Development of novel hybrid models for the prediction of COVID-19 in Kuwait

Ahmad Al-Dousari[1, *], Maria Qurban[2], Ijaz Hussain[2,] Maha Al-Hajeri[3]

[1]Dept. of Geography, Kuwait University, Kuwait City, Kuwait.
[2]Dept. of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
[3]Dept. of Health Information and information management, College of Ailed Health,
Kuwait University, Kuwait City, Kuwait.
*Corresponding author: dr_ahmad@hotmail.com

Abstract

The first case of COVID-19 in Kuwait was reported on February 24, 2020. There is a need to develop a prediction model for estimating this epidemic size. In this study, we aimed to develop and compare several prediction models using real-time data of COVID-19 from February 24 to June 12, 2020. We modeled the uncertainty and non-stationary real-time data of COVID-19 cases using a multilayer model with different decomposition techniques. We applied our proposed hybrid methodology to predict COVID-19 cases in Kuwait. We further evaluated the performance of the novel hybrid model with others using mean relative error, mean absolute error, and mean square error. We found that our proposed hybrid approach performed better than others for predicting COVID-19 cases.

Keywords: Complete Ensemble Empirical Mode Decomposition with Adaptive Noise; COVID-19; Empirical Mode Decomposition; Ensemble Empirical Mode Decomposition; hybrid model.

1. Introduction

Different studies on modelling and forecasting of COVID-19 are carried out by different researchers who provided predictions of this epidemic. Boccaletti *et al.* (2020) used the Susceptible–Exposed–Infected–Recovered (SEIR) model to predict the peak of this epidemic. Koczkodaj *et al.* (2020) forecasted the increase in the number of cases of COVID-19, reaching one million outside of China by specifying a date according to WHO reports. Almeshal *et al.* (2020) used deterministic and stochastic models for the prediction of the epidemic peak in Kuwait by using confirmed cases. Chimmula & Zhang (2020) proposed a prediction model using genetic programming (GP) for the established and death cases in India.

Chimmula & Zhang (2020) proposed long short-term memory (LSTM) networks to forecast the COVID-19 cases in Canada. However, many researchers ignored the time-varying characteristics of COVID-19 data. Extensive measures are needed to control the spread and transformation of COVID-19 in this pandemic outbreak. Prediction of COVID-19 cases can help to control the current outbreak. So that special efforts and attention may be applied to reduce the person-to-person transformation of COVID-19.

Kuwait faced this pandemic of COVID-19 on February 24, 2020, with five cases. These cases were brought to Kuwait by travelers. These numbers keep on increasing over time. To control or slow down the spread of this pandemic many efforts and steps are taken, i.e., quarantine efforts and

rigorous containments by the government of Kuwait. It is identified that most of the cases are imported by citizens of Kuwait who came back to Kuwait, and some cases are locally transmitted.

 The main objectives of our current study are

1. To develop a novel strategy to estimate and predict the spread of COVID-19 in Kuwait for the estimation of epidemic size by considering the time-varying and nonlinearity of the Corona cases, which is mostly ignored in previous studies.
2. To provide an efficient way to forecast the spread of COVID-19 to policymakers for planning the intervention measures.

2. Proposed method

The hybrid models are used for the prediction and forecasting of the spread of different pandemics and diseases such as the SARS pandemic and dengue outbreak (Yoneyama *et al.*, 2010; Husin *et al.*, 2012). The hybrid models have been used with machine learning and with Autoregressive Integrated Moving Average (ARIMA) models. Two models are commonly used for prediction: a process-based model and a data-driven model. The process-based model needs a large amount of calibration and validation data set (Chimmula & Zhang, 2020). The data-driven model considered the physical mechanism and scientific knowledge of the COVID-19 stochastic process (Islam, 2011). The lack of scientific knowledge, unavailability of such data, and overfitting make it challenging to understand a model. Hybrid methodologies with different decomposition techniques are used for modelling and accurate prediction of the current pandemic. The multi-models are combined with other data processing techniques to increase the prediction accuracy. The primary strategy of the hybrid technique is based on decomposition and denoising, prediction, and ensemble stage (Peng *et al.*, 2017; Di *et al.*, 2014; Chen *et al.*, 2017). The main objective of decomposition and ensemble is the simplification of the forecasting process and to evaluate its performance, respectively. In many time-series data, noises are specific components that sometimes reduce the prediction accuracy. In the present study, Empirical Model Decomposition (EMD) and Wavelet Analysis (WA) (Dybała & Zimroz, 2014; Liu *et al.*, 2010; Science *et al.*, 2020), based thresholds are used for the reduction of noises from the COVID-19 data. The EMD is combined with ANN in many past studies, especially in hydrology (Liu *et al.*, 2019). To overcome the mode mixing problem, Wu *et al.* (2009) proposed a new technique of decomposition called Ensemble Empirical Mode Decomposition (EEMD) in which they used white Gauss noise. Many hybrid methods based on EEMD are used for prediction in past studies (Santhosh *et al.*, 2018)**.** Torres *et al.* (2011) proposed the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) technique. In their proposed model, at each decomposition stage, a specific noise is added, and a residue is calculated to get each mode, which is an improved form of EEMD (Wei *et al.*, 2016).

