Journal of Industrial and Systems Engineering
Vol. 13, Special issue: 17th International Industrial Engineering Conference
Summer (July) 2021, pp. 37-54

(IIEC 2021)
MASHHAD, IRAN

Integrating PSO-GA with ANFIS for predictive analytics of confirmed cases of COVID-19 in Iran

Amir Eshaghi Chaleshtori[1], Abdollah Aghaie[1*]

[1]*Department of Industrial Engineering, K.N Toosi University of Technology, Tehran, Iran*

amir.isaqi@yahoo.com, aaghaie@kntu.ac.ir

Abstract

The first case of the unknown coronavirus, referred to as COVID-19, was detected in Wuhan, China, in late December 2019, and spread throughout China and globally. The total confirmed cases globally are rising day by day. This study proposes a novel prediction model to estimate and predict the total confirmed cases of COVID-19 in the next two days, according to Iran's confirmed cases reported before. The proposed model is an improved adaptive neuro-fuzzy inference system (ANFIS) using a co-evolutionary PSO-GA algorithm. PSO-GA is generally used to strike a balance between exploration and exploitation capabilities enhanced further by integrating the genetic operators, i.e., mutation and crossover in the PSO algorithm. The proposed model (i.e., PSO-GA-ANFIS) thus aims to enhance the efficiency of the ANFIS model by determining ANFIS parameters using PSO-GA. The model is assessed by utilizing epidemiological data provided by John Hopkins University to forecast the COVID-19 epidemic prevalence trend of Iran in 02.20.2020-06.10.2020-time window. A comparison was also made between the proposed model and a couple of available models. The results indicated that the proposed model outperforms the other models regarding MSE, RMSE, MAPE, and R^2.

Keywords: ANFIS, PSO-GA, COVID-19, prediction model, time series

1-Introduction

As a big family of viruses, coronavirus (CoVs) are well-established pathogens of humans, known to cause hepatic, gastrointestinal, neurological, and respiratory infections. They can be transmitted among bats, humans, livestock, mice, birds, and other wild animals (Chen et al. 2020), (Ge et al. 2013) and (Wang et al. 2006). The outbreak of SARS-CoV in 2003 and MERS-CoV in 2012 confirmed animal-to-animal and human-to-human transmission, respectively (Cauchemez et al. 2013). On December 2019, the WHO was notified of many respiratory infection cases in China that had previously gone to a Wuhan seafood market (Sohrabi et al. 2020). The new coronavirus (COVID-19) is now spreading across Wuhan. As concluded by the authors. In Lu et al. (2020) COVID-19 is likely to have its origins in a bat species, as it is more akin to the two of bat-derived coronaviruses. COVID-19 origins, however, have not been approved yet, requiring further examination. On January 30, 2020, the Centers for Disease Control and Prevention (CDC) affirmed human-to-human COVID-19 transmission.

[*]Corresponding author

As reported by CDC, COVID-19 can spread through the air, intimate personal contact, touching objects or surfaces containing viral particles, and seldom by fecal contamination. As indicated in Mahase (2020) COVID-19 is notorious for its incubation period, almost two weeks, during which it can transmit to other people. In Guan et al. (2020), a Chinese group estimated the incubation period to be three days, within the 0-24-day range, with the median age estimated to be 47.0 years. The confirmed cases have thus risen day by day. Many countries have already enacted and adopted stricter plans and policies against the hazardously rapid spread of COVID-19. Thus, predicting the confirmed cases in the days ahead is of vital importance to implement the required plans.

