Current Journal of Applied Science and Technology

39(24): 79-90, 2020; Article no.CJAST.60134
ISSN: 2457-1024
(Past name: British Journal of Applied Science & Technology, Past ISSN: 2231-0843, NLM ID: 101664541)

Application of Weighted Fuzzy Time Series Model to Forecast Epidemic Injuries

Hala Ahmed Abdul-Moneim[1,2*]

[1]*Department of Mathematics, Faculty of Science, Minia University, Egypt.*
[2]*Department of Mathematics, Al Dar University College, Jazan University, Kingdom of Saudi Arabia.*

Author's contribution

The sole author designed, analyzed and interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/CJAST/2020/v39i2430875
Editor(s):
(1) Dr. Koji Nagata, Korea Advanced Institute of Science and Technology, Korea.
(2) Dr. Santiago Silvestre, Universitat Politècnica de Catalunya, Spain.
Reviewers:
(1) Mohammad Reza Omidi, Iran.
(2) Jean Pierre Namahoro, China University of Geosciences, China.
(3) Sambit Satpathy, National Institute of Technology Agartala, India.
Complete Peer review History: http://www.sdiarticle4.com/review-history/60134

Method Article

Received 29 July 2020
Accepted 14 August 2020
Published 20 August 2020

ABSTRACT

Aims: It is important to predict the amount of COVID-19 injuries. Since the first suspected case of novel coronavirus (2019-nCoV) on December 1st, 2019, in Wuhan, Hubei Province, China, a total of 40,235 confirmed cases and 909 deaths have been reported in China up to February 10, 2020, evoking fear locally and internationally. Here, based on the large amounts of daily publicly available epidemiological data and the need to make an accurate prediction of future behavior requires the definition of powerful and effective techniques capable of inferring random dependency between the past and the future from observations. In this paper, we apply a rewarding model to predict injuries in areas where COVID-19 is, especially in the Arab region. This forecast uses epidemic injuries data from March 2nd, 2020 to July 20th, 2020 in Saudi Arabia.
Methodology: We propose the use of weighted fuzzy time series techniques (WFTS) and weighted non-stationary fuzzy time series techniques (WNSFTS) to be compared with the classical Auto-Regressive Integrated Moving Average (ARIMA) statistical method. The available data is not a stationary and should therefore be converted first to stationary to forecast it with (ARIMA) and (WFTS) techniques. We do a log transform and differencing on our injuries dataset.
Results: When we examine the original data by Dickey-Fuller Test (DFT) to get p-value, we find it is equal to 0.646, it is more than 0.05 which implies the non-stationarity. The mean square error

**Corresponding author: E-mail: h4tallaabdulmoneim@yahoo.com;*

(MSE), the root mean square error (RMSE) and normalization root mean square error (NRMSE), are applied to compare the accuracy of the methods. The results show that WFTS methods give good services for predicting epidemic injuries in the territory by COVID-19.

Conclusion: The use of Weighted Non Stationary Fuzzy Time Series (WNSFTS) in forecasting epidemic injuries problem can provide significantly better results because it is able to predict the infected cases at the next time and achieve great predictive accuracy.

Keywords: Injuries; Arabic Zone; WFTS; ARIMA; WNSFTS; RMSE.

1. INTRODUCTION

1.1 Context

The newest zoonotic Coronavirus disease that crossed species to affect humans and spread in an unprecedented manner is COVID-19, as named by World Health Organization (WHO) on 11th February 2020. More than 300,000 cases have been confirmed across 166 countries in six continents, causing more than 13,000 deaths in less than three months.

For Arabic Zone in Africa as mentioned in [1] the first case in Egypt was discovered at Cairo International Airport involving a Chinese national on 14th February, then on 6th March, the Egyptian Health Ministry and WHO confirmed 12 new cases of coronavirus infection. The infected persons were among the Egyptian staff aboard the Nile cruise ship which was travelling from Aswan to Luxor. Also on March 7th, 2020, health authorities announced that 45 people on board had tested positive, and that the ship had been placed in quarantine at a dock in Luxor.

The first case in Morocco was confirmed to have spread on March 2nd, 2020, when the first case COVID-19 case was confirmed in Casablanca where it involved a Moroccan expatriate residing in Bergamo, Italy who arrived from Italy on 27th February. An 89-year-old woman Moroccan residing in Italy who had returned to Morocco on 25th February from Bologna, Italy was confirmed as the second case. In mid-March the Government closed schools and suspended international passenger flights as the outbreak widened in Morocco.

