six-day ahead. In this part, our object is to choose the fit ARIMA model of each low-frequency IMFs and residual. We go over all the possible models for each of these components and select the ARIMA model which has the lowest AIC-value. Table 2 presents the order and the AIC- value for the fit ARIMA Model of residual and IMFs components. The AIC value of a model is estimated by applying the Equation 10, such that; k is the number of estimated parameters in the model, $\hat{L}$ is the maximized value of the likelihood function for the model.

$$\mathrm{AIC} = 2k - 2\ln(\hat{L}) \qquad (10)$$

Table 2: The order and the AIC- value for the fit ARIMA Model of residual and IMFs components.

Component	ARIMA(p, d, q)	Suitable AIC -value
Residual	ARIMA(0,2,5)	−4098.39
IMF6	ARIMA(0,0,0)	11720.89
IMF5	ARIMA(0,1,4)	4061.32
IMF4	ARIMA(4,1,1)	2377.79

To show the performance of the proposed forecasting method (Enhanced Hybrid EMD-ARIMA) for the new cases of the COVID-19 pandemic in Jordan, forecasting evaluation methods will be used, which are deeply have been employed in the literature. Table 3 presents seven accuracy evaluations with their mathematical form. These formulas will be employed to estimate the forecasting error for each technique. Such that y represents the forecast amount of the actual value y at time period i from the input real-time series data.

Table 3: Error measures are used in study

	Name of measure error	Formula of measure error
1	Mean Error [48]	ME 1n y y
2	Root Mean Squared Error [49]	= n ∑(i− ^i)
3	Mean Absolute Error [50]	i=1 –
4	Mean Percentage Error	RMSE =q 1 ∑n (yi− y^i)2
5	Mean Absolute Percentage Error [51]	n i=1 MAE = 1 ∑n |yi− yi|
6	Mean Absolute Scaled Error [52]	MPE n i=1 ^ = 100% ∑n yi−y^i
7	Theil's U-statistic [53]	n i=1 yi

		$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{\lvert y_i-\hat{y}_i \rvert}{y_i} .100\%$ $MASE = \frac{1}{h}\sum_{i=1}^{h} \frac{\lvert y_i-\hat{y}_i \rvert}{\frac{1}{h-1}\sum_{i=2}^{h}\lvert y_i-y_{i-1} \rvert}$ $TheilU = \frac{\sqrt{\sum_{i=1}^{h-1}\left(\frac{\hat{y}_{i+1}-y_i}{y_i}-\frac{y_{i+1}-y_i}{y_i}\right)^2}}{\sqrt{\sum_{i=1}^{h-1}\left(\frac{y_{i+1}-y_i}{y_i}\right)^2}}$

Table 4 displays the accuracy of eight forecasting methods by seven forecasting measures. These functions are ME, RMSE, MAE, MPE, MAPE, MASE, and TheilU. While the forecasting models are the proposed method (Enhanced Hybrid EMD-ARIMA) with seven selected forecasting methods. These values (in Table 4) present the forecasting accuracy of eight methods for forecasting at $h = 6$ for the new daily COVID-19 cases time series data of Jordan. Based on the error values in Table 4, the proposed method (Enhanced Hybrid EMD-ARIMA) has achieved the minimum error values in each of the seven error functions in Table 4. For example, in the MAPE column values, the Enhanced Hybrid EMD-ARIMA method is equal to 23.979, which is the lowest value of this column (38.744, 38.728, 35.552, 38.817, 39.015, 38.809, 45.583, 23.979). As the rest of the columns have the same feature, the error value of the proposed forecasting method has the lowest value of error compared to the seven existing forecasting methods. These values (in Table 4) appear that the forecast accuracy for EMD-ARIMA is superior to the eight selected forecasting methods. This can suggest that the Enhanced Hybrid EMD-ARIMA model captures all of the patterns in the new cases of the COVID-19 pandemic in Jordan. This indicates the proposed forecasting method is very suitable to be implied to forecast this type of data, which is the new cases for COVID-19 pandemic in Jordan.