   In the current study, we proposed two novel methods for the enhancement of the prediction accuracy of the time-series data. Both methods have the same map and formation except in the denoising stage, where the approaches for the elimination of the noises from the data are different. At the decomposition stage of both methods, EEMD and CEEMDAN are used to identify the oscillation. At the prediction stage, the multi models are used after considering the nature of IMFs for the accurate prediction of the obtained IMFs. There are two main objectives of using the multi models; the first objective is to predict the IMF by looking up their nature accurately. The second objective is to determine the attainments of the simple and complex models after the reduction of

the complexity of time series data by using decomposition methodology. All the IMFs obtained through prediction are combined for the final prediction of the time series data. The proposed hybrid models consist of denoising, which is denoted as D-step, decomposition, and prediction, which are denoted as Decompose-step and P- step, respectively (Nazir *et al.*, 2019). A short description of these steps is given below:

1. To get rid of the noise from the time series data of COVID-19, WA and EMD methods are used.
2. In this step, different methodologies, i.e., EEMD, and CEEMDAN, are used for the decomposition of the denoised series into the n IMF components and one residual term.
3. The series obtained after denoising is decomposed into IMFs components and one residual term, which is predicted by using linear stochastic and non-linear machine learning methods. The model, which has a lower rate of error for prediction than other models, is utilized for further consideration based on three performance evaluation criteria. In the last step, the predicted results are combined to get the eventual prediction.

For better understanding and easiness of the reader, by combining all the steps mentioned above, the proposed strategy can be named EMD-CEEMDAN-MM, and WA-CEEMDAN-MM is exhibited in Figure 1.
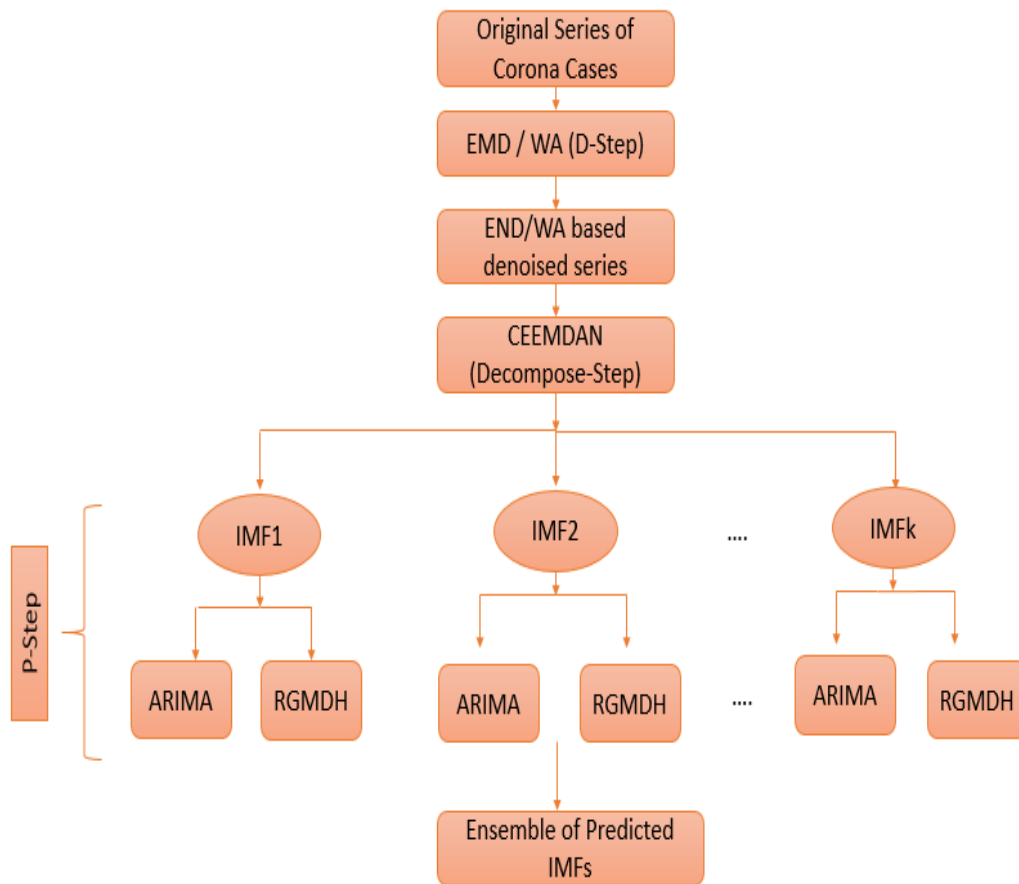

Fig. 1. The suggested structure of EMD/WA-CEEMDAN-MM for the prediction of COVID-19 cases.