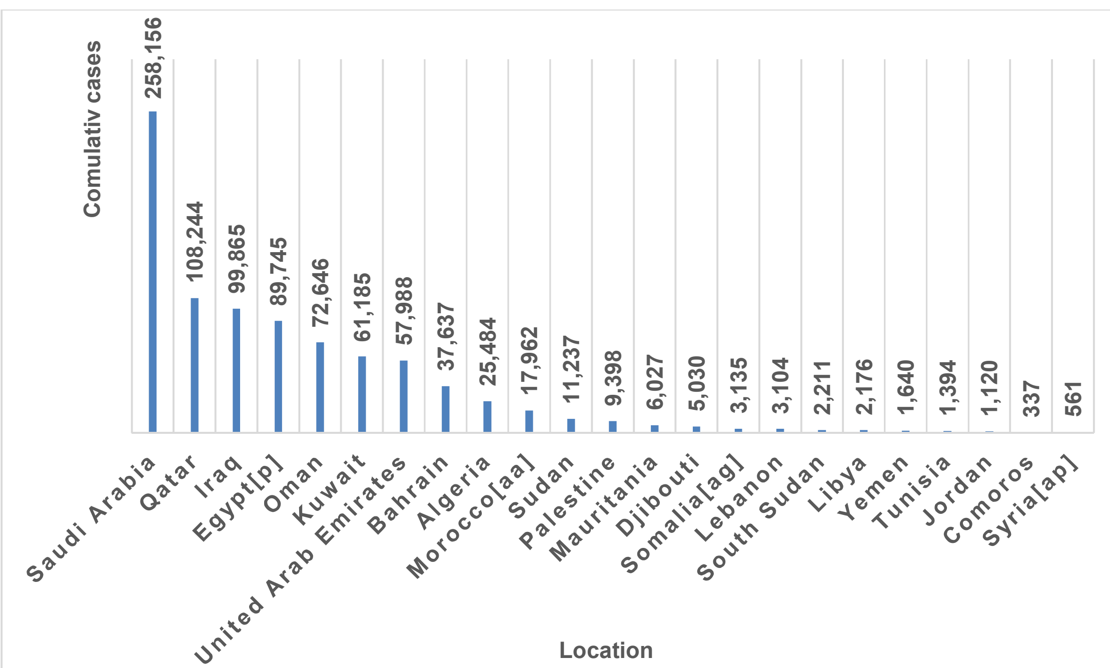


Fig. 1. Cumulative Cases in Arabic Countries until July 23rd 20**20**

The first case in Algeria was confirmed to have spread in February 2020 that was an Italian man who arrived on 17th February. After that on 28th February, Algeria deported him back to Italy via a special flight. Fig. 1 shows the cumulative number of confirmed cases in 23 Arabic countries until 23rd July 2020.

In Asia for the biggest, most important Arabic Zone, Saudi Arabia, the first case, as mentioned in [2] was confirmed on 2nd March that is a Saudi national returning from Iran via Bahrain. So the government agencies started to take the necessary measures to reduce the spread of the virus. Then on Friday, 20th March, Saudi Arabia announced it is going to suspend all domestic flights, buses, taxis and trains for 14 days amid global coronavirus outbreak and after a week on 26th March, authorities announced a total lockdown in Riyadh, Mecca and Medina, plus a nation-wide curfew.

When we analyze the effect of lockdowns in several countries we find that the complete lockdowns helped to reduce the number of affected cases significantly. However, although the lockdown has been gradually lifted, the use of the media and social media to disseminate awareness and information about prevention, screening, testing and policies in different countries and the power of health care services has been more effective in reducing the spread of the disease. Such as the lockdown in Saudi Arabia was considered to be on 15th March 2020. Fig. 2 shows that the maximum injured number was happened after the partially lockdown began in June and then the numbers started to fluctuate day after day in July.

1.2 Fuzzy Time Series Forecasting

Corona Statistics around the world and My health statistic to directly monitor cases globally and locally are modeled by a simple visibility graph. Fig. 3 shows the daily real-time growth rates estimating detected cases in territories where COVID-19 until July 20th in Saudi Arabia. Time dependencies that cause identical points of time to belong to different categories or predict different behavior are constituting a so called “Time series”. Time Series model is one of the commonly used statistical models in analytics where, previous data and recent data are studied in order to make predictions. It has wide applications in finance, construction costs prediction, and some other fields. Autoregressive integrated and moving average (ARIMA) and seasonal ARIMA models are developed that can study linear and stationary time series [3].

It became increasingly clear that linear models are not adapted to many real applications [4]. Song, et al.[5] defined and studied the observations of special dynamic process with linguistic values. This dynamic process is referred to “fuzzy time series”. Fuzzy models or sets are mathematical tools of representing vagueness and imprecise information (hence the term fuzzy).

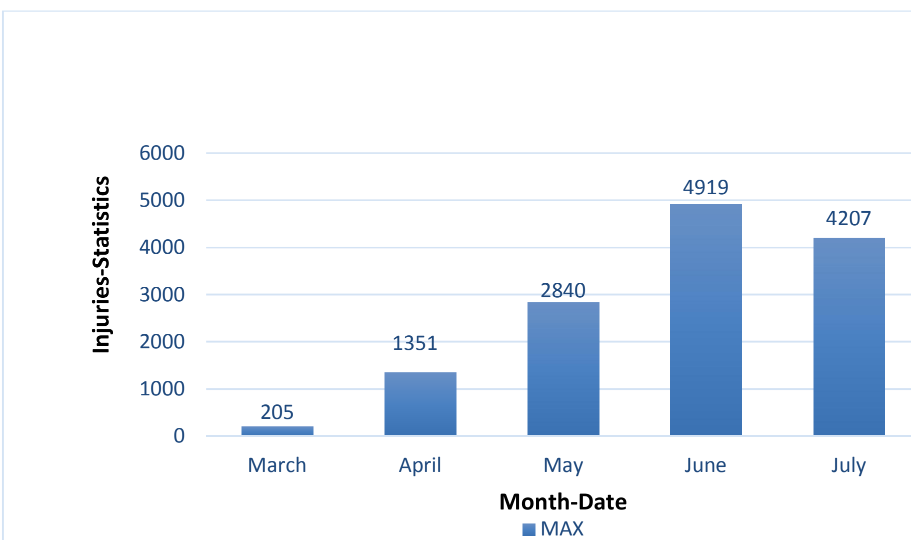


Fig. 2. The maximum number 4919 of injures was in June in territories where COVID-19 in Saudi Arabia