Zhao et al. (2020) suggested a mathematical model to predict the actual COVID-19 case number two weeks after the onset of 2020, not previously reported. They found 469 unreported cases in the first half of January 2020. They also estimated a 21-fold increase in cases following January 17, 2020. Nishiura et al. (2020) suggested a model to estimate the COVID-19 infection rate in Wuhan by employing the information of a total of 565 Japanese evacuees from Wuhan from 29 to January 31, 2020. As reported by them, the estimated infection rate was 9.5%, and the mortality rate was in the 0.3%-0.6% range. Nevertheless, the estimated infection and mortality rates cannot be valid due to the small number of Japanese evacuees from Wuhan. Tang et al. (2020) suggested a mathematical model to estimate COVID-19 transmission risk. They estimated the basic reproduction number to be 6.47. Furthermore, they forecasted total confirmed cases in seven days. They also anticipated that coronavirus would hit a peak in 14 days. In Thompson (2020) the information of 47 cases was utilized to predict sustained human-to-human COVID-19 transmission. The transmission rate was estimated to be 0.4, dropping to 0.012 if the median time from symptom onset to hospitalization is 1/2 of the tested data. In Jung et al. (2020) a model was proposed to estimate COVID-19 mortality risk. According to the results, the mortality risk rates were 5.1% and 8.4% under two distinct scenarios. Moreover, reproduction number was estimated for both scenarios to be 2.1 and 3.2, respectively. Accordingly, COVID-19 is likely to become pandemic.

Iran reported its first confirmed cases of COVID-19 on February 19, 2020. Total lab-confirmed-associated COVID-19 cases reported per week have risen daily. Failure in controlling the pandemic in Iran can considerably affect other countries in the region or beyond. Hence, the confirmed COVID-19 cases must be studied, and the COVID-19 prevalence trend must be predicted to implement effective strategies nationally. Scant research has been done so far in the Middle East region, together with some outstanding work in this respect Ghaffarzadegan and Rahmandad, (2020). Based on reviewing preliminary studies, COVID-19 challenges have been scantily dealt with to forecast the illness. Recently, numerous modeling techniques have been introduced for various countries, including China (Roosa et al. 2020), (Li et al.2020), (Hu et al. 2020), (Liu et al. 2020) and (Peng et al. 2020). Italy (Fanelli and Piazza 2020), (Grasselli et al. 2020), (Russo et al. 2020) and (Wangping et al. 2020). France (Fanelli and Piazza 2020), (Grasselli et al. 2020), (Russo et al. 2020) and (Wangping et al. 2020) and (Massonnaud et al. 2020). USA (Liu et al. 2020), (Lover and McAndrew, (2020) and (Wise et al. 2020) and South Korea (Kim, 2020) and (Mandal et al. 2020). Analytically, several studies have been carried out using time-series forecast models, including autoregressive integrated moving average (ARIMA) (Dehesh et al. 2020), (Shi and Fang, 2020) and (Benvenuto et al. 2020) and exponential smoothing (Elmousalami and Hassanien, 2020), (Wu et al. 2020) and (Zheng et al. 2020). They are conventional methods that properly forecast time-series data rapidly. They have been chosen for this purpose since they are widely used by researchers and rapidly implemented by different stakeholders. Aside from the mathematical modeling above, no forecasting models are available for this application. Unpredictable and complex problems entail using highly efficient techniques, including artificial intelligence (AI), to figure out such problems. Because of the significance of the subject, this study examined the role of an AI method, i.e., GA-PSO-ANFIS, in forecasting and analyzing the trend of Iran's confirmed COVID-19 cases.

The adaptive neuro-fuzzy inference system (ANFIS) is frequently used in time-series forecasting problems, suggesting an excellent performance in lots of available applications (Jang, 1993). It functions flexibly in detecting nonlinearity in time-series data, integrating the properties of artificial neural networks (ANNs) and fuzzy logic systems (FLSs). It has been utilized in several forecasting applications. For instance, Wei (2016) proposed a stock-price prediction model using empirical mode decomposition (EMD) and ANFIS. Chen et al. (2013) proposed Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) time-series forecast model in accordance with hybrid ANFIS-OWA (ordered weighted averaging). In Pousinho (2012), another time-series forecasting technique was proposed for price of electricity using ANFIS. Svalina et al. (2013) suggested an ANFIS-based forecast model for closing-price indices for a stock market for 5 days. Ekici and Aksoy (2011) also proposed an ANFIS-based forecast model for estimating building energy consumption. ANFIS is also used for electricity load forecasting (Cheng and Wei, 2010). Kumar et al. (2014) introduced an ANFIS-based model to predict product returns. Ho and Tsai (2011) used ANFIS to predict the performance of new product development (NPD). Nonetheless, ANFIS parameter estimation is a difficult task requiring to be addressed. Earlier investigations thus applied several separate swarm intelligence (SI) techniques to ANFIS parameters to improve the performance of time-series forecasting as these parameters substantially affect the performance of ANFIS. These SI techniques included PSO (Catalão et al. 2010), Social-Spider Optimization (SSO) (Bagheri et al. 2014), Sine-Cosine Algorithm (SCA) (Ewees et al. 2017), and Multi-Verse Optimizer (MVO) (Al-Qaness et al. 2018). For instance, (Ewees et al. 2017) applied SCA to increase the performance of the ANFIS model in forecasting petroleum consumption in Canada, Germany, and Japan. In Al-Qaness et al. (2018) MVO was employed to improve the performance of the ANFIS model in predicting petroleum consumption in two countries. Likewise, in El Aziz et al. (2017), PSO-ANFIS was applied to forecast the biochar yield. Nevertheless, separate SI algorithms may get stuck in local optima. A solution is, therefore, to utilize hybrid SI algorithms to overcome such a problem. In Elaziz et al. (2019) a hybrid GA-SSA (Salp Swarm Algorithm) method was proposed to increase the performance of the ANFIS model. The novel proposed model (i.e., GA-SSA-ANFIS) was used for forecasting price of crude oil for long-term time-series data.

