Int. J. Nonlinear Anal. Appl. 13 (2022) No. 1, 627-641
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.5553

A hybrid ARFIMA-fuzzy time series (FTS) model for forecasting daily cases of Covid-19 in Iraq

Saif Adnan Salman[a,*], Emad Hazim Aboudi[a]

[a]*Department of Statistics, College of Management and Economics, Baghdad University, Baghdad, Iraq*

(Communicated by Madjid Eshaghi Gordji)

Abstract

Most time series are characterized in practice that they consist of two components, linear and non-linear, and when making predictions, the single models are not sufficient to model these series. Recently several linear, non-linear and hybrid models have been proposed for prediction, In this research, a new hybrid model was proposed based on the combination of the linear model Auto-Regressive Fractionally Integrated Moving Average (ARFIMA) with the non-linear model fuzzy time series model (FTS). The proposed hybrid model analyzes the linear component of the specified time series using the ARFIMA model, calculates the estimated values, and then calculates the residuals for this model by subtracting the estimated values from the original time series. The nonlinear component is analyzed using the (FTS) model for the computed residuals, which inherently contain the nonlinear patterns of the time series. The final values for the prediction by applying the hybrid model (ARFIMA-FTS) are obtained by combining the predictions of the (ARFIMA) model of the original series with the predictions of the model (FTS) for the residual series. The new hybrid model was used to predict those infected with Covid-19 virus in Iraq for the period from 24/2/2020 to 11/8/2021. The proposed model was more efficient in the prediction process than the single (ARFIMA) model using a number of comparison criteria, including (RMSE), (MAPE) and (MAE). The final results showed that the proposed model has the ability to predict time series that contain linear and nonlinear components.

Keywords: ARFIMA, Fuzzy Time Series, Long Memory Time Series, Hybrid Model.

*Corresponding author

Email addresses: saif.salman0901@coadec.uobaghdad.edu.iq (Saif Adnan Salman), emadhazim@coadec.uobaghdad.edu.iq (Emad Hazim Aboudi)

Received: August 2021 *Accepted:* September 2021

1. Introduction

The process of forecasting time series is one of the most important methods of statistical analysis that plays a major role in making decisions in uncertainty. The time series is defined as a set of measurements and historical observations that are linked to each other for the phenomenon and for certain periods of time and are usually equal in length [6]. Over the past decades, researchers have been interested in building a models of time series for their importance and ability to explain phenomena and data in many fields including (economic, social, medical...etc) and even at the individual level. One of the most important modern problems facing the world, which has occupied the attention of researchers, is the spread of the (covid-19) virus, as the world is facing great pressures due to the pandemic, and the increasing number of daily infections with the virus puts human life at great risk. It became necessary to predict the number of future cases to prepare a plan control this setting, recently many methods and models have been presented to predict the numbers of people infected with this virus, including time series prediction methods. The most important of these models are the long memory time series and the FTS model, and the FTS model has a flexible mechanism that sometimes outperforms the basic models of time series because it does not require in the prediction process the basic assumptions for building the model. This model is characterized by high-accuracy predictions for its use of many artificial intelligence algorithms. In this research, a hybrid model was proposed between the (ARFIMA) model and the (FTS) model, and this research aims to build a hybrid model to predict the numbers of people infected with the (covid-19) virus in Iraq.

2. Long Memory Process

The time series that is represented by a long-term process, the continuity in its observations may make the behavior of the autocorrelation coefficients of the series observations not decreasing exponentially towards zero when the displacements increase [4], as is the case in the short-term time series , but the autocorrelation function will be In the form of hyperbola and decreasing very slowly and at a polynomial rate, and this process is referred to as the Long Memory Time Series. [10] is considered the first to notice the phenomenon of long memory time series in the field of irrigation, and then subsequent studies included the phenomena Economic, as it turned out to have a long memory. [13] confirmed the idea of [10] by constructing the Brown fractal movement. So, it can be said that a process has a long memory in the time field if:

$$\lim_{n\to\infty} \sum_{i=-n}^{n} |\rho_i| = \infty \tag{2.1}$$

Since the absolute values of the autocorrelation are not aggregated.

2.1. Testing of Long Memory

The Hurst exponent (H), which is produced by the rescaled range (R/S) analysis [10], is used to determine if a time series contains long memory. The Hurst exponent measures long term non periodic dependency and shows the average duration the reliance may persist for a given time series. The R/S analysis begins by calculating the R range for a given n

$$R(n) = \max_{1\le j\le n} \sum_{j=1}^{n} \left(X_j - \overline{X}\right) - \min_{1\le j\le n} \sum_{j=1}^{n} \left(X_j - \overline{X}\right) \tag{2.2}$$

where, R(n) is the range of accumulated deviation of X(t) over the period of n and $\overline{X}$ is the overall mean of the time-series. Let S(n) be the standard deviation of X(t) over the period of n.

$$\begin{aligned} Q(n) &= R(n)/S(n) \\ Q(n) &= \alpha n^{H}, \qquad \text{where } \alpha \text{ is a constant, } n \ \rightarrow \infty \end{aligned} \tag{2.3}$$