7 Conclusion

In this study, a new enhanced hybrid forecasting technique was presented. This technique contains the empirical mode decomposition (EMD) with the auto-regressive integrated moving average model (ARIMA); it's denoted by (EMD-ARIMA). This method was applied by one of the most critical data, which is the daily new COVID-19 cases in Jordan. Seven forecasting accuracy measures are employed in this study to compare the new enhanced hybrid technique with seven existing forecasting methods. The comparison showed that the results of the EMD-ARIMA forecasting method are more accurate than the results of the seven selected methods.

Conflict of interest

The authors declare that there is no conflict regarding the publication of this paper.

8 References

[1] N. A. Alessa., & H. S. Shalhoob. Medium-to Long-Term Socioeconomic Impact of Covid-19 in Arab Countries, 2022. Journal of Statistics Applications & Probability, **12**(2), 329-345 (2023)

[2] Salgotra, R., Gandomi, M., &Gandomi, A. H. Time series analysis and forecast of the COVID-19 pandemic in India using genetic programming. Chaos, Solitons & Fractals, **138**, 109945. (2020).

[3] Abuhammad, S., Khabour, O. F., Alzoubi, K. H., Hamaideh, S., Khassawneh, B. Y., Alsmadi, B. F., &Alqudah, S. Mental health of COVID-19 recovered individuals: A national study from Jordan. Electronic Journal of General Medicine, **20** (2): em451. (2023).

[4] Bertozzi, A. L., Franco, E., Mohler, G., Short, M. B., & Sledge, D. The challenges of modeling and forecasting the spread of COVID-19. Proceedings of the National Academy of Sciences, **117**(29), 16732-16738. (2020).

[5] Hussein, T., Hammad, M. H., Fung, P. L., Al-Kloub, M., Odeh, I., Zaidan, M. A., & Wraith, D. COVID-19 pandemic development in Jordan—short-term and long-term forecasting. Vaccines, **9**(7), 728. (2021).

[6] Ayoobi, N., Sharifrazi, D., Alizadehsani, R., Shoeibi, A., Gorriz, J. M., Moosaei, H., ... &Mosavi, A. Time series

forecasting of new cases and new deaths rate for COVID-19 using deep learning methods. Results in physics, **27**, 104495. (2021).

[7] Jaber, A. M., Ismail, M. T., &Altaher, A. M. Empirical mode decomposition combined with local linear quantile regression for automatic boundary correction. In Abstract and Applied Analysis **2014**, 1-8. (2014).

[8] Kaur, J., Parmar, K. S., & Singh, S. Autoregressive models in environmental forecasting time series: a theoretical and application review. Environmental Science and Pollution Research, **30**(8), 19617-19641. (2023).

[9] Abbasimehr, H., Paki, R., &Bahrini, A. A novel approach based on combining deep learning models with statistical methods for COVID-19 time series forecasting. Neural Computing and Applications, **34** 3135–3149. (2022).

[10] Yang, Q., Wang, J., Ma, H., & Wang, X. Research on COVID-19 based on ARIMA modelΔ —Taking Hubei, China as an example to see the epidemic in Italy. Journal of Infection and Public Health, **13**(10), 1415-1418. (2020).

[11] Alakkari, K. M., Mishra, P., Rawat, D., Abotaleb, M., & Ghazi, A. M. Using ETS State Space Model for forecasting on third wave on COVID19 in India. Mathematical Sciences Letters,v**(11)**2 59-64 (2022).

[12] ArunKumar, K. E., Kalaga, D. V., Kumar, C. M. S., Chilkoor, G., Kawaji, M., &Brenza, T. M. Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA). Applied soft computing, **103**, 107161. (2021).

[13] Chimmula, V. K. R., & Zhang, L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solitons & Fractals, **135**, 109864. (2020).

[14] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... & Liu, H. H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, **454**(1971), 903-995. (1998)

[15] Trybek, P., Sobotnicka, E., Wawrzkiewicz-Jałowiecka, A., Machura, Ł., Feige, D., Sobotnicki, A., & Richter-Laskowska, M. A New Method of Identifying Characteristic Points in the Impedance Cardiography Signal Based on Empirical Mode Decomposition. Sensors, **23**(2). 675. (2023).