As indicated in the literature review, novel metaheuristic algorithms have been formulated to increase the performance of the ANFIS model. Notwithstanding, methods presented hitherto are limited, so that they may influence the performance of forecasting output. For instance, slow convergence and succeeding in striking a balance between the exploitation phase and the exploration phase may affect the quality of the final output. GA and PSO are two general examples that have been employed satisfactorily to solve different engineering design optimization problems effectively. Both algorithms, however, have their own pros and cons. In GA, the data stored by an individual is lost if it is not selected, while PSO has the required memory to store the information.

Nevertheless, in the absence of a selection operator, PSO may waste its recourses on poor individuals. In contrast, GA may find it difficult to find a precise solution, showing a good performance in achieving a global region. Group interactions in PSO boost the search for optimal solutions. Thus, exploiting the strengths of both algorithms encouraged the presentation of an alternative hybridization-based forecasting method. This concept overlooks the limitations of conventional SI methods by integrating the advantages of various methods, thereby generating novel SI methods that outperform conventional methods. Hence, this paper mainly aims to formulate an efficient hybrid PSO-GA-ANFIS approach to forecast and analyze Iran's confirmed COVID-19 cases. The present study thus investigates the degree to which co-evolutionary algorithms can be applied to enhance the performance of the ANFIS model. To do so, the PSO-GA approach was proposed by exploiting both algorithms involved to solve nonlinear time-series forecasting problems. Here, PSO is applied to enhance the vector, whereas GA is applied to modify decision vectors by employing genetic operators. Notably, past studies have failed to use this kind of analysis. Reportedly, no research has been conducted to forecast and analyze the COVID-19 prevalence trend by utilizing a combination of ANFIS and coevolutionary algorithms, including PSO-GA. Therefore, to forecast and analyze the trend of Iran's confirmed COVID-19 cases, the PSO-GA algorithm was proposed to define ANFIS parameters, referred to as PSO-GA-ANFIS.
The primary contributions of this study include:

1. An effective forecast model has been proposed to forecast Iran's confirmed COVID-19 cases over the next 2 days in accordance with earlier confirmed cases.

2. An improved ANFIS model has been proposed combined with a coevolutionary algorithm, i.e., PSO-GA-ANFIS.

3. A comparison has been made between the proposed model and the main ANFIS and available modified ANFIS model, PSO-ANFIS and GA-ANFIS.

The remainder of this paper is organized as follows. In section 2, ANFIS-based theories and general optimization problem formulation are introduced. In section 3, the proposed approach used in this study is presented. In section 4, the sources of data and the experimental setup is described. In section 5, the experimental results are analyzed and discussed. In section 6, the paper is concluded, and future research directions are suggested.