Taking logarithm on both sides of the Hurst exponent equation (2.3), we have:

$$\log Q(n) = \alpha + H \ \log(n), \qquad as\ n \ \rightarrow \infty.$$

Then

$$H \approx \frac{\log Q(n)}{\log(n)} \tag{2.4}$$

The parameter (H) relates the average values of sub-samples of equal length to the series with the number of observations for each sub-sample of equal length, and the value of the parameter (H) is always greater than zero. In general, the relationship between the Hearst exponent and the long memory parameter (d) is represented by the relationship [9, 12].

$$H = 1 - d$$

3. ARFIMA Model

The idea of time series is simply crystallized in estimating a mathematical model that enables it to simulate reality through the historical gradation of that phenomenon, as the model estimates the special parameters of the prediction process, obtaining estimated values for the time series, and using this model to obtain future values for the phenomenon studied. ARFIMA [8] is one of the modern methods that contribute to analysis because it is characterized by a high ability in modeling time series with long memory and present in the majority of time series, and it is an extension of ARIMA models for [3] when the differential coefficient (d) takes real values that are limited Between 0.5 and -0.5, and its importance is that it allows modeling the short-term behavior of the time series through autoregressive and moving averages parameters, and the long-term behavior through the fractional integration parameter. It can be said that the time series (X_t) is subject to the (ARFIMA) model if:

$$\phi_p(L)(1-L)^d X_t = \theta_q(L)\,\varepsilon_t \tag{3.1}$$

Where:

$$\begin{aligned} \phi_p(L) &= 1 - \sum_{i=1}^{p} \phi_i L^i = 1 - \phi_1 L - \cdots - \phi_p L^p \\ \theta_q(L) &= 1 - \sum_{j=1}^{q} \theta_j L^j = 1 - \theta_1 L - \cdots - \theta_q L^q \end{aligned}$$

(ϕ_p, θ_q)They represent polynomials in L for the AR(p) and MA(q) parts of degree p and q of the model, respectively. (L) is called the backshift operator.

3.1. Estimation of ARFIMA Model

3.1.1. Semi-Parametric Estimation Methods

In this study, the GPH estimator proposed by [7] was used to estimate the long memory parameter. This approach is based on an approximation of a regression equation derived from the spectral density function's logarithm. The GPH estimate process consists of two steps, the first of which is the estimation of (d). This approach, which is based on spectral least squares regression, takes use of the sample form of the spectral density pole at the origin: $f_y(\lambda) \sim \lambda^{-2d}, \lambda \to 0$ We may write the spectral density function of a stationary model to illustrate this approach.

$$f_y(\lambda) = \left[4\sin^2\left(\frac{\lambda}{2}\right)\right]^{-d} f_\varepsilon(\lambda) \tag{3.2}$$

$f_\varepsilon(\lambda)$denotes the spectral density of (ε_t), which is considered to be a finite and continuous function on the interval $[-\pi, \pi]$. Equation (3.3) can be used to express the log-spectral density.

$$\log(f_y(\lambda)) = \log(f_\varepsilon(0)) - d\log\left[4\sin^2\left(\frac{\lambda}{2}\right)\right] + \log\frac{f_\varepsilon(\lambda)}{f_\varepsilon(0)} \tag{3.3}$$

The linear regression can be written as:

$$\log[I_x(\lambda_j)] = \alpha - d\log\left[4\sin^2\left(\frac{\lambda_j}{2}\right)\right] + e_j, \quad where\ \ e_j \sim \text{ iid } \left(-c, \pi^2/6\right)$$

Let $(I_x)(j)$ be the Fourier frequency-evaluated periodogram. $\lambda_j = \frac{2\pi_j}{T}; j = 1, 2, \ldots, m; T$ is the number of observations, and m is the number of Fourier frequencies.
Let $x_j = -\log\left[4\sin^2\left(\frac{\lambda}{2}\right)\right]$ the OLS estimate of the regression $\log(I_x(\lambda_j))$ on the constant α and x_j is the GPH estimator, It can be written as in equation (3.4)

$$\widehat{d}_{\text{GPH}} = \frac{\sum_{j=1}^{m}(x_j - \overline{x})\log[I_x(\lambda_j)]}{\sum_{j=1}^{m}(x_j - \overline{x})^2}, \qquad where,\ \overline{x} = \sum_{j=1}^{m} x_j/m \tag{3.4}$$

3.1.2. Parametric Estimation Methods

All parameters of the ARFIMA process are estimated using parametric techniques in a single phase. [14] proposed EML (Exact Maximum Likelihood). For estimating the parameters of the ARFIMA model, the EML technique is used in this study.
Let $X = (X_1\ , \ldots, X_T)\ ,\ X \sim N(0, \Sigma)$ be the original time series. The log-likelihood of the estimation is simply, Both the long memory parameter and the ARMA parameters can be estimated simultaneously using the EML method. The objective function of maximum likelihood is written as,

$$L_E(\Phi, \theta, d; X) = -\frac{T}{2}\log|\Sigma| - \frac{1}{2}X'\Sigma^{-1}X$$

and the EML estimator of d can be derived as:

$$\widehat{d}_{\text{EML}} = \arg\max\left[-\frac{T}{2}\log|\Sigma| - \frac{1}{2}X'\Sigma^{-1}X\right] \tag{3.5}$$

4. Artificial Neural Network

Artificial neural networks are considered one of the most important applications of artificial intelligence, which reflects the way of human thinking. It includes computational techniques designed to simulate the way the human brain performs a particular task and produces solutions that exceed the capabilities of traditional methods of prediction.