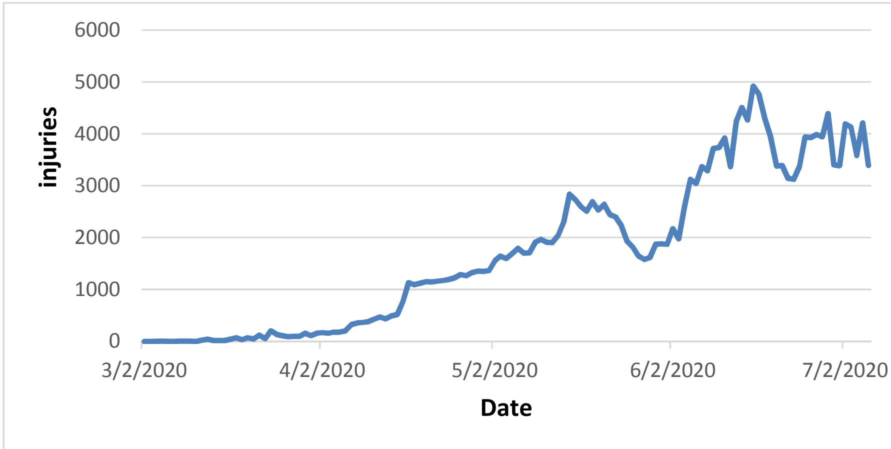


Fig. 3. The daily real-time growth rates estimating detected cases in territories where COVID-19 to July 20th 2020

Some key features distinguish the Fuzzy Time Series: simplicity and scalability [6]. The use of fuzzy logic has been applied in the literature for time series forecasting and provides a significant improvement over the traditional statistical methods because it is non-linear and it is able to approximate any complex dynamical systems better than linear statistical models [7]. In academic enrollments, Preetika Saxena and Santhosh Easo [8] proposed a method based on fuzzy time series, which gives the higher forecasting accuracy rate than the existing methods. On an Internet of Things (IoT)-based analysis system, Satpathy, et al. [9] developed a fuzzy classifier which results in a lower execution time in comparison with that of previous models, such as K-nearest neighbor, decision tree, support vector machine, and naive Bayes. In the field-programmable gate array (FPGA) analysis system with the help of fuzzy system, Satpathy, et al. [10] designed a novel method for prediction multiple diseases in a rural area. In cooking, Fuzzy logic control enables to obtain a cooking time for different type of rice and for different quantity of water [11]. A fuzzy logic technique and matlab graphical user interface technique helped Sambit and Ambika [12] to checking out the on, off status of rice cooker. If there is any trouble then it smartly detect it and solve it automatically. In a solar forecasting problem the use of Fuzzy Time Series (FTS) provides significantly better results than other approaches in the literature [13]. Based on Weighted Fuzzy Time Series (WFTS), K. A'yun,et al.[14] predict Trans Jogja passenger and stated that forecasting process with WFTS model is better than forecast with fuzzy time series model. FTS Markov chain (FTSMC) model urged Mahmod Othman, et al. [15] to predict a daily air pollution index model based on a grid method with an optimal number of partitions, which can greatly develop the model accuracy for air pollution and stated that the proposed forecasting method has produced a higher prediction accuracy as compared to some FTS models. There have been still more areas where FTS have been applied.

1.3 Our Contribution

In this paper, we propose and investigate the use of some methods based on Fuzzy Time Series(FTS) to perform epidemic injuries forecasting in Arabic Zone in territories where COVID-19 epidemic injuries forecasting is a very hard forecasting problem. FTS methods are providing good results for hard forecasting problems [16] or stock market forecasting [17]. We introduce the application of one-variable FTS in epidemic injuries forecasting, describing the data and discussing preprocessing steps. The high-order FTS methods and the Weighted FTS methods are compared with ARIMA forecasting models widely used to approach epidemic injuries forecasting. We evaluate the performance of FTS methods and different forecasting techniques to solve epidemic injuries forecasting problem. The results show that FTS methods are able to achieve significant improvements in forecasting accuracy and performance if compared to the other methods taken for the comparison in the present study. We forecast the epidemic injuries in Saudi Arabia, from 2nd March 2020 to 20th July 2020.

2. REVIEW

2.1 Prediction and Forecasting

In statistics, prediction is a part of statistical inference. One particular approach to such inference is known as predictive inference, but the prediction can be undertaken within any of the several approaches to statistical inference. When information is transferred across time, often to specific points in time, the process is known as forecasting. Prediction is more general term. Both refer to formal statistical methods employing time series.

2.2 Time Series

Time Series is a forecasting method . It is a sequence of numerical data points in successive order. Such as values detecting the cases that has affected by a disease, over a specified period with data points recorded at regular intervals. We use historical values and associated patterns to predict future activity. Malhotra et al. [18] defined Time-Series as a vector. A stationary time series is one whose statistical properties such as mean, variance, autocorrelation are all constant over time. Such statistics are useful as descriptors of future behavior only if the series is stationary. Non-stationary data should be first converted into stationary data. In this study, we use Dickey-Fuller Test to test stationarity. When the p-value is less than a certain value (α=.05), the data is stationary.

2.3 ACF and PACF

ACF is an auto-correlation function which gives us values of auto-correlation of any series with its lagged values, it describes how well the present value of the series is related with its past values. It's a 'complete auto-correlation plot' [19].

PACF is a partial auto-correlation function, it finds correlation of the residuals with the next lag value hence 'partial' and not 'complete' as we remove already found variations before we find the next correlation. So if there is any hidden information in the residual which can be modeled by the next lag, we might get a good correlation and we will keep that next lag as a feature while modeling [19].