2- Material and methods

2-1- Adaptive neuro-fuzzy inference system (ANFIS)

This section introduces ANFIS principles. In an ANFIS model, fuzzy logic is connected to a neural network (Jang, 1993). Here, a mapping is generated between input and output by implementing 'if-then' rules (aka Takagi-Sugeno fuzzy inference model). In Fig. 1, an ANFIS model is demonstrated, where y and x denote inputs to the first layer, while o_{1i} is the output of node i, which can be calculated as follows:

The principle of the ANFIS are given in this section. The ANFIS model links the fuzzy logic and neural network (Jang et al. 1993). It generates a mapping between the input and output by applying IF-THEN rules (it is also called Takagi-Sugeno inference model). Figure (1) illustrate the ANFIS model where, y and x define the inputs to layer 1 whereas,o_{1i} is its output of node i that is computed as follows:

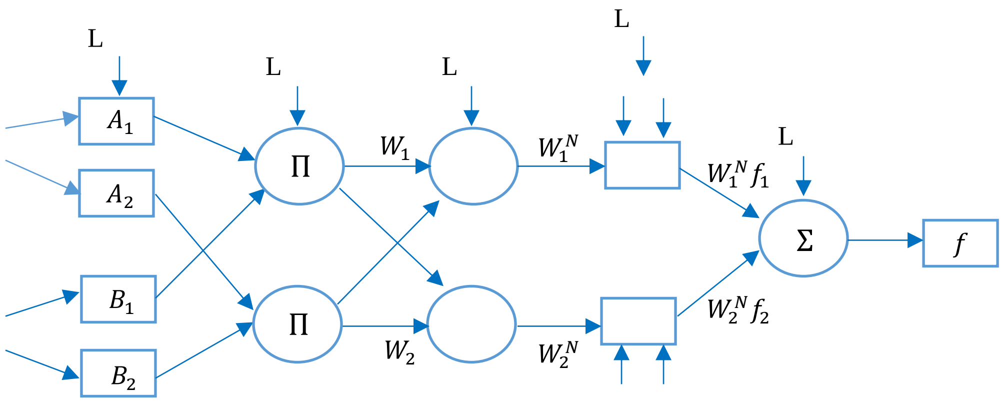

Fig 1. ANFIS model structure (Al-qaness et al. 2020)

$$o_{1i} = \mu_{A_i}(x)\ , i = 1,2\ \ , o_{1i} = \mu_{B_{i-2}}(y)\ , i = 3,4 \quad (1)$$

$$\mu(x) = e^{-(\frac{x-\varphi_i}{\alpha_i})^2} \quad (2)$$

where μ denotes the generalized Gaussian membership functions. A_i and B_i define the membership values of μ. α_i and φ_i denote the premise parameters set.

The output of the second layer (aka the firing strength of a rule) is calculated as follows:

$$o_{2i} = \mu_{A_i}(x) * \mu_{B_{i-2}}(y) \quad (3)$$

In the meantime, the output of the third layer (aka the normalized firing strength) is calculated as follows:

$$o_{3i} = \overline{w_i} = \frac{w_i}{\sum_{i=1}^{2} w_i'} \quad (4)$$

The output of the fourth layer (aka adaptive node) is measured as follows:

$$o_{4i} = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i) \tag{5}$$

In the equation above, r_i, q_i, and p_i indicate the following parameters of Node i. The fifth layer contains a single node whose output is calculated as follows:

$$o_5 = \sum_i \overline{w_i} f_i \tag{6}$$

2.2. Hybrid PSO-GA approach

In this subsection, a novel hybrid PSO-GA approach is described to increase the performance of the ANFIS model.

2-2-1- Genetic algorithm

GA is an adaptive heuristic search algorithm proposed in evolutionary themes in natural selection. It has been designed principally to model a natural system's processes needed for evolution, especially those that conform to Charles Darwin's principles to identify the survival of the fittest (Goldberg, 1989). GA represents an intelligent expansion of a random search in a predetermined search space to solve a certain problem. GA was first developed in the 1960s by John Holland. It has been extensively examined, tested, and used in many engineering fields. GAs was proposed as a computational analogy of adaptive systems. They are generally modeled in accordance with evolutionary principles by natural selection, using a population selected in the presence of variability-inducing operators, including crossover and mutation. A fitness function is employed to assess individuals and reproductive success changes with fitness.