Artificial neural networks are defined as a computational system consisting of several (neurons) interconnected with each other and characterized by its dynamic nature in processing incoming data [11]. The neural network consists of three main layers: the input layer, the hidden layer, and the output layer (1). To use neural networks as a model to predict the future values of the time series, the variable at time (t) is a non-linear function of the same variable at previous time displacements $(x_{t-1}, x_{t-2}, \ldots, x_{t-k})$.

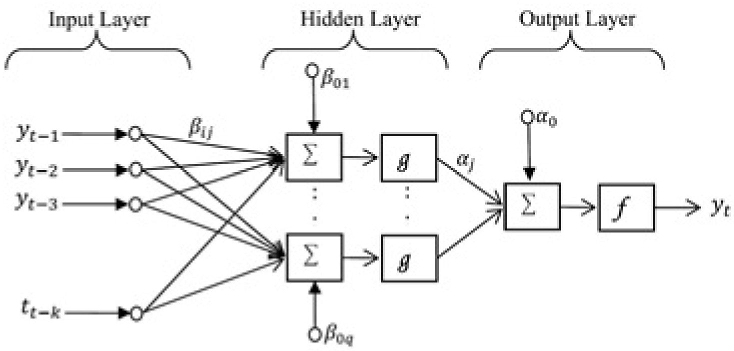

Figure 1: A schematic diagram of a three-layered feedforward network model with a single output and two hidden units.

So before modeling the time series using neural networks, the time series is converted into a set of patterns with degree (n-k+1) and each pattern has k inputs $(x_{t-1}, x_{t-2}, \ldots, x_{t-k})$ and one output is (xt). The data is usually divided into a training group, a validation group, and the remainder for testing. Several algorithms are used to train the neural network, the most important of which is the back-propagation algorithm to get the best weights for the neural network. Then the model is evaluated using a set of loss functions like MSE to get the best results.
The relationship between the input variable and the output variable as in equation (4.1).

$$x_t = f\left(\alpha_0 + \sum_{j=1}^{q} \alpha_j g\left(\beta_{0j} + \sum_{i=1}^{k} \beta_{ij} x_{t-i}\right)\right) \tag{4.1}$$

$$\text{sigmoid } (x) = \frac{1}{1+e^{-x}} \tag{4.2}$$

where k is the number of inputs of the neural network, α_j $(j = 0, 1, 2, \ldots, q)$ is the weight between jth hidden neuron and output neuron, α_0 is the bias unit of output neuron, $\beta \circ j$ is the bias at jth hidden neuron, β_{ij} is the weight between i th input and j th hidden neuron and q is the number of hidden neurons. Generally, sigmoid activation function (as in Eq. (4.2)) is used at hidden layer whereas linear activation function is used at output layer

5. Fuzzy time series model

[15] proposed the theory of fuzzy sets, which was the theoretical basis for the analysis of fuzzy time series. On the basis of this theory, (Song & Chissom ,1993) proposed the first fuzzy time series

model to predict the numbers of students enrolled at the University of Alabama. The main ability of the FTS model lies in its ability to deal with the problems of predicting linguistic variables. The basic definitions and concepts of the FTS model are summarized as follows (Song & Chissom ,1993): Define U as the universe of discourse, where $U = \{u_1, u_2, \ldots, u_n\}$. A fuzzy set A_i is defined as

$$A_i = f_{A_i}(u_1)/u_1 + f_{A_i}(u_2)/u_2 + \ldots + f_{A_i}(u_n)/u_n,$$

where f_{A_i} represents the membership function of the corresponding fuzzy set.

Definition 5.1 (Fuzzy time series). *Let $X(t)(t = \ldots, 0, 1, 2, \ldots)$ one of the subsets of real numbers, be the universe of discourse on which fuzzy sets $A_i(t)(t = \ldots, 0, 1, 2, \ldots)$ are defined. Then the $F(t)$ which is a collection of $A_i(t)$ is called a fuzzy time series on $Y(t)$.*

Definition 5.2. *Le $F(t)$ be a fuzzy time series. If $F(t)$ is a caused by $F(t-1)$,then this fuzzy logical relationship is represented by $F(t-1) \rightarrow F(t)$. and it is called first order fuzzy time series model.*

Definition 5.3. *Let $F(t)$ be a fuzzy time series. If $F(t)$ is caused by $F(t-1)$, $F(t-2)$,..., $F(t-m)$, then this fuzzy logical relationship is represented by: $F(t-m), \ldots, F(t-2, F(t-1) \rightarrow F(t)$.*

The prediction process according to the proposed FTS model is summarized in the following steps:

Step 1: Partition the universe of discourse into fuzzy intervals:

In this step fuzzy c-means (FCM) clustering is using to partitioning the universe of discourse. The FCM clustering technique was first introduced by [2]. Let the number of fuzzy sets (c) and $2 \leq c \leq n$, where n is the number of observations. In the FCM algorithm, the time series whose observations are crisp values are divided into fuzzy sets, and the center of each fuzzy set is determined, the degree to which each observation belongs to a fuzzy set is calculated depending on the centers of the fuzzy sets.