[16] Awajan, A. M., Ismail, M. T., & Wadi, S. A. L. A review on empirical mode decomposition in forecasting time series. Italian Journal of Pure and Applied Mathematics, **43**, 301-323. (2019).

[17] Razif, N. R. A., &Shabri, A. Application of empirical mode decomposition in improving group method of data handling. In AIP Conference Proceedings**2500** (2023).

[18] Mohd Ismail and Ahmad M Awajan. “A new hybrid approach EMD-EXP for short-term forecasting of daily stock market time series data”. Electronic Journal of Applied Statistical Analysis **10**.2 307–327, (2017).

[19] Al-Jawarneh, A. S., Ismail, M. T., &Awajan, A. M. Elastic Net Regression and Empirical Mode Decomposition for Enhancing the Accuracy of the Model Selection. International Journal of Mathematical, Engineering and Management Sciences, **6**(2), 564. (2021).

[20] Al-Jawarneh, A. S., Ismail, M. T., Awajan, A. M., &Alsayed, A. R. Improving accuracy models using elastic net regression approach based on empirical mode decomposition. Communications in Statistics-Simulation and Computation, **51**(7), 4006-4025. (2022).

[21] Awajan, A. M., Ismail, M. T., & Wadi, S. A. Stock market forecasting using empirical mode decomposition with holt-winter. AIP Conference Proceedings **2184,** 050006. (2019).

[22] Box, G. E., & Jenkins, G. M. Time series analysis: Forecasting and control San Francisco. Calif: Holden-Day. (1976).

[23] Alabdulrazzaq, H., Alenezi, M. N., Rawajfih, Y., Alghannam, B. A., Al-Hassan, A. A., & Al-Anzi, F. S. On the accuracy of ARIMA based prediction of COVID-19 spread. Results in Physics, **27**, 104509. (2021).

[24] Chyon, F. A., Suman, M. N. H., Fahim, M. R. I., &Ahmmed, M. S. Time series analysis and predicting COVID-19 affected patients by ARIMA model using machine learning. Journal of Virological Methods, **301**, 114433. (2022).

[25] Hyndman, R. J., &Athanasopoulos, G. Forecasting: principles and practice. OTexts. (2018).

[26] Fawzy, H., Rady, E. H. A., & Fattah, A. M. A. A Comparative Simulation Study of ARIMA and Computational Intelligent Techniques for Forecasting Time Series Data. Journal of Statistics Applications & Probability Letters, **(11)**1. 1-7. (2022).

[27] Venables, W. N., & Ripley, B. D. Modern applied statistics with S-PLUS. Springer Science & Business Media. (2013).

[28] Hurvich, C. M., & Tsai, C. L. Bias of the corrected AIC criterion for underfitted regression and time series models. Biometrika,**78**(3), 499-509. (1991).

[29] Kufel, T. ARIMA-based forecasting of the dynamics of confirmed Covid-19 cases for selected European countries. Equilibrium. Quarterly Journal of Economics and Economic Policy, **15**(2), 181-204. (2020).

[30] Katoch, R., & Sidhu, A. An application of ARIMA model to forecast the dynamics of COVID-19 epidemic in India. Global Business Review, 0972150920988653. (2021).

[31] Tandon, H., Ranjan, P., Chakraborty, T., & Suhag, V. Coronavirus (COVID-19): ARIMA-based Time-series Analysis to Forecast near Future and the Effect of School Reopening in India. Journal of Health Management, **24**(3), 373-388. (2022).

[32] Donoghue, W. F. Distributions and Fourier transforms. Academic Press. (2014).

[33] Awajan, A. M., Ismail, M. T., & Al Wadi, S. Improving forecasting accuracy for stock market data using EMD-HW bagging. PloSone, **13** (7), e0199582. (2018).

[34] Van Loan, C. Computational frameworks for the fast Fourier transform. Society for Industrial and Applied Mathematics. (1992).

[35] Ramasamy, K., Jayakumar, S., &Chinnu, G. COVID-19 Situation at Chennai City–Forecasting for the Better Pandemic Management. International Journal of Current Research and Review, **12**(12), 37-47, (2020).