2-2-2- Particle swarm optimization

PSO was developed by Eberhart and Kennedy (1999) and Kennedy and Eberhart (1995). It is a population-based stochastic optimization (SO) strategy that is inspired by the social behavior of flocks of birds, schools of fish, bee swarms, and even at times, the social behavior of humans. PSO is identical to GA in population initialization with random solutions and finding global optima in consecutive generations. However, mutation and crossover are not applied to PSO, while particles move across the problem space following the current optimum particles. The basic idea is that the velocity of each particle (aka the potential solution) at any instant of time varies between its personal-best position and global-best position. Arithmetically, a particle swarm is randomly initialized in the search space and move across a D-dimensional space to find new solutions. Let x_k^i be the position of i^{th} particle and v_k^i its velocity over the search space at k^{th} iteration. Then, the position and velocity of this particle at $(k+1)^{th}$ iteration are updated using the following equations:

$$v_{k+1}^i = w v_k^i + c_1 r_1 (p_k^i - x_k^i) + c_2 r_2 (p_k^g - x_k^i) \tag{7}$$

$$x_{k+}^i = x_k^i + v_{k+1}^i \tag{8}$$

where r_1 and r_2 represent random numbers between 0 and 1, c_1 and c_2 are constant, p_k^i represents the best ever position of i^{th} particle, and p_k^g corresponds to the global best position in swarm up to k^{th} iteration.

2-2-3- Hybrid PSO-GA approach

The PSO-GA approach was developed aimed at integrating GA and PSO advantages. By incorporating genetic operators into the standard PSO, the balance between exploitation and exploration capabilities is further enhanced. Nonetheless, both models have their own advantages and disadvantages. In GA, the data stored by an individual is lost if it is not selected, while PSO has the required memory to store the information. Nevertheless, in the absence of a selection operator, PSO may waste its recourses on poor individuals. Thus, the underlying concept of PSO-GA is to integrate the local search capability of GA with the social thinking capability of PSO. Since both PSO and GA are population-based, the PSO-GA approach is population-based as well, thereby proceeding to search for a global solution. The PSO-GA approach starts with initialization, during which swarm particles, along with their respective velocities, are generated randomly in the search space. The initial population x_0^i of i^{th} particles is taken $x_0^i \sim U(x_{min}, x_{max})$ randomly from a uniform distribution within the range of$[x_{min}, x_{max}]$. Here, x_{min} and x_{max} denote lower and upper bounds of the decision variables, respectively. Nonetheless, the particle's current position in the swarm is updated using the velocity vector in regard to the "memory" obtained by each particle and the knowledge acquired by the entire swarm. Thus, the particle's position in the swarm is determined according to its own experience ($pbest$) and its neighbors' (gbest). Following each iteration, the particle's position is updated based on equation (8). After a new generation is formed in PSO iterations, a certain number of particles are selected among the new population on which GA is applied individually. The particle population size is huge. Thus, for the sake of saving time, GA is not applied to the entire population. Here arises the question as to how many swarms need to be evolved in a PSO renewal generation? To answer this question, among all swarm particles, the number determined by GA_{Num} in Eq. (9) are taken, which is evolved in each PSO generation [58].

$$GA_{Num} = GA_{NumMax} + (\frac{PSO_i}{PSO_{MaxIter}})^{\gamma} * (GA_{NumMax} - GA_{NumMin}) \tag{9}$$

In the equation above, PSO_i represents the current PSO's iteration and $PSO_{MaxIter}$ represents the highest number of generations in PSO. First, the best individual is selected out of the population. Then, a new population is generated by replacing points in the existing population with better points through genetic principles. That is, after being selected and mutation/crossover operators are applied to them, a one-point crossover operator is used to recombine two parents using roulette wheel selection. Afterward, a kind of elitism is conducted to preserve the optimum solutions in the population using equation (10):

$$x_{i+1} = \begin{cases} y_i & if\ f(y_i) < f(x_i) \\ x_i & otherwise\ i = 1,2,\dots, GA_{PS} \end{cases} \tag{10}$$

Upon evaluating the new population, its size, and the highest number of iterations for GA changes regarding PSO's iteration and their relationship is defined as follow:

$$GA_{PS} = GA_{MinPS} + (\frac{PSO_i}{PSO_{MaxIter}})^{\gamma} * (GA_{MaxPS} - GA_{NumMin}) \tag{11}$$

$$GA_{MaxIter} = GA_{MinIter} + (\frac{PSO_i}{PSO_{MaxIter}})^{\beta} * (GA_{MaxIter} - GA_{MinIter}) \tag{12}$$

The population is guided to a global optimum via a repeated reproduction process. Figure 2 shows a full representation of the proposed method.