In this technique, fuzzy clustering is applied by minimizing the least squared errors within groups, The main step is to minimize the objective function which is defined by:

$$J_\beta(X, V, U) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{\beta} d^2(x_j, v_i) \tag{5.1}$$

Where X matrix of the data, V matrix centers of clusters and U membership values matrix, u_{ij} is the membership values, β is a constant $\beta > 1$ called the fuzzy index and $d(x_j, v_i)$ is a measure of similarity between an observation and the center of the corresponding fuzzy cluster. The constraints to minimized the objective function J_β is:

$$0 \leq u_{ij} \leq 1\ , \qquad \forall i, j,$$

$$0 < \sum_{j=1}^{n} u_{ij} \leq n, \qquad \forall i$$

$$\sum_{i=1}^{c} u_{ij} = 1, \qquad \forall j.$$

An iterative algorithm is used to update the values of (v_i, u_{ij}) and solve the minimization problem given above

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^{\beta} x_j}{\sum_{i=1}^{n} u_{ij}\beta} \tag{5.2}$$

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left[\frac{d(x_j,v_i)}{d(x_j,v_k)}\right]^{\frac{2}{(\beta-1)}}} \tag{5.3}$$

Step 2: Fuzzify the observations

After partitioning the universe of discourse into c clusters (fuzzy sets), the crisp time-series observations are converted into linguistic variables depending on the maximum degree of belonging for each observation within the U-matrix [1].

Step 3: Establish the fuzzy relationship with artificial neural network.

In this step we use tow order fuzzy time series and an example will be used to explain this step [1]. Because of using tow order fuzzy time series, two inputs are used in the neural network model (k=2) to get a lagged variable $F(t-2)$ and $F(t-1)$ from fuzzy time series $F(t)$.These series are shown in Table (1). The index (i) is taken for $F(t-2)$ and $F(t-1)$ series as input values for the neural network model and they are defined as in Table (1) by the whose titles input (1) and input (2), and the index (i) for $F(t)$ series is taken as target values for the neural network model. It is defined as in Table (1) by the whose titles target value. When taking the third observation, for example, the input values for the learning sample [A5, A6] are 5 and 6, and the target value for the learning sample is 2.

Step 5: Defuzzify results

The defuzzyfied forecasts are the center of the fuzzy cluster which correspond to fuzzy forecasts obtained by neural networks in the previous stage.

Table 1: Notations for two order fuzzy time series

Observation No.	***F(t-2)***	***F(t-1)***	***F(t)***	***Input (1)***	***Input (2)***	***Target***
1			A_5	-	-	-
2		A_5	A_6	-	-	-
3	A_5	A_6	A_2	***5***	***6***	***2***
4	A_6	A_2	A_8	***6***	***2***	***8***
5	A_2	A_8	A_4	***2***	***8***	***4***
6	A_8	A_4	A_3	***8***	***4***	***3***

6. Proposed hybrid (ARFIMA-FTS) model

In practice, most time series are characterized by containing linear and non-linear patterns, and for the difficulty of modeling these patterns, many models have been proposed that can explain these patterns through the process of combining linear and non-linear components, and they are called hybrid prediction models. Applications of hybrid methods appear in the literature by combining different methods, it can be an effective way to improve predictions. A lot of research was presented in this field, the most important of which is the hybrid model presented by the [16] to predict time series, where he clarified that the basic idea of combining linear and non-linear models is that the single model cannot adequately employ time series data and cannot clarify Precisely the properties and characteristics of the time series.

Therefore, to model time series that have linear and non-linear properties, a hybrid method is proposed to model long memory time series through [16] approach called ARFIMA-FTS. It is supposed that a time series can be considered composing of two components, which are a linear autocorrelation structure part and a non-linear part respectively. The model can be given as follows:

$$X_t = L_t + N_t \tag{6.1}$$

Where X_t is the originals time series and L_t defined as linear component and N_t defined as Non-linear component. The prediction steps according to the ARFIMA-FTS hybrid model can be summarized as follows:

(i) Estimation of the best linear ARFIMA model for the specified time series.

(ii) Calculating the residual values of the ARFIMA model calculated in the previous step according to the following equation:

$$e_t = X_t - \widehat{L}_t \tag{6.2}$$

Where e_t are the residuals of the ARFIMA model at time t

(iii) In this step, the non-linearity of the ARFIMA model's residuals is tested. as the process of building the hybrid model requires that these residuals be non-linear, and the test is done using graphs and the BDS test [5] for non-linearity

(iv) Estimating the non-linear residuals using the FTS model and obtaining the non-linear component of the hybrid model. The non-linear FTS model can be written as follows:

$$e_t = f\left(e_{t-1}, e_{t-2}, \ldots, e_{t-n}\right) + \varepsilon_t \tag{6.3}$$

Where f is the nonlinear function estimated using the FTS model and ε_t is the random error of the model. Using equation (6.3) we get the non-linear component of the time series $\widehat{N}$, so the predictive values of the time series are calculated according to the following equation:

$$\widehat{X}_t = \widehat{L}_t + \widehat{N}_t \tag{6.4}$$

The prediction process according to the ARFIMA-FTS hybrid model can be summarized in the following scheme:

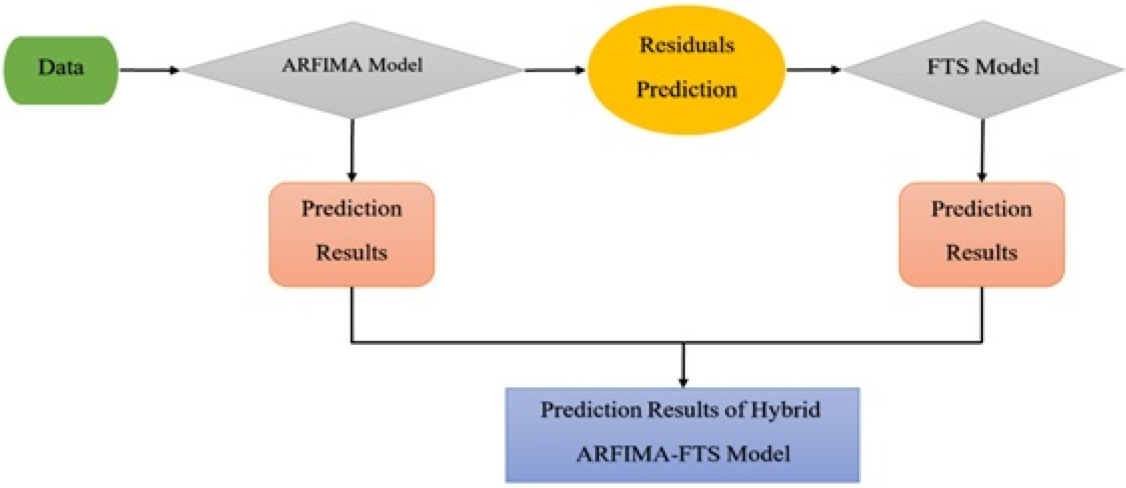

Figure 2: Schematic diagram of the Hybrid ARFIMA - FTS model.

7. Performance evaluation

Different indicators are employed in order to measure the forecasting performance of the proposed hybrid model.

I. RMSE It ascertains the amount of differences between values predicted by the model and the actual values. In other words, the RMSE shows the standard deviation of differences between forecasted and realized values. RMSE is a criterion for measuring accuracy to compare prediction errors of various models for a particular variable, as it is scale-dependent. The RMSE equation is as follows:

$$RMSE = \left(n^{-1} \sum_{t=1}^{n} \left(X_{(t)} - \widehat{X}_{(t)} \right)^2 \right) \tag{7.1}$$

II. MAPE MAPE is applied to measure a model forecasting accuracy. Describing accuracy as a percentage, this method is formulated as shown in the following equation

$$MAPE = n^{-1} \sum_{t=1}^{n} \left| X_{(t)} - \widehat{X}_{(t)} / X_{(t)} \right| \tag{7.2}$$

III. MAE The Mean Absolute Error (MAE)

$$MAE = n^{-1} \sum_{t=1}^{n} \left| X_{(t)} - \widehat{X}_{(t)} \right| \tag{7.3}$$

8. Applications

8.1. Time series Description

Our time series was obtained from the official website of the Iraqi Ministry of Health, and it represents the number of daily infections of the covid-19 virus for the period from 24/2/2020 to 11/8/2021 with 535 views. Table 2 shows the most important descriptive statistics for the time series.

Table 2: descriptive statistics for covid-19 time series in Iraq

Observations	***Kurtosis***	***Skewness***	***Std. Dev.***	***Minimum***	***Maximum***	***Median***	***Mean***
535	***4.026541***	***1.017997***	***2764.24***	***0***	***13515***	***2961***	***3254.08***

From the above table, it is noted that the highest infections were 13515, and the average is 3254. Figure 3 represents the graph of the time series, and it is noted that it is a non-linear series, as there are clear highs and lows for varying periods.

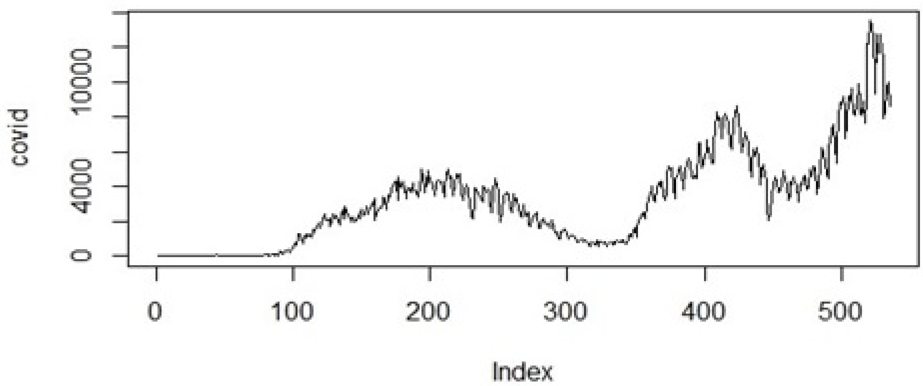

Figure 3: Plot of covid-19 time series in Iraq.

8.2. Long memory test

The existence of long memory for the time series is tested using equation (2.4).
The result of Hurst exponent (H = 0.831), which is close to one, as it proves that the series has a long memory, and this is also confirmed by the graph of the ACF and PACF functions which are shown in Figure 4.