2.1 Denoising Step (D-Step)

In this study, EMD and WA-based thresholds are used for the reduction of noises from COVID-19 cases of Kuwait.

Wavelet analysis (WA): Wavelet analysis is a useful and powerful tool for converting a signal into a stationary signal with specific effectiveness. Discrete Wavelet Transform (DWT) is used to eliminate the noises from the signal, and Symlet 8 mother function is used to decompose the signal into approximation and detailed coefficients by following equations, respectively.

$$a_{i,l} = \sum_{i=0}^{2^{N-i}-1} 2^{-i/2} \phi(2^{-i}p - l) \tag{1}$$

$$\text{and} \quad d_{i,l} = \sum_{i=1}^{I} \sum_{l=0}^{2^{N-i}-1} 2^{-i/2} \theta(2^{-i}p - l) \tag{2}$$

Hard and Soft Thresholds

In the current study, to apply the wavelet denoising techniques, soft and hard thresholds are applied to the data, which are listed below (Nazir *et al.*, 2019). The soft threshold is:

$$d'_{i,l} = \begin{cases} d_{i,l} & |c_{i,l}| \geq Th_i \\ 0 & |c_{i,l}| < Th_i \end{cases} \tag{3}$$

and a hard threshold is

$$d'_{i,l} = \begin{cases} sgn(c_{i,l})(|c_{i,l}| - Th_i) & |c_{i,l}| \geq Th_i \\ 0 & |c_{i,l}| < Th_i \end{cases} \tag{4}$$

where Th_i is the threshold which is calculated as,

$$Th_i = a\sqrt{2\widehat{Md}_i ln(N)}, i = 1, 2, 3, ...I,$$

where, a is constant, between 0.4 and 1.4 with the jump of 0.1 and $\widehat{Md}_i$ is the median deviation, i.e.,

$$\widehat{Md}_i = \frac{median(\{d_{i-1,l}|l=1,2,3,...,2^{L-1}-1\})}{0.6745} \tag{5}$$

By utilizing the following Equation, the decomposed signal is reconstructed by using the approximations and noise-free details:

$$\widetilde{x}(p) = \sum_{l=0}^{2^{N-i}-1} a'_{i,l} 2^{-i/2} \phi(2^{-i}p-l) + \sum_{i=1}^{I} \sum_{l=0}^{2^{N-i}-1} d'_{i,l} 2^{-i/2} \theta(2^{-i}p-l)$$

(6)

where $a'_{i,l}$ and $d'_{i,l}$ are the threshold approximation coefficient and detailed threshold coefficient, respectively.

Empirical Mode Decomposition (EMD): EMD is a decomposition technique of flexible kind used to decompose the sophisticated kind of signals into different parts. The primary purpose of EMD is to extract the IMFs from the signal by satisfying two conditions.

(a) In the real data, the number of zero crossings and the number of extrema should either be equal or differ by at most one.

(b) At any point, the value of the envelope defined by local maxima and minima should be zero.

The procedure of the EMD for an original time series $x(p),(p=1,2,3,...,N)$ is described as below:

1. Identify all the local extrema from the original signal.
2. Estimate the lower $G(p)$ and upper $U(p)$ envelopes.
3. Find the average of the upper envelope and lower envelope

    $m(p) = (U(p) + G(p))/2$.

4. Compute the difference $d(p) = x(p) - m(p)$.

5. Identify the properties of mean $m(p)$ and difference $d(p)$ obtained in the last two steps.
6. Repeat steps 1-5 until the number of extrema is less than or equal to one so that no more IMF can be extracted.

Finally, the signal can be shown

$$x(p) = \sum_{j}^{n} d_j + e(p)$$,

(7)

Except for the last two IMF (which are used without denoising because of the low frequencies), the rest of the procedure of denoising process is the same as in wavelet-based denoising by using equations (3), (4), and (7). Before the reconstruction of the signal, a smooth input signal can be obtained by thresholding the IMFs. The reconstruction of the denoised signal is generalized as:

$$\widetilde{x}(p) = \sum_{j=1}^{n_1-2} h_j(p) + \sum_{j=n_1-2}^{n} h_j(p) + e(p)$$

(8)

Where the parameter n_1 is the number of IMFs, which provide us with the easiness of the elimination of the low order IMF which are noisy, and also of higher-order IMFs which are a little

bit noisy in Gaussian noise conditions as $(j = 1, 2, 3, ..., n)$, $h_j(p)$ and $e(p)$ is the jth IMF and trend of the signal respectively.

2.2 Decomposition Step (Decompose Step):

We used the EEMD technique in the decomposition step to mitigate the mode mixing problem.