2.4 Models

Models can be classified into:

2.4.1 Compartmental models

Compartmental models its basis is the assumption that the entire population is divided into groups, ie, compartments [20]. Compartmental models simplify the mathematical modelling of infectious diseases as in [21]. The simplest compartmental model is the SIR model, and many models are derivatives of this basic form. SIR model is reasonably predictive for infectious diseases that are transmitted from human to human. SEIR is an epidemiological model consists of four possible states: Susceptible [S], Exposed or latent [E], Infectious [I] or Removed [R]. The proportion of a population in each state is governed by the rate of change between each, β ([S] to [E]), σ ([E] to [I]) and γ ([I] to [R]). Richa Tripathi, et al. [22] trained the machine learning regression models with features as networks parameters on various model networks and the corresponding parameter of stability of the disease stage R_0 values as output labels. After training the others predicted R_0test networks. Zifeng Yang , et al.[23] used a modified susceptible exposed-infected-removed (SEIR) epidemiological model that incorporates Kthe domestic migration data before and after 23rd January and the most recent COVID-19 epidemiological data to predict the epidemic progression. Yi-Cheng Chen,et al.[24] presented mathematical and numerical analyses that address important questions for COVID-19. Can Hou, et al.[25] investigated that, by reducing the contact rate of latent individuals, interventions such as quarantine and isolation can effectively reduce the potential peak number of COVID-19 infections and delay the time of peak infection. Abdullah Murhaf Al-Khani and Mohamed Abdelghafour Khalifa [26] estimate that by the time the Hajj season commences in Saudi Arabia, the pandemic will be in the midst of its deceleration phase (phase 3). Haitham khoj and Alaa F Mujallad [27] employ SIR model to forecast the peak of COVID-19 progression and an estimation of it is end in Saudi Arabia. Peiliang Sun and ang Li [28] present a new model produces some consistent estimates on the total death toll if community spread is greatly suppressed after 23rd March in UK.

2.4.2 Statistical models

Models for time series data can represent different stochastic processes. There is a specific collection of methods and techniques that well suited for predicting the value of a dependent variable according to time such as the autoregressive (AR) models, the integrated (I) models, and the moving average (MA) models. These three classes depend linearly on previous data points [29]. Combinations of these ideas produce autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) models. The autoregressive fractionally integrated moving average (ARFIMA) model generalizes the former three. Rob J Hyndman and George Athanasopoulos [30] provided a comprehensive introduction to forecasting methods and presented enough information about each method.

2.4.3 Fuzzy time series representation

Not all time series are known to be stationary, meaning that their mean values are (more or less) constant. In non-stationary series, the mean changes over time. Zadeh [31] proposed the Fuzzy Logic and state that a certain element may belong and simultaneously do not belong to the same set at certain levels, such that the membership is a value in the interval [0, 1]. Song and Chisson [32] are the pioneers in working on fuzzy time series. Chen [33] presented the idea of dividing the Universe of Discourse from time series in intervals/partitions (the fuzzy sets), and how each area behaves (extracting rules through the time series patterns) that is create a linguistic variable to represent the numerical time series, and these areas will be the linguistic terms of our variable.

We drive a simple Fuzzy time series prediction example from [7] as follows: given X, a Numerical Variable, X ∈ℝ — for instance an height measure — its Universe of Discourse, abbreviated to U, such that (U = [min(X), max(X)]). We define (U = [20,220]) and the linguistic variable Ã as: (Ã = {“very small”, “small”, “short”, “medium”, “tall”, “very tall”}) or when we use a 10 partitions scheme, (Ã = {A0, A1, , A9 }). Using the Chen’s method the fuzzyfied values (for example, from 20 ,,220) can be F(t) = { A1, A2, A2, A3, A4, A4, A4, A4, A5, A5, A5, A4, A4, A4, A4, A4, A5, A7, A8, A8, A8, A7 }. A temporal patterns appear sequentially on fuzzy time series F(t) and have the format Precedent →Consequent, where the precedent indicates a fuzzy set on time t and the consequent the fuzzy set that appears soon after on time t+1.The generated temporal patterns will be: A1→A2, A2→A2, A2→A3, A3→A4, A4→A4, …, A8→A8, A8→A7 and we will group them by its precedents :

(A1 →A2), (A2 →A2,A3), (A3 →A4), (A4 →A4,A5), (A5 →A4,A5,A7), (A7 →A8), (A8 →A7,A8). These groups are the generated rules. Now we want to predict the next instant, x (t + 1). For example, for (t = 1/4/2020) the value is (x(t) = 157). Fuzzyfying x (t) the set is A7, so (f(t) = A7) then we have the rule A7 → A8. Then (f (t + 1) = A8). This simple model has the advantages: a) it is very easy to parallelize/distribute, what makes it very attractive for big data; b) it is very easy to update, what makes it very attractive for frequently-changing data.

3. MATERIALS AND METHODS

3.1 Data Set Description and the Applied Methods

Data on COVID-19 were obtained from the Saudi Arabian Ministry of Health and the World Health Organization. Statistics from March 2nd, 2020 to July 20th, 2020 were retrieved from [34]. The dataset consists of daily case reports and daily time series summary tables. Data must be stationared for the applied methods included in the following (i) and (ii). Methods included in the following (iii), it is for non-stationared data. The applied methods that we will implement (individually) include:

- The classical statistical methods (AR), (MA) and ARIMA (as described in [35]). We applied these methods to compare the performance of the other used method with it.
- The probability weighted fuzzy time series (PWFTS), the high order fuzzy time series (HOFTS) and weighted fuzzy time series (WFTS) methods (as described in [13]).
- The high order non-stationary fuzzy time series (HONSFTS) and weighted non-stationary fuzzy time series (WNSFTS).