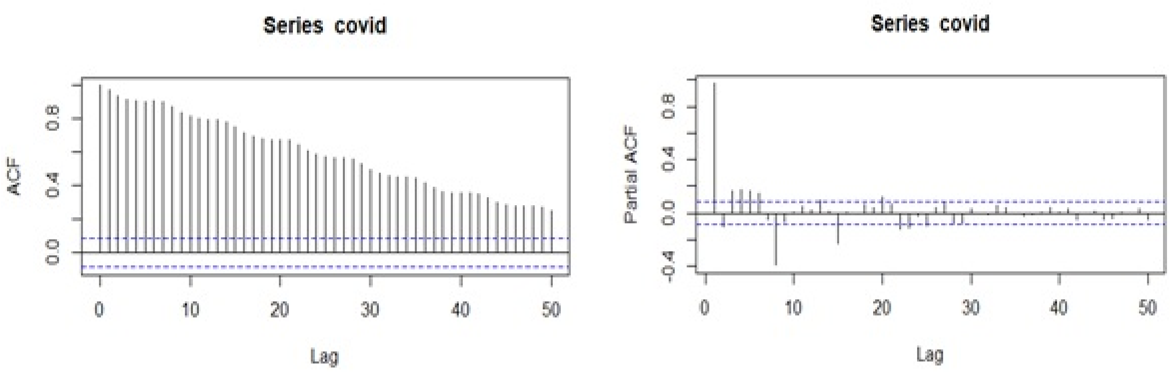

Figure 4: ACF &PACF for covid-19 time series in Iraq.

8.3. Nonstationary Testing

To test the stationarity of the time series, the ADF and PP tests are used. The null hypothesis says that the time series has a unit root. And the results shown in Table 3 show that the P-value for

Table 3: Result of the ADF &PP tests for stationarity

P-Value	Critical Value	Test Sta.	Test
0.452	-1.941	-0.163	ADF
0.574	-1.941	-0.309	PP

the two tests is greater than 0.05, so we cannot reject the null hypothesis and that the time series has a unit root and is not stationary.

Therefore, fractional differences will be taken to make the time series stationary because it has the long memory property. The value of the fractional difference (d) was estimated using the EML and GPH methods as in equations (3.4) and (3.5), the two results were close to both methods, the estimated value using the EML method = 0.495 and using the GPH method = 0.49. The results are shown in Table 4 the time series stability test after taking the fractional difference.

Table 4: Result of the ADF &PP tests after taking the fractional difference

P-Value	Critical Value	Test Sta.	Test
0.0003	-1.941	-2.983	ADF
0.0000	-1.941	-15.142	PP

The above results confirm that the time series became stationary after taking the fractional difference.

8.4. Applying ARFIMA Model

After the time series has become stationary, the appropriate model for the time series is identified by reconciling a number of the proposed models, and a comparison is made between these models using the comparison criteria AIC, BIC, HQ. The results in Table 5 showed that the model ARFIMA (1,0.495,2) has the lowest values for these criteria.

Table 5: Values of the comparison criteria for the candidate models

Model	AIC	BIC	HQ
ARFIMA(0,d,0)	15.90463	15.92064	15.91089
ARFIMA(0,d,1)	15.59934	15.62335	15.60873
ARFIMA(0,d,2)	15.54364	15.57566	15.55617
ARFIMA(1,d,0)	15.53278	15.55679	15.54217
ARFIMA(1,d,1)	15.53252	15.56453	15.54504
ARFIMA(1,d,2)	15.45748	15.4975	15.47314
ARFIMA(2,d,0)	15.53303	15.56505	15.54556
ARFIMA(2,d,1)	15.53606	15.57608	15.55172
ARFIMA(2,d,2)	15.45823	15.50626	15.47702

The autoregressive (p) and moving averages (q) parameters of the model ARFIMA(1,0.495,2) were estimated using the Maximum Likelihood method and the estimated values of these parameters are shown in Table 6

Table 6: Result of the estimated ARFIMA(1,0.495,2) parameters

Variable	***Coefficient***	***Std. Error***	***t-Statistic***	***Prob.***
AR(1)	***0.988211***	***0.006379***	***154.9051***	***0.000***
MA(1)	***-0.50264***	***0.032974***	***-15.2436***	***0.000***
MA(2)	***-0.363731***	***0.027336***	***-13.30598***	***0.000***

The above table shows that all parameters of the model are significant. The coefficients of the moving averages are always stationary, to be invertible the following conditions must be proven:

$$\theta_1 + \theta_2 = -0.866371 < 1$$
$$\theta_2 - \theta_1 = 0.138909 < 1$$
$$|\theta_2| = 0.3637 < 1$$

The autoregressive coefficients are always invertible, to be stationary the following conditions must be proven:

$$|\varnothing_1| = 0.988211 < 1$$

When stationarity and invertibility are confirmed, the model can be considered efficient in the estimation process. At the final of applying ARFIMA model, the residuals of the chosen model must be analyzed and its validity for future prediction. Because the time series is non-linear, it is assumed that the non-linear patterns of the series appear in the residuals of the chosen model.
The BDS test was conducted to find out the presence of non-linear patterns and the results were shown in Table 7.