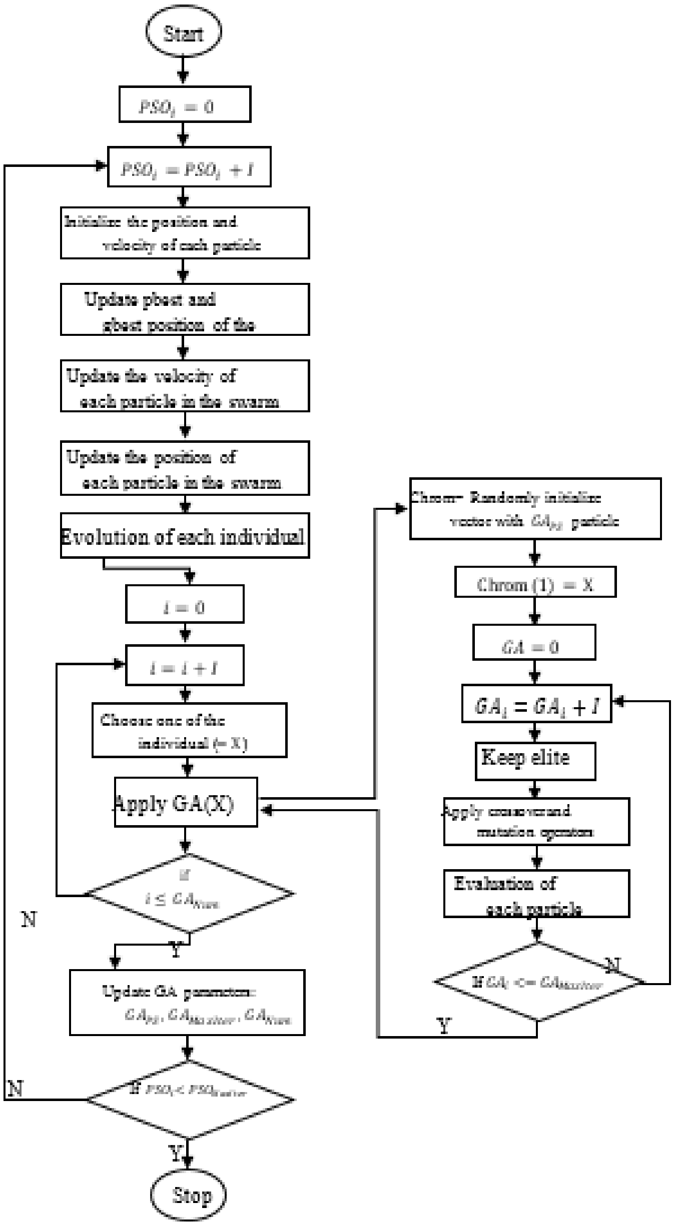

Fig 2. Hybrid PSO-GA algorithm (Garg, 2016)

3- Theory and calculation

In this section, the proposed PSO-GA-ANFIS method is explained. This is a time-series approach to predicting Iran's confirmed COVID-19 cases, as illustrated in figure 3.