In order to evaluate the results, we apply (individually) the statistical metric mean square error (MSE), root mean square error (RMSE), and its normalized form (NRMSE). They are defined by (1)-(3).

$$MSE = \frac{1}{N}\sum_{t=1}^{N}[\hat{y}(t) - y(t)]^2 \tag{1}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}[\hat{y}(t) - y(t)]^2} \tag{2}$$

$$NRMSE = \frac{\sqrt{\frac{1}{N}\sum_{t=1}^{N}[\hat{y}(t) - y(t)]^2}}{y_{max} - y_{min}} \tag{3}$$

3.2 Determining Stationary Data

Before predicting data must be stationared, raw data taken from daily rates estimating detected injuries cases from 2nd March 2020, to 20th July 2020 in Saudi Arabia. Another method to examine data if it is stationary, it is divided into groups and compare its mean and its standard deviation. Table 1 shows the statistics properties for each group.

It is clear that the mean, variance, are not the same so data are not stationer. Fig. 4(a) shows non stationary data. Thus data has to be subject to some transfers. Injuries values can be seen as time series. Better results in forecasting can be obtained if these components are eliminated or mitigated [35]. Fig. 4(b) shows both trend and seasonality are modeled separately and we process the remaining part of the series (Residuals). Fig.4(c) shows the time series after had stationaried by using a quick log transform and differencing on our dataset. Test stationarity by Dickey-Fuller Test gives that the p-value of the residuals data was less than 0.05 it was 0.00042. Fig. 4(d) shows the auto correlation function ACF and partial auto correlation function PACF of injuries data after the differencing process.

4. EXPERIMENTS

This work was implemented using python programming language version 3.7. Our experiments were carried out using a number of python libraries some of which include:

- Pandas Library – For reading excel data.
- PyFTS Library – To implement fuzzy time series models (FTS).
- Matplot Library – For visualizing the applied algorithms.
- Numpy Library – For calculations.
- Statsmodels Library – For viewing the correlation between our dataset by ACF and PACF, testing stationarity and to implement ARIMA models.

4.1 Forecasting Using ARIMA

In this study, the experiments were applied to evaluate the performance of the applied forecasting methods in terms of accuracy and elapsed time. We apply ARIMA and special cases of ARIMA models are AutoRegression (AR) which is $ARIMA(p, 0,0)$ and Moving Average (MA) which is $ARIMA(0,0, q)$. Table 3 shows the results of forecasting in terms of accuracy sorted by NRMSE. Fig. 5 (c) shows the forecasting in red colors using the applied ARIMA methods.

4.2 Forecasting Using Fuzzy Time Series

The forecasting process of injuries with stationary data that has been obtained is conducted by the steps as stated in [13] and [14]. Fuzzy time series (FTS) has a different representation of a time series. If a time series are represented by real numbers, fuzzy time series are composed by fuzzy sets which form the universe of discourse for the forecasting problem. The universe of discourse is obtained from the range of values observed in the time series. In our study we use a grid scheme (GridPartitioner) with 35 partitions, separated into 5 subgroups, VL— Very Low, L— Low, M — Medium, H— High and VH— Very High, each subgroup with 7 levels. We applied the high order monovariate(one-variable) methods with and without weights and all of them were tested with the order of 1 to 3. The order parameter is how much past information is needed to describe future events. The methods are implemented using python version 3.7 and the models package are in the “pyFTS” library. Training Data from March 2nd to June 2nd and testing data from June 3rd to July 20th.

4.2.1 Forecasting using stationary fuzzy time series

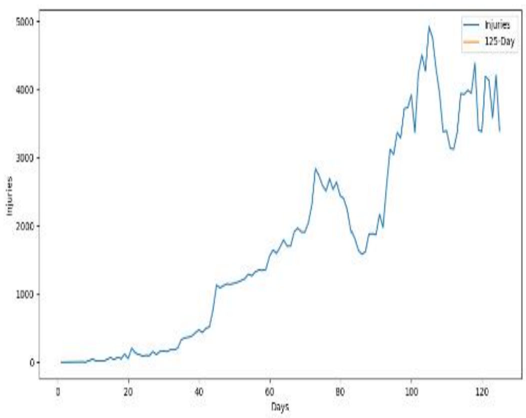
We applied High Order FTS (HOFTS), High Order Weighted FTS (WHOFTS) and Probabilistic Weighted FTS (PWFTS) methods. Table 3 shows the results of forecasting in terms of accuracy. Fig. 5 (b) shows the forecasting in red and orange colors using the applied methods.

4.2.2 Forecasting using non-stationary fuzzy time series

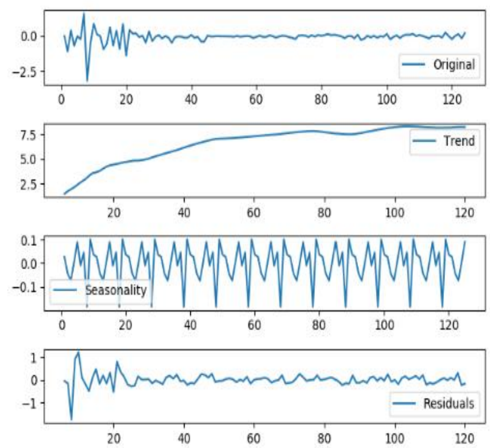
We applied the non-stationary time series forecasting models. We have provided a comparative evaluation of the performance of the non-stationary fuzzy time series (NSFTS), weighted non-stationary fuzzy time series (WNSFTS) and high order non-stationary fuzzy time series (HONSFTS) models. The input data for the methods was a time series raw-data of epidemic Injuries of Saudi Arabia. Table 4 shows the results of forecasting in terms of accuracy sorted by NRMSE. Fig. 5 (a) shows the forecasting in red and orange colors using the applied methods.