Ensemble Empirical Mode Decomposition (EEMD): To bring some improvements in the EMD and mitigate the mode mixing, EEMD is developed. In this technique, the white noise added by EEMD is distributed equally among all the time-frequency space, which helps in the separation of the frequency scales and decreases the occurrence of mode mixing. The procedure is well explained in (Wei *et al*., 2016; Zhang *et al*., 2019):

Complete Ensemble Empirical Mode Decomposition with added Noise (CEEMDAN): Although EEMD can bring down the problem of mode mixing to a certain degree with included white noise sequence, the error cannot be eliminated after the computation of the averaging to a finite number. It affects the sequence of reconstruction. For the elimination of the mode mixing, CEEMDAN adds the adaptive white noise smoothing pulse interference in decomposition, and for making the decomposition of the data more complete, it uses the properties of the mean Gaussian white noise whose mean is zero. The detailed procedure of the CEEMDAN is given in (Nazir *et al*., 2019).

2.3 Prediction Step (P-Step):

In the prediction stage, the denoised IMFs are used as input for the prediction of the COVID-19 by using machine learning time series and stochastic methods. The reason for the usage of two different types of models in prediction is that the IMFs with high frequencies are predicted better through ML methods. Stochastic models offer better outcomes for the prediction of the IMFs with low frequencies. The models used for the prediction purpose are described as follows:

Autoregressive Integrated Moving Average (ARIMA) Model: For the prediction of the IMF, the Autoregressive moving average model is used as follow:

$$IMF_k^j = \psi_1 IMF_{k-1}^j + ... + \psi_p IMF_{k-p}^j + \varepsilon_k^j + \varphi_1 \varepsilon_{k-1}^j + ... + \varphi_q \varepsilon_{k-q}^j \tag{9}$$

Here IMF_k^j is the jth IMF and ε_p^j it's j^{th} residual obtained through CEEMDAN, p and q are the lag values of the autoregressive and moving average terms. If the time series data is non-stationary an appropriate degree of differencing (d) is used to make the series stationary. In this case, the model is denoted by $ARIMA(p, d, q)$.

Group Method of Data Handling (GMDH) type NN Model: GMDH is a type of new neural network. The GMDH-NN models are established by considering the evolutionary method of modelling, which is a program that generates a polynomial type of neural network for modelling the data. The input variables, hidden layers containing neurons, the most appropriate model structure, and the number of layers are determined automatically in these types of networks. By

considering the evaluation criteria, some of the neurons are chosen then the output of these selected neurons turns into the input of the next layer.

The procedure is repeated until the final layer. In the last layer, only one predicted neuron is considered. But GMDH-NN selects the relation of only two variables and ignores the effect of an individual variable.

A refined form of GMDN-NN is the Architecture Group Method of Data Handling (RGMDH-NN), which considers not only two variables but also considers them individually; the remaining procedure of RGMDH is the same as GMDH.

3. Study area and experimental design

Selection of area for study: The confirmed COVID-19 cases from February 24, 2020, when 5 cases of COVID-19 were reported in Kuwait for the first time to June 12, 2020, is considered here for the implementation of our model and prediction purposes. The time series plot of COVID-19 cases in Kuwait is presented in Figure. 2, whose average is not zero and variance is not constant over time.
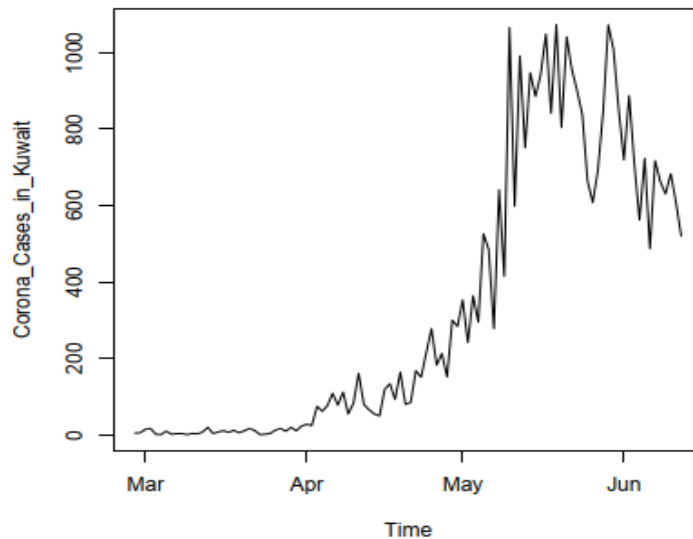

Fig. 2. Time series plot of Covid-19 cases in Kuwait.

Comparison of the proposed hybrid model with other models: Both suggested models are compared with other existing 1-Stage, 2-Stage, and 3-stage models.

***Accuracy Measure Techniques*:** The performance of the models is compared by using the four evaluation measures i.e., Mean Relative Error (MRE), Mean Absolute Error (MAE), Mean Square Error (MSE), and Mean Absolute Percentage Error (MAPE) (Kim & Kim, 2016).

4. Results and discussions:

Some of the previous studies (Knight *et al.*, 2016) used linear methods in their studies, which neglects the temporal components of the data. The data of infectious diseases is usually non-linear. Many statistical models are proposed to model the transmission dynamic of COVID-19 (Benvenuto *et al.*, 2020; *Forecasting of COVID-19 Confirmed Cases in Different Countries with ARIMA Models*, 2020). In some cases, these statistical models are not able to capture the nonlinearity and non-stationarity of the COVID-19 data.