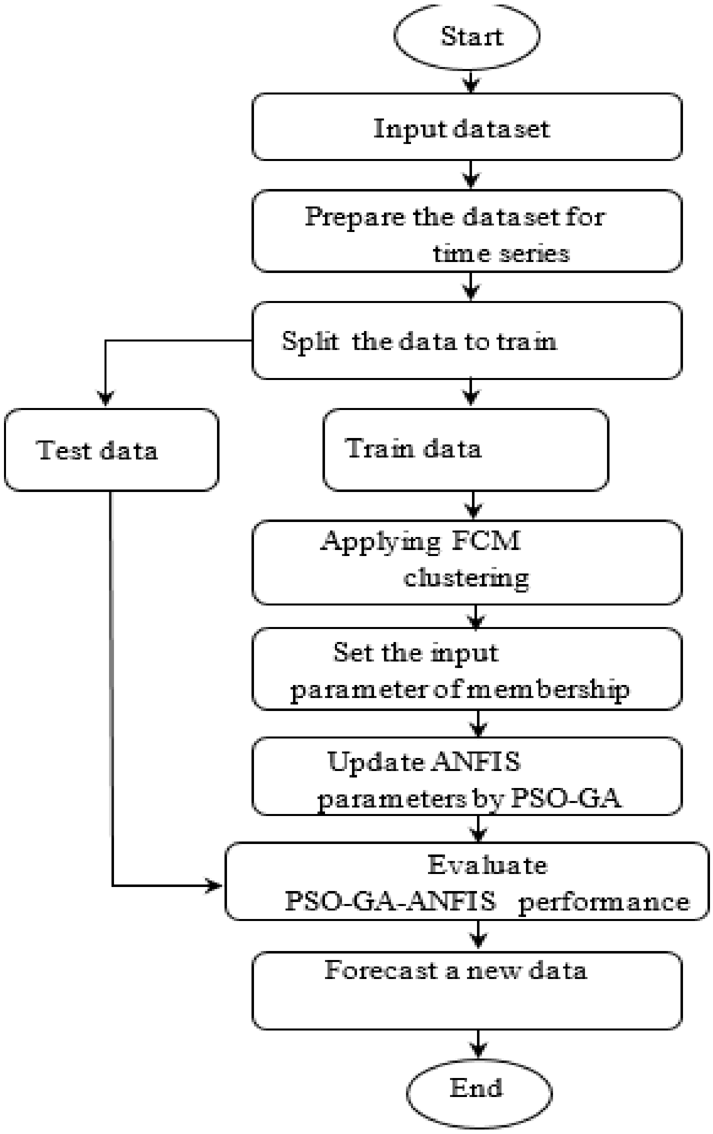

Fig 3. The proposed PSO-GA-ANFIS method (Al-qaness et al. 2020)

Using PSO-GA, PSO-GA-ANFIS train the ANFIS model through the optimization of its parameters. PSO-GA-ANFIS consists of 5 layers identical to the traditional ANFIS model. The first layer includes input variables, i.e., historical confirmed COVID-19 cases. The fifth layer, however, generates the forecasted values. PSO-GA-ANFIS begins with formatting input data in a time-series format. We used autocorrelation function (ACF) for our case, which can be used to identify patterns in data. It provides information on the correlation between points that are separated by different time lags. The number of autoregressive (AR) terms usually employed for real-world applications has not been stated for nonlinear time-series problems (Zhang and Hu, 1998). Thus, the present study tested with a fairly greater number (2) for the order of AR terms.

Each dataset is split into a training set (55%) and a test set (45%). The initial number of clusters is determined by fuzzy C-means (FCM) to develop the ANFIS model. The parameters in the ANFIS model are adjusted by PSO-GA. In the training stage, the calculation error, as in equation (13), between the actual and predicted data, is utilized to assess the quality of parameters.

$$RMSE = \sqrt{\frac{1}{N_s}\sum_{i=1}^{N_s}(T_i - P_i)^2} \tag{13}$$

In the equation above, T is real data, P is predicted data, and N_s is the sample length. The smaller values of the objective function (OF) suggest excellent ANFIS parameters. PSO-GA-ANFIS typically begins with population (X) generation, followed by the calculation of the OF for each solution. The solution with minimum error value is stored in the subsequent iteration. This sequence is repeated until the stopping condition has been met (i.e., the highest number of iterations in this paper). Afterward, the optimum solution is passed in order to train the parameters of the ANFIS model. Upon completing the training stage, the testing stage begins with the optimum solution to calculate the final output. The performance of the proposed method is assessed by making a comparison between real and predicted data using performance measures.

4- Experiments

In this section, the utilized dataset, parameter setting for all methods, performance measures, experimental results, and discussion are described.

4-1- Dataset description

In the present study, the original dataset consists of Iran's confirmed COVID-19 cases. The required data was gathered from the epidemiological data repository of John Hopkins University[1] to forecast Iran's COVID-19 epidemic prevalence trend in the time window 02.20.2020 to 06.10.2020. 55% of the dataset was utilized for training the model, and the remaining 45% to test the model.