Table 1. Statistical properties for the daily detected Injuries cases in territories where COVID-19 in Saudi Arabia for 138 days from March 02, 2020, to July 20, 2020

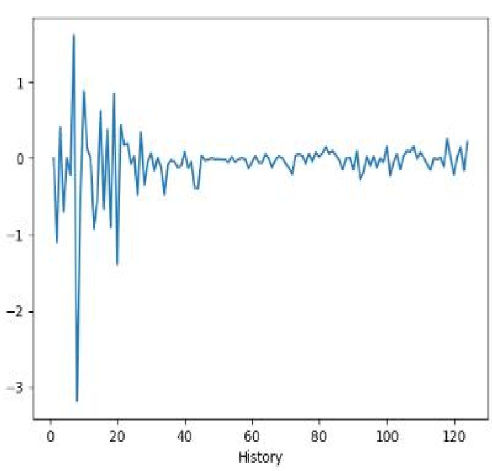
Statistic	The all , days from 1:138	Group2, days from 1:41	Group3, days from 42:83	Group4 , days from 84:125	The stationared (residual) data:
count	138.000000	41	41	41	128
mean	1835.862319	130.9512	1728.39	3379.683	0.006672
std	1430.229185	137.5545	597.9733	915.3556	0.268151
min	1.000000	1	518	1581	-1.770218
max	4919.000000	472	2840	4919	1.227400



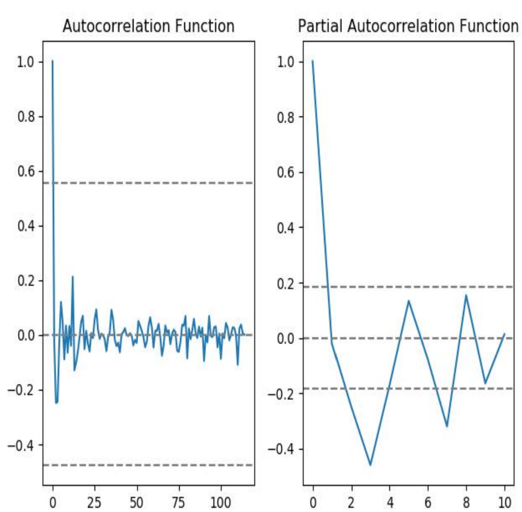
(a)



(b)



(c)



(d)

Fig. 4. (a) The raw 138 data taken from daily rates. (b) Both trend and seasonality are modeled separately and the remaining part of the series. (c) Remaining part of the series (residuals). (d) The auto correlation function ACF and PACF

Table 2. ARIMA prediction error and execution time

Methods in "statsmodels" library	ARIMAOrder	MSE	RMSE	NRMSE	Time in Seconds
AR	(12,0,0)	0.06	0.246	0.082	42.9584
AR	(2,0,0)	0.067	0.259	0.086	0.084
AR	(1,0,0)	0.071	0.267	0.089	0.11893

Method	ARIMAOrder	MSE	RMSE	NRMSE	Time in Seconds
MA	(0,0,12)	0.056	0.237	0.079	210.748
MA	(0,0,2)	0.064	0.253	0.084	0.1569
MA	(0,0,1)	0.071	0.267	0.089	0.267

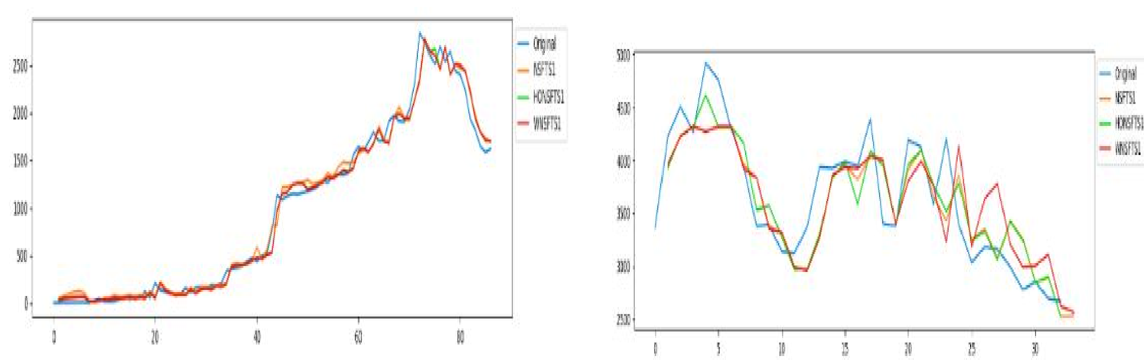
Method	ARIMAOrder	MSE	RMSE	NRMSE	Time in Seconds
ARIMA	(12,0,12)	0.0588	0.2425	0.0809	1888.45
ARIMR	(2,0,2)	0.06169	0.24837	0.08285	0.3548
ARIMA	(1,0,1)	0.0645	0.2541	0.0848	0.39677

Table 3. Prediction error and execution time

Method in"pyFTS" library	Order	MSE	RMSE	NRMSE	Time in Seconds
PWFTS3	3	0.0049	0.07	0.02335	0.2708
HOFTS3	3	0.0064	0.08	0.02669	0.1769
WHOFTS3	3	0.0064	0.08	0.02669	0.3008
PWFTS2	2	0.0324	0.18	0.06005	0.1079
WHOFTS2	2	0.0441	0.21	0.0776	0.084
HOFTS2	2	0.0484	0.22	0.07339	0.1379
WHOFTS1	1	0.0625	0.25	0.0834	0.057
PWFTS1	1	0.0576	0.24	0.0801	0.049
HOFTS1	1	0.0676	0.26	0.0867	0.042