To check the linearity and non-stationarity of the COVID-19 cases of Kuwait, we used the Augmented Dickey-Fuller (ADF) (Taylor & Test, 2012) test on our data. The ADF test was not significant, and we concluded the series is said to be non-stationary.

D-stage results: by using two noise removal filters, the results of denoising are described as follows:

Wavelet-based denoising: By using equations (1) and (2), the approximations are calculated on the COVID-19 cases in Kuwait, to remove the noises from the coefficients of data, soft and hard thresholding are used. Rules of soft and hard thresholding are calculated by using equations (3) and (4), respectively. Based on the lower value of the MSE denoised series, hard thresholds are reconstructed for wavelet analysis.

EMD based denoising: For removing the noises from the series of COVID-19 cases using EMD, IMFs are calculated. For denoising, these calculated IMFs except the last two IMFs, soft and hard thresholding rules, are used. The denoised IMFs based on hard thresholding is used for the reconstruction of noise-free COVID-19 time-series data. The WA and EMD based denoised are combined in Figure. 3.
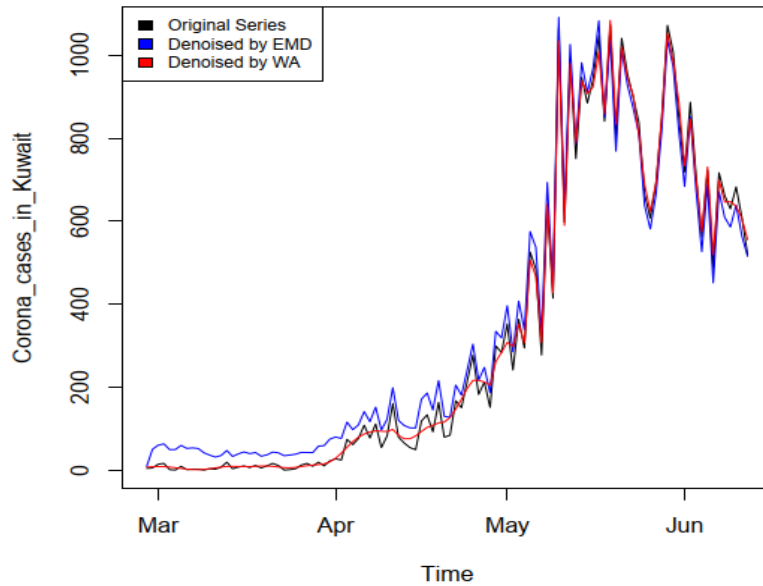
Fig. 3. The denoised series of COVID-19 cases in Kuwait using EMD (in blue color) and WA (in red color) techniques.

The statistical measures of original series and denoised series for COVID-19 cases, i.e., Mean, Standard deviation, MRE, MAE, and MSE for EMD and Wavelet, noise removal techniques are shown in Table 1. Table 1 and Figure. 1 suggest that WA behaves better than EMD. The results show that WA and EMD behave differently for COVID-19 cases. There is no significant difference in the mean and standard deviation of the original series, denoised series by EMD, and WA. However, WA performed better than EMD by attaining the minimum value of MRE, MAE, and MSE. In general, WA performed well to denoise the COVID-19 cases. In the decomposition stage, EMD and WA both denoising techniques are used separately as input to get those characteristics that change in terms of varying frequencies, i.e., high frequencies and low frequencies.

Table 1. Statistical measures of WA and EMD based denoised Series of COVID-19 Cases in Kuwait.

Statistical measures of WA and EMD based denoised Series of COVID-19 Cases in Kuwait.						
COVID-19 Cases in Kuwait	**Mode**	**Mean**	**Sigma**	**MRE**	**MAE**	**MSE**
Confirmed Cases	Original series	329.7358	360.3322			
	EMD	351.1726	338.9776	37.11843	3.249895	1480.283
	WA	329.7597	357.9186	16.30251	0.369013	497.3974