Table 7: BDS test results on residuals from ARFIMA model.

Dimension	***BDS Statistic***	***Std. Error***	***z-Statistic***	***Prob.***
2	***0.026918***	***0.004725***	***5.69745***	***0.000***
3	***0.062655***	***0.007541***	***8.308275***	***0.000***
4	***0.091455***	***0.009023***	***10.13566***	***0.000***
5	***0.110727***	***0.009451***	***11.71568***	***0.000***
6	***0.120449***	***0.009161***	***13.14828***	***0.000***

The above table shows that the P-value is less than 0.05 for all dimension levels, and this means that the null hypothesis is not rejected, which states that the residuals of the model are non-linear. Therefore, these patterns must be analyzed to take advantage of them in the forecasting process.

8.5. Applying FTS Model

8.5.1. Partition the universe of discourse into fuzzy intervals

The fuzzy time series model (FTS) is applied to the residuals of the ARFIMA (1,0.495,2) model considering these residuals as a time series that bears the non-linear patterns of the model. In the first step, the universe of discourse is defined and divided based on the FCM algorithm into a number of clusters (c), as in equations (5.1) and (5.2), In this research, (25) clusters were identified to fit the data of the residual series. Table 8 shows the centers of the clusters.

Table 8: Centers of the (25) clusters.

Cluster Number	*Cluster Center*	*Cluster Number*	*Cluster Center*	*Cluster Number*	*Cluster Center*	*Cluster Number*	*Cluster Center*	*Cluster Number*	*Cluster Center*
cluster1	***-2055.14***	***cluster6***	***-513.057***	***cluster11***	***-35.4872***	***cluster16***	***269.364***	***cluster21***	***767.3102***
cluster2	***-1333.33***	***cluster7***	***-371.913***	***cluster12***	***4.520552***	***cluster17***	***345.2759***	***cluster22***	***934.9524***
cluster3	***-1002.67***	***cluster8***	***-269.272***	***cluster13***	***52.50986***	***cluster18***	***436.2801***	***cluster23***	***1240.573***
cluster4	***-806.05***	***cluster9***	***-172.044***	***cluster14***	***125.4693***	***cluster19***	***562.879***	***cluster24***	***1614.334***
cluster5	***-631.293***	***cluster10***	***-99.074***	***cluster15***	***203.5642***	***cluster20***	***654.1948***	***cluster25***	***2784.946***

Fuzziness of the time series depending on the degree of membership in the matrix (U) which calculated by equation (5.3). The crisp values of the residual series are converted to the corresponding cluster number to the maximum degree of belonging to each deterministic value in the time series, so that we have a time series whose observations are linguistic values. Table 9, represents the first ten values and the maximum degree of belonging of each observation to any cluster. And converting the crisp time series into a (FTS).

Table 9: fuzzifying first ten values of residual series.

Actual	***max membership degree***	***cluster***	***FTS***
-97.255	***0.997698649***	***10***	***A10***
-2.22187	***0.934918967***	***12***	***A12***
2.597373	***0.994925302***	***12***	***A12***
-3.1807	***0.914534119***	***12***	***A12***
-2.10501	***0.937209181***	***12***	***A12***
4.761055	***0.999922362***	***12***	***A12***
-0.99515	***0.956859121***	***12***	***A12***
1.556711	***0.987824689***	***12***	***A12***
-1.36968	***0.950652326***	***12***	***A12***
-3.0063	***0.918453163***	***12***	***A12***

8.5.2. Establish the fuzzy relationship with artificial neural network

This step is considered the most important stage of applying the fuzzy time series model. After obtaining the FTS from the previous step, we find FTS-1 and FTS-2, which represent the inputs of the artificial neural network. The input layer of the network in this research consists of two nodes, the hidden layer of the network consists of five nodes, and the output layer consists of one node, which represents the predicted value. for the fuzzy time series.

8.5.3. Defuzzify results

After obtaining the FTS predictions, the predictions are defuzzed by replacing the prediction value with their corresponding cluster centers for all time series observations. Table 10 shows the first ten predicted values that have been defuzzed.

Table 10: the first ten predicted values for residual series.

Actual FTS	*Predict FTS*	*Cluster Index*	*Predict value*
A10	*A12*	*12*	*4.520552*
A12	*A12*	*12*	*4.520552*
A12	*A12*	*12*	*4.520552*
A12	*A12*	*12*	*4.520552*
A12	*A12*	*12*	*4.520552*
A12	*A12*	*12*	*4.520552*
A12	*A12*	*12*	*4.520552*
A12	*A12*	*12*	*4.520552*
A12	*A12*	*12*	*4.520552*
A12	*A12*	*12*	*4.520552*

8.6. Applying Hybrid Model

After obtaining the predictive values using the ARFIMA model (1,0.495,2), which represents the linear component $\widehat{L}_t$ and the predictive values of the residuals using the FTS model, which represent the nonlinear part $\widehat{N}_t$, The linear and non-linear component are collected as in equation (6.4) to get the final predictions of the hybrid model. Figure 5 shows the plot of the real time series with its predictive values using the hybrid model