Table 4. Prediction error and execution time

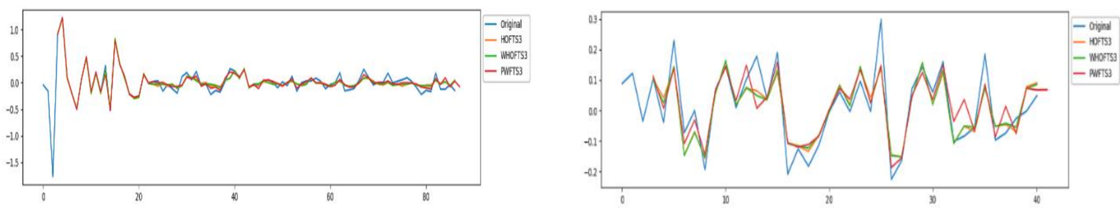
Method in "pyFTS" library	Order	MSE	RMSE	NRMSE	Time in Seconds
WNSFTS1	1	11278.4400	106.20	0.021594	0.06498
HONSFTS1	1	11493.9841	107.21	0.021800	0.103940
NSFTS1	1	13614.2224	116.68	0.023725	0.074978
HONSFTS3	3	23796.1476	154.26	0.031366	0.301846
HONSFTS2	2	136789.0225	369.85	0.075203	0.173919



training data

Test data

(a)

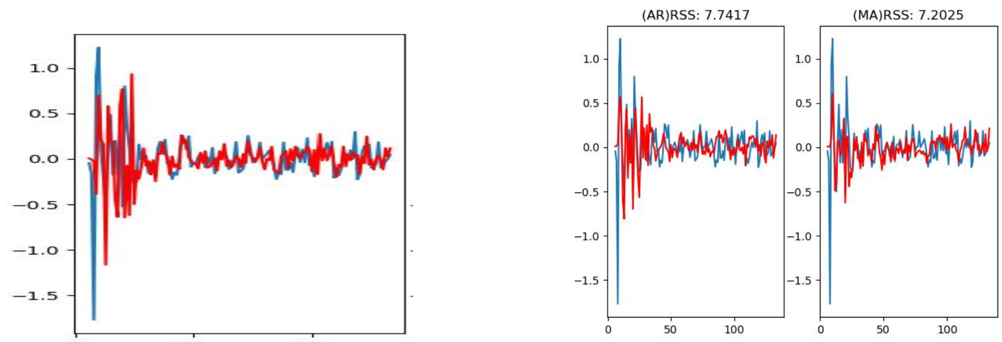


training data

Test data

(b)

ARIMA



(c)

Fig. 5. The predicted values (red and orange curves) are close to the actual ones (blue curve) (a) Forecasting using non-stationared data using FTS.(b) Forecasting using residuals (stationared data) using FTS.(c) Forecasting using residuals (stationared data) using ARIMA

5. CONCLUSION

Based on the results, this study suggests that the use of weighted Fuzzy Time Series (WFTS) in the epidemic injuries forecasting problem can provide significantly better results than the classical statistical methods. To better illustrate the prediction performance of the method, it is compared with classical time series prediction methods including autoregressive integrated moving average (ARIMA), Auto Regressive (AR) and Moving Average (MA) and other prediction fuzzy time series methods(FTS) includes, high order fuzzy time series(HOFTS) and probability weighted fuzzy time series(PWFTS). Tables 2, 3, 4 listed the forecasting errors of the comparison methods and they also present the average time in seconds. In terms of accuracy, the methods PWFTS, HOFTS and WHOFTS of order 3 have comparable results. They have the smallest MSE error. The weighted non-stationary fuzzy time series (WNSFTS) model of order 1 presented the faster testing time, while ARIMA model took a long time to implement. This can be an important improvement for forecasting problems, where usually large amount of non-stationary data needs to be processed at a reduced time. The results show that the proposed method can be used to predict the epidemic injuries. This result is of relevance in the context of epidemic injuries forecasting. Future work will be to use case-level data. Therefore, future work is to evaluate techniques that could bring improvements to the multivariable FTS methods (MVFTS). We could try different advanced algorithms like: neuro-fuzzy, weighted sum algorithms etc.

DISCLAIMER

Data used in this paper is available on (https://sehhty.com/sa-covid/).

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. Available:https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Africa.
2. Available:https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Asia#Saudi_Arabia.
3. Mao S, Xiao F. Time Series Forecasting Based on Complex Network Analysis. Digital Object Identifier 10.1109/ACCESS.2019.2906268.