9

Decompose-stage results: To get the local changing features for the time from the denoised series of COVID-19 cases by WA/EMD are decomposed further into five IMF components and one residual term to get the local changing features for the time from the denoised series of COVID-19 cases. The EMD and WA with CEEMDAN decomposition methods are used for the extraction of IMFs from the COVID-19 patients in Kuwait. The decomposition results of EMD- CEEMDAN and WA-CEEMDAN techniques of COVID-19 cases in Kuwait are presented in Figure. 4. The drawn-out IMFs represent the characteristics of COVID-19 cases where the starting IMFs show a higher frequency. In contrast, the last two IMFs represent lower frequency and residual shown as a trend.
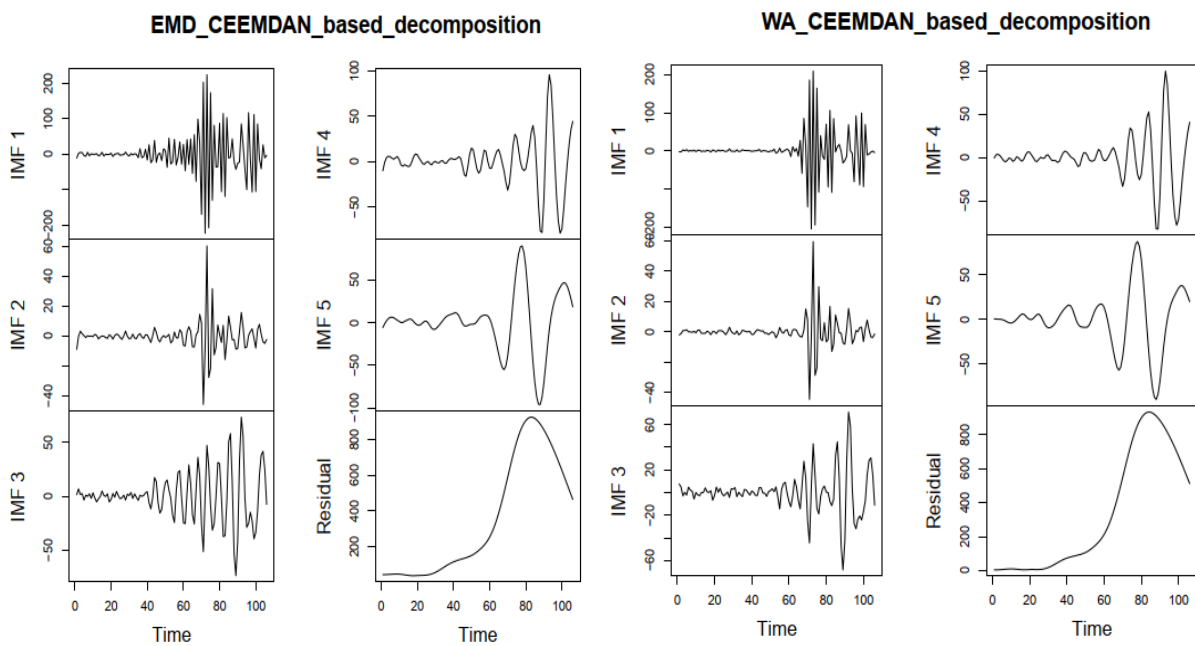

Fig. 4. The decomposed series by EMD-CEEMDAN and WA-CEEMDAN of COVID-19 cases in Kuwait.

The amplitude's value of the white noise is chosen as 0.2, and a maximum number of the ensemble members (1000) are selected. Almost all IMFs and residuals for a series of COVID-19 cases showed identical characteristics for EMD-CEEMDAN and WA-CEEMDAN decomposition methods.

P-step results: Two methods, i.e., ARIMA (p, d, q) and GMDH-NN models, are used to predict all extracted IMFs and residuals to get precise results.

The series of COVID-19 cases is divided into training data set and testing data sets using 70% and 30% ratios, respectively. The parameters of the model and its structure are estimated using 77 observations. The suggested model and other models used for comparison purposes are tested in

terms of their validity using 33 observations. After assessing multi-models for every IMF component and residual, the model with the least value of MRE, MAE, and MSE is considered as the most appropriate and selected for the prediction of each IMF. The findings for the training and testing data set of the suggested model and all other models in comparison for COVID-19 cases are given in Table 2 and Table 3, respectively. The prediction results of proposed models EMD-CEEMDAN-MM and WA-CEEMDAN-MM illustrated the effectiveness with a minimum value of MRE, MAE, and MSE in comparison with 1-stage, 2-stage, and 3-stage evaluation models for both testing and training data sets. However, the suggested model WA-CEEMDAN-MM acquired the lowest value of MSE than the other proposed EMD-CEEMDAN-MM model.

Table 2: The evaluation of the prediction error of suggested (EMD-CEEMDAN-MM and WA-CEEMDAN-MM) in comparison with other models for COVID-19 cases in Kuwait having a training data set.