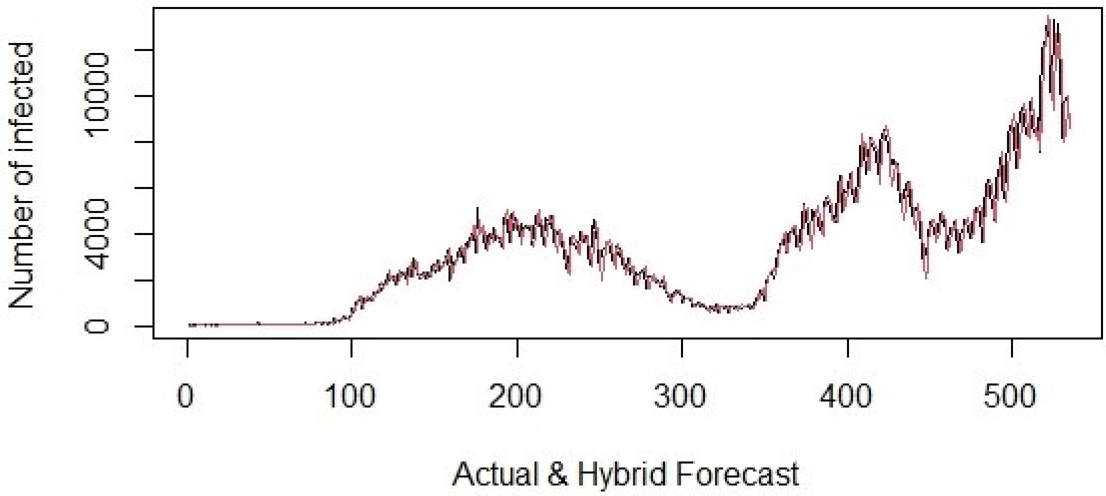

Figure 5: Actual & forecasting time series by hybrid ARFIMA-FTS model.

9. Results and discussion

In this research a new hybrid model is proposed to predict the number of people infected with covid-19 virus in Iraq. By combining the linear component of the time series, which was analyzed using the ARFIMA (1,0.495,2) model with the non-linear component analyzed by the FTS model for the residuals of the ARFIMA model (1,0.495,2). It was found through the research that the time series under study is not stationary and contains of linear and non-linear patterns. The not stationarity was analyzed by taking the fractional difference d = 0.495 which was estimated by EMLE and GPH methods. The prediction results using the hybrid model were more accurate than the predictions obtained from the ARFIMA (1,0.495,2) model. Table 11 shows the results of the prediction accuracy criteria RMSE, MAPE, and MAE.

Table 11: The comparison of the results.

Model	***RMSE***	***MAPE***	***MAE***
ARFIMA (1,0.495,2)	***532.5365***	***0.2274474***	***334.2731***
ARFIMA-FTS	***514.7011***	***0.1755291***	***320.1998***

The ARFIMA-FTS hybrid model has the lowest values for these criteria than the singular ARFIMA (1,0.495,2) model. Therefore, the hybrid model can be used to predict the future series of the number of people infected with COVID-19 virus in Iraq. It can also be used to predict time series that contain linear and non-linear components.

References

[1] C.H. Aladag, U. Yolcu and E. Egrioglu, *A high order fuzzy time series forecasting model based on adaptive expectation and artificial neural networks*, Math. Comput. Simul. 81 (2020) 875–882.

[2] JC. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, NY., 1981.

[3] G.E.P. Box and G.M. Jenkins, *Time Series Analysis; Forecasting and Control*, 2nd ed., Holden-Day, San Francisco, 1976.

[4] P.J. Brockwell and R.A. Davis,*Time Series Theory and Method*, Springer, 1995.

[5] W.A. Broock, J.A. Scheinkman, W.D. Dechert and B. LeBaron, *A test for independence based on the correlation dimension*, Econom. Rev. 15 (1996) 197–235.

[6] P.H. Franses, D.J.C. van Dijk and A. Opschoor, *Time Series Models for Business and Economic Forecasting*, 2nd Edition, Cambridge University Press, 2014.

[7] J. Geweke and S. Porter-Hudak, *The estimation and application of long memory Time series models*, J. Time Ser. Anal. 4 (1983) 221–238.

[8] C. W. J. Granger and R. Joyeux, *An Introduction to Long Memory Time Series Models and Fractional Differencing*, J. Time Ser. Anal. 1 (1980) 15–29.

[9] J. Hosking,*Fractional differencing*, Biomet. 68 (1981) 165–176.

[10] H.R. Hurst,*Long-term storage in reservoirs*, Trans. Amer. Soc. Civil Eng. 116 (1951) 770–799.

[11] S. Kumar, *Neural Networks: A Classroom Approach*, Tata McGraw-Hill Publishing Company Limited, New Delhi, 2004.

[12] A.W. Lo, *Long-term memory in stock market prices*, Econom. 59 (1991) 1279–1313.

[13] B.B. Mandelbrot and J.R. Wallis, *Noah, Joseph and operational hydrology*, Water Resources Res. 4 (1968) 909–918.

[14] F. Sowell, *Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models*, J. Econom. 53 (1992) 165–188.

[15] L.A. Zadeh, *Fuzzy sets*, Inf. Cont. 8(1965) 338–353.

[16] G.P. Zhang, *Time series forecasting using a hybrid ARIMA and neural network model*, Neurocomput 50 (2003) 159–175.