4. De Gooijer JG. Hyndman R.J. 25 years of time series forecasting. International Journal of Forecasting. 2006;22(3):443–473.
5. Song, Qiang, Chissom, Brad S. Fuzzy time series and its models. Fuzzy sets and systems. 1993;54(3):269–277. Available:https://doi.org/10.1016/0165-0114(93)90372-O.Acess in 25/07/2018.
6. Available:https://towardsdatascience.com/a-short-tutorial-on-fuzzy-time-series-dcc6d4eb1b15.
7. Imo Eyoh, Robert John and Geert De Maere. Time Series Forecasting with Interval Type-2 Intuitionistic Fuzzy Logic Systems. Conference Paper; 2017. DOI:10.1109/FUZZ-IEEE.2017.8015463. Available:https://www.researchgate.net/publication/315401738
8. Preetika Saxena, Santhosh Easo. A New Method for Forecasting Enrollments based on Fuzzy Time Series with Higher Forecast Accuracy Rate. Preetika Saxena et al,Int.J.Computer Technology & Applications; 3(6):2033-2037.
9. Satpathy, Sambit, Prakash Mohan, Sanchali Das, and Swapan Debbarma. "A new healthcare diagnosis system using an IoT-based fuzzy classifier with FPGA." The Journal of Supercomputing 2019;1-13.
10. Satpathy, Sambit, M. Prakash, Swapan Debbarma, Aditya S. Sengupta, and Bidyut K. Bhattacaryya. "Design a FPGA, fuzzy based, insolent method for prediction of multi-diseases in rural area. Journal of Intelligent & Fuzzy Systems. 2019;37(5):7039-7046.
11. Pradhan MC, Satpathy S, Bhoi BK. An Intelligent Fuzzy Based Technique of Making Food Using Rice Cooker. Asian Journal of Electrical Sciences. 2016;5(1):1-7.
12. Satpathy, Sambit, Ambika Prasad Sahu. A Graphical User Interface, Fuzzy Based Intelligent Rice Cooker.In 2015 International Conference on Computational Intelligence and Communication Networks (CICN). 2015;1216-1220.
13. Severiano SA, Jr, Silva PCL, Sadaei HJ, Guimarães FG. Very Short-term Solar Forecasting using Fuzzy Time Series.IEEE International Conference on Fuzzy Systems; 2017. DOI:10.1109/FUZZ- IEEE.2017.8015732
14. A'yun K, Abadi AM, Saptaningtyas FY. Application of Weighted Fuzzy Time Series Model to Forecast Trans Jogja's Passenger. International Journal of Applied Physics and Mathematics. 2015;5(2).
15. Mahmod Othman , Yousif Alyousifi , Rajalingam Sokkalingam , Ibrahima Faye and Petronio C. L. Silva. Predicting Daily Air Pollution Index Based on Fuzzy Time Series Markov Chain Model. Available:https://www.mdpi.com/2073-8994/12/2/293
16. Sadaei HJ, Enayatifar R, Abdullah AH, Gani A. Short-term load forecasting using a hybrid model with a refined exponentially weighted fuzzy time series and an improved harmony search. International Journal of Electrical Power & Energy Systems. 2014;62:118– 129.
17. Talarposhti FM, Sadaei HJ, Enayatifar R, Guimaraes FG, Mahmud M, Eslami T. Stock market forecasting by using a hybrid model of exponential fuzzy time series. International Journal of Approximate Reasoning. 2016;70:79–98.
18. Malhorta P, Shroff G, Vig L, Agarwal P. Long Short Term Memory Networks for Anomaly Detection in Time Series. in 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2015, At Bruges, Belgium. 2015:22–24.
19. Available:https://towardsdatascience.com/significance-of-acf-and-pacf-plots-in-time-series-analysis-2fa11a5d10a8.
20. Branimir K. Hackenberger. M for measles, M for math, M for mod data analysis in medical research: fromfoe to friend, Croat Med J. 2019;60:463. Available:https://doi.org/10.3325/cmj.2019.60.463.
21. Available:https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology .
22. Richa Tripathi, Amit Reza K, Dinesh Garg. Prediction of the disease controllability in a complex network using machine learning algorithms.Researcha Gate; 2019. Available: https://www.researchgate.net/publication/331397417.
23. Zifeng Yang, Zhiqi Zeng, Ke Wang, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. Journal of Thoracic Disease. All rights reserved. J Thorac Dis. 2020;12(3):165-174. Available:http://dx.doi.org/10.21037/jtd.2020.02.64.
24. Yi-Cheng Chen, Ping-En Lu, Graduate Student Member, IEEE, Cheng-Shang

Chang, Fellow, IEEE, and Tzu-Hsuan Liu. A Time-dependent SIR model for COVID-19 with Undetectable Infected Persons. Available:arXiv:2003.00122v6 [q-bio.PE] 28 Apr 2020.
25. Can Hou, Jiaxin Chen,Yaqing Zhou, Lei Hua, Jinxia Yuan, Shu He, Yi Guo,Sheng Zhang, Qiaowei Jia,Chenhui Zhao Jing Zhang,Guangxu Xu,Enzhi Jia. The effectiveness of quarantine of Wuhan city against the Corona Virus Disease 2019 (COVID-19): A well-mixed SEIR model analysis. wileyonlinelibrary.com/journal/jmv. J Med Virol. 2020;92:841–848.
26. Abdullah Murhaf Al-Khani, Mohamed Abdelghafour Khalifa. The SARS-CoV-2 pandemic course in Saudi Arabia: A dynamic epidemiological model; 2020. Available:https://doi.org/10.1101/2020.06.01.20119800.
27. Khoj H, Alaa F Mujallad. Epidemic Situation and Forecasting of COVID-19 in Saudi Arabia using the SIR model. Available:https://doi.org/10.1101/2020.05.05.20091520.
28. Zhang PG, Qi M. Neural network forecasting for seasonal and trend time series. European Journal of Operational Research. 2005;160:501–514.
29. Ruder S. An overview of gradient descent optimization algorithms. 2016;1–14. Available:https://pdfs.semanticscholar.org/e2dc/8810671f76927d862e63faa29c401bdec5da.pdf
30. Rob J Hyndman, George Athanasopoulos. Forecasting: Principles and Practice. Monash University, Australia. Available:https://otexts.com/fpp2/AR.htmlhttps://otexts.com/fpp2/AR.html
31. Zadeh LA. Fuzzy sets. Information and Control. 1965;8(3):338–353. Available:https://doi.org/10.1016/S0019-9958(65)90241-X. Accessed in 25/07/2018.
32. Song, Qiang, Chissom, Brad S. Fuzzy time series and its models. Fuzzy sets and systems1993;54(3):269–277. Available:https://doi.org/10.1016/0165-0114(93)90372-O. Access in 25/07/2018.
33. Shyi-Ming C. Forecasting enrollments based on fuzzy time series. Fuzzy sets and systems. 1996;81(3):311–319. Access in 25/07/2018. Available:https://doi.org/10.1016/0165–0114(95)00220–0.
34. Available:https://sehhty.com/sa-covid.
35. Box G, Jenkins G, Reinsel G. Time Series Analysis: Forecasting and Control. Fourth ed. Wiley; 2008.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/60134