Here 1-stage ARIMA model, as compared to 2-stage ARIMA, performed worse as the 1–stage ARIMA model attains maximum value of MSE without applying the techniques of denoising and decomposition on the series of COVID-19 cases. The predicted graph of the suggested model WA and EMD-CEEMDAN-MM in comparison with all 2-stage models are presented in Figure 5.
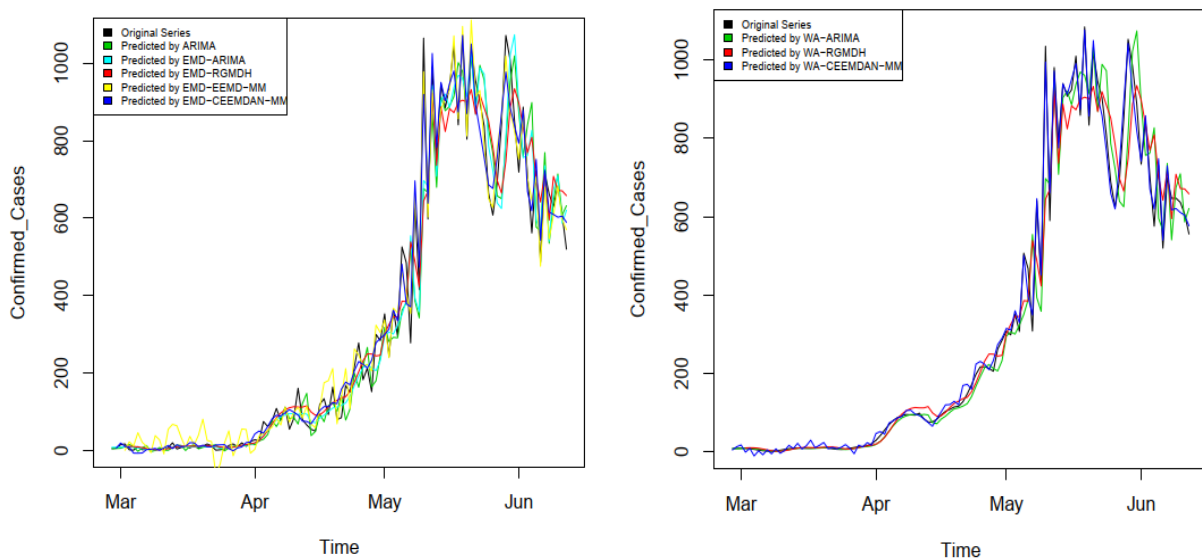

Fig. 5. Prediction results of COVID-19 cases in Kuwait using suggested EMD-CEEMDAN-MM and WA-CEEMDAN-MM models in comparison with 1-stage (ARIMA), 2-stage (EMD/WA-ARIMA, EMD/WA-RGMDH), and 3-stage (EMD-EEMD-MM) predicted models.

However, it is concluded from Table 2 and Figure 5 that the suggested models, i.e., EMD-CEEMDAN-MM and WA-CEEMDAN-MM perform well to predict the COVID-19 cases, by decreasing its intricacy and increasing the performance of prediction over 1-stage, 2-stage, and 3-

stage models. Also, from both suggested models, WA-CEEMDAN-MM performed better than the EMD-CEEMDAN-MM model by attaining a minimum value of MRE, MAPE, and MSE.

Table 2. The evaluation of the prediction error of EMD-CEEMDAN-MM and WA-CEEMDAN-MM in comparison with other models for COVID-19 cases in Kuwait having a training data set.

COVID-19 Cases	Model Name	Models	MRE	MAPE	MSE
Confirmed Cases	1-S	ARIMA	58.31	0.67	8935.98
	2-S	WA-ARIMA	44.62	0.16	6567.60
		WA-RGMDH	44.54	0.21	6612.13
		EMD-ARIMA	55.86	0.37	8293.34
		EMD-RGMDH	67.70	1.02	9795.46
	3-S	EMD-EEMD-MM	103.47	2.48	16228.10
		EMD-CEEMDAN-MM	96.84	0.43	20571.46
		WA-CEEMDAN-MM	**23.03**	**0.15**	**1531.45**

Table 3. The evaluation of the prediction error of EMD-CEEMDAN-MM and WA-CEEMDAN-MM in comparison with other models for COVID-19 cases in Kuwait having a testing data set.

COVID-19 Cases	Model Name	Models	MRE	MAE	MSE
Confirmed Cases	1-S	ARIMA	115.93	0.15	18307.11
	2-S	WA-ARIMA	9.77	0.01	194.91
		WA-RGMDH	10.15	0.01	196.78
		EMD-ARIMA	115.75	0.15	18149.25
		EMD-RGMDH	109.86	0.15	16563.53
	3-S	EMD-EEMD-MM	43.16	0.06	3121.96
		EMD-CEEMDAN-MM	48.89	0.06	3648.31
		WA-CEEMDAN-MM	**5.84**	**0.01**	**51.29**

5.   Conclusions

Two models were carried out for the forecast of COVID-19 spread in the whole world, including Kuwait, using the CEEMDAN decomposition technique. The forecast was based on the data of COVID-19 cases in Kuwait from February 24, 2020, to June 12, 2020. The main objective of the study was to establish an efficient model for the forecast of COVID-19 cases in Kuwait.

References

Almeshal, A. M., Almazrouee, A. IAlenizi, M. R., & Alhajeri, S. N. (2020). Forecasting the spread of COVID-19 in Kuwait using compartmental and logistic regression models. *Applied Sciences*, **10**(10). https://doi.org/10.3390/APP10103402

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Data in brief Application of the ARIMA model on the COVID- 2019 epidemic dataset. *Data in Brief*, *29*, 105340. https://doi.org/10.1016/j.dib.2020.105340

Boccaletti, S., Ditto, W., Mindlin, G., & Atangana, A. (2020). Modeling and forecasting of epidemic spreading: The case of COVID-19 and beyond. *Chaos, Solitons and Fractals*, **135**. https://doi.org/10.1016/j.chaos.2020.109794

Chen, W., Zhang, D., & Chen, Y. (2017). Random noise reduction using a hybrid method based on ensemble empirical mode decomposition. *J. Seismic Explor*e, **26**(3): 227-249.