[36] Ramasamy K. and S Jayakumar. The Trend of COVID-19 at Bengaluru: Prediction to Continue the Better Epidemic Management. International Journal of Current Research and Review **12**.13 56-60,(2020).

[37] Pełka, P. Analysis and Forecasting of Monthly Electricity Demand Time Series Using Pattern-Based Statistical Methods. Energies,**16**(2), 827, (2023).

[38] Trull, O., Garc´ıa-D´ıaz, J. C., &Troncoso, A. (2020). Initialization methods for multiple seasonal Holt–Winters forecasting models. Mathematics, **8**(2), 268, (2020).

[39] Almazrouee, A. I., Almeshal, A. M., Almutairi, A. S., Alenezi, M. R., &Alhajeri, S. N. Long-term forecasting of electrical loads in kuwait using prophet and holt–winters models. Applied Sciences, **10**(16). 5627, (2020).

[40] Luo, J., Chen, W., Ray, J., & Li, J. Short-Term Polar Motion Forecast Based on the Holt-Winters Algorithm and Angular Momenta of Global Surficial Geophysical Fluids. Surveys in Geophysics, **43**(6) 1929–1945, (2022).

[41] Yelchuri, S., Rangaraj, A. G., Xie, Y., Habte, A., Joshi, M. C., Boopathi, K., ... &Balaraman, K. . A short-term solar forecasting platform using a physics-based smart persistence model and data imputation method (No. NREL/TP-5D00-81421). National Renewable Energy Lab.(NREL), Golden, CO (United States). (2021).

[42] Wang, J., Mei, H., Li, K., Zhang, X., & Chen, X. Collaborative Filtering Model of Graph Neural Network Based on Random Walk. Applied Sciences, **13**(3), 1786 (2023).

[43] Levantesi, S., Nigri, A., Piscopo, G., &Spelta, A. Multi-country clustering-based forecasting of healthy life expectancy. Quality & Quantity, 1-27. (2023).

[44] Freitas, W. B., & Junior, J. R. B. Random walk through a stock network and predictive analysis for portfolio optimization. Expert Systems with Applications, **215**. 119597 (2023)

[45] Banerjee, A., Dolado, J. J., Galbraith, J. W., & Hendry, D. Co-integration, error correction, and the econometric analysis of non-stationary data. Oxford university press. (1993).

[46] Ramsey, J. B. Tests for specification errors in classical linear least-squares regression analysis. Journal of the Royal Statistical Society: Series B (Methodological), **31**(2), 350-371 (1969).

[47] Breusch, T. S., & Pagan, A. R. A simple test for heteroscedasticity and random coefficient variation. Econometrica: Journal of the econometric society, **47**(5) 1287–1294. (1979)

[48] Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. Remote sensing of environment, **268**,112760. (2022).

[49] Mia, M. S., Rahman, M. S., & Das, S. Forecasting the BTD/USD exchange rate: An accuracy comparison of artificial neural network models and different time series models. Journal of Statistics Applications & Probability Letters: An International Journal, **4**(3), 131-138. (2017).

[50] Pandhiani, S. M., &Shabri, A. B. A hybrid model for monthly time series forecasting. Applied Mathematics & Information Sciences, **9**(6), 2943. (2015).

[51] Alrweili, H., &Fawzy, H. Forecasting crude oil prices using an ARIMA-ANN hybrid model. Journal of Statistics Applications &Probability, **11**(3), 845-855. (2022).

[52] Awajan, A. M., Ismail, M. T., & Al Wadi, S. A hybrid EMD-MA for forecasting stock market index. Italian Journal of Pure and Applied Mathematics, **38**(1), 1-20. (2017).

[53] Akintunde, M. O., Amusan, A. S., & Olawale, A. O. On Empirical Comparison of Forecast Performance of Autoregressive Moving Average Model and Generalized Autoregressive Conditional Heteroscedasticity Model. Journal of Statistics Applications &Probability Letters, **(8)**1 41-49 (2021).

[54] Awajan, A. M., & Ismail, M. T. A hybrid approach emd-hw for short-term forecasting of daily stock market time series data. AIP Conference Proceedings **1870**(1), 060006, (2017)