Dybała, J., & Zimroz, R. (2014). Rolling bearing diagnosing method based on Empirical Mode Decomposition of the machine vibration signal. *Applied Acoustics*, **77**: 195–203. https://doi.org/10.1016/j.apacoust.2013.09.0 01

Dehesh, T., Mardani-Fard, H. A., & Dehesh, P. (2020). *Forecasting of COVID-19 Confirmed Cases in Different Countries with ARIMA Models*. (2020). 1–12. https://doi.org/10.1101/2020.03.13.20035345doi: medRxiv

Husin, N. A., Mustapha, N., Sulaiman, N., & Yaakob, R. (2012). *A Hybrid Model using Genetic Algorithm and Neural Network for Predicting Dengue Outbreak*. (September), 2–4.

Islam, Z. (2011). Literature Review on Physically Based Hydrologic Modeling. *Alberta Canada: Department of Civil and Environmental Engineering, University of Alberta*.

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, **32**(3): 669–679. https://doi.org/10.1016/j.ijforecast.2015.12.003

Knight, G. M., Dharan, N. J., Fox, G. J., Stennis, N., Zwerling, A., Khurana, R., & Dowdy, D. W. (2016). International Journal of Infectious Diseases Bridging the gap between evidence and policy for infectious diseases: How models can aid public health decision-making. *International Journal of Infectious Diseases*, **42**: 17–23. https://doi.org/10.1016/j.ijid.2015.10.024

Koczkodaj, W. W., Mansournia, M. A., Pedrycz, W., Wolny-Dominiak, A., Zabrodskii, P. F., Strzałka, D., … Mazurek, J. (2020). 1,000,000 cases of COVID-19 outside of China: The date predicted by a simple heuristic. *Global Epidemiology*, *2*, 100023. https://doi.org/10.1016/j.gloepi.2020.100023

Liu, H., Tian, H., Chen, C., & Li, Y. (2010). A hybrid statistical method to predict wind speed and wind power. *Renewable Energy*, **35**(8): 1857–1861. https://doi.org/10.1016/j.renene.2009.12.011

Nazir, H. M., Hussain, I., Faisal, M., Shoukry, A. M., Gani, S., & Ahmad, I. (2019). *Development of Multidecomposition Hybrid Model for Hydrological Time Series Analysis*. Volume 2019, Article ID 2782715, https://doi.org/10.1155/2019/2782715

Peng, T., Zhou, J., Zhang, C., & Fu, W. (2017). Streamflow forecasting using empirical wavelet transform and artificial neural networks. *Water*, **9**(6): 406

Santhosh, M., Venkaiah, C., & Kumar, D. M. V. (2018). Ensemble empirical mode decomposition based adaptive wavelet neural network method for wind speed prediction. *Energy Conversion and Management*, **168**(May): 482–493. https://doi.org/10.1016/j.enconman.2018.04.099

Science, N., Phenomena, C., Kumar, V., Chimmula, R., & Zhang, L. (2020). *Chaos, Solitons, and Fractals Time series forecasting of COVID-19 transmission in Canada using LSTM networks R. 135*. https://doi.org/10.1016/j.chaos.2020.109864

Taylor, P., & Test, D. F. (2012). *Bag Order and Ceitica s f the Augmented Dic*. (December 2014): 37–41. https://doi.org/10.1080/07350015.1995.10524601

Torres, M. E., Colominas, M. A., Schlotthauer, G., & Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. *ICASSP, IEEE International Conference on Acoustics, Speech, and Signal Processing – Proceedings*.

Wei, X., Lin, R., Liu, S., & Zhang, C. (2016). *Improved EEMD Denoising Method Based on Singular Value Decomposition for the Chaotic Signal. 2016.*

Wu, Z., Huang, N. E., & Chen, X. (2009). *The multi-dimensional ensemble empirical mode decomposition method*. **1**(3): 339–372.

Xu, S., & Niu, R. (2018). Computers and Geosciences Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. *Computers and Geosciences*, **111**(March 2017): 87–96. https://doi.org/10.1016/j.cageo.2017.10.013

Yoneyama, T., Das, S., & Krishnamoorthy, M. (n.d.). *A Hybrid Model for Disease Spread and an Application to the SARS Pandemic*. 1–24.

Zhang, J., Jiang, R., Li, B., & Xu, N. (2019). Computers and Geosciences An automatic recognition method of microseismic signals based on EEMD-SVD and ELM. *Computers and Geosciences*, **133**(September): 104318. https://doi.org/10.1016/j.cageo.2019.104318 https://doi.org/10.1109/ICASSP.2011.5947265