4-2- Performance measures

Using a group of performance measures, the quality of the proposed method was assessed as follows:

- Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N_s}\sum_{i=1}^{N_s}(T_i - P_i)^2} \tag{14}$$

Where T is the real data, and P is the predicted data, N_s is the sample length.

- Mean Square Error (MSE)

It is square form of RMSE mentioned in equation (14)

- Coefficient of determination (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{N_s}(Y_i - YP_i)^2}{\sum_{i=1}^{N_s}(Y_i - \overline{Y}_i)^2} \tag{15}$$

Where Y is the real data, and YP is the predicted data, N_s is the sample length, and $\overline{Y}$ represent the average of Y. Smaller values of $RMSE$ and MSE indicate the superiority of the method and larger values of R^2 indicate better correlation for the method.

[1] https://github.com/CSSEGISandData/COVID-19

- Mean absolute percentage error (MAPE)

$$MAPE = 100 * \frac{\sum_{i=1}^{N_s} \left|\frac{Y_i - YP_i}{Y_i}\right|}{N} \tag{16}$$

Smaller values of and $MAPE$ indicate the superiority of the method.

4-3- Parameter settings

The purpose of this paper is to evaluate the capability of PSO-GA-ANFIS to predict COVID-19 by comparing it with other techniques, such as the ANFIS model and the trained ANFIS model, in terms of its performance, using PSO and GA individually. Table 1 lists the parameters set for this model. Conventional parameters (e.g., population size) have been set at 25-100 iterations are applied. Each algorithm is then implemented for 30 independent runs for a fair comparison. Certain parameters are selected since they lead to desirable behaviors in the earlier experiments, such as (Catalão et al. 2010), (Al-Qaness et al. 2018), (Ahmed et al. 2016), (Alameer et al. 2019) and (Garg, 2016).

Table 1. Parameters' setting

Algorithm	Parameters setting
ANFIS	Max epochs=100, Error goal=0, Initial step=0.01 Decrease rate=0.9, Increase rate=1.1
GA-ANFIS	Crossover type=1, Crossover rate=0.4, Mutation rate=0.1
PSO-ANFIS	wMax=0.9, wMin=0.2, C1=1, C2=2, mp=0.01
PSO-GA-ANFIS	$GA_{MinPS} = 10, GA_{MinIter} = 10, GA_{NumMax} = 20$ $,\gamma = 1,\beta = 15 GA_{NumMin} = 1$

5- Results and discussion

Table 2 depicts the results of the comparison between the proposed model (i.e., PSO-GA-ANFIS) and others to predict COVID-19. It may be inferred that the proposed model had a better performance than other models. For instance, by examining E , MSE and $MAPE$ results, PSO-GA-ANFIS notably obtains the minimum value among other comparison algorithms, indicating excellent quality of PSO-GA-ANFIS. Additionally, R^2 value indicates a strong correlation between the prediction by the proposed model (i.e., PSO-GA-ANFIS) and the original COVID-19, which is close to 1. This is also demonstrated in figures (4-11), illustrating the performance of algorithms using the historical COVID-19 data along with their prediction performance for 2 days. In light of the above, it may be inferred that the proposed model (i.e., PSO-GA-ANFIS) can significantly predict the COVID-19 dataset. The results above ignore the limitations of the conventional ANFIS model since it is combined with PSO-GA.

Table 2. Performance metric for each algorithm

Performance Metrics	$ANFIS$		$GA-ANFIS$		$PSO-ANFIS$		$PSO-GA-ANFIS$	
	$train$	$test$	$train$	$test$	$train$	$test$	$train$	$test$
MSE	$3.7412*10^{-6}$	0.02214	0.032564	0.0071505	0.023983	0.0057813	0.0026981	0.0055163
$RMSE$	0.0019342	0.14882	0.18045	0.084561	0.15486	0.0766035	0.051943	0.074272
$MAPE$	0.2754	0.2789	0.0983	0.01583	0.0675	0.0842	0.0483	0.0423
R^2	0.999	0.27053	0.911	0.79529	0.90849	0.83708	0.97468	0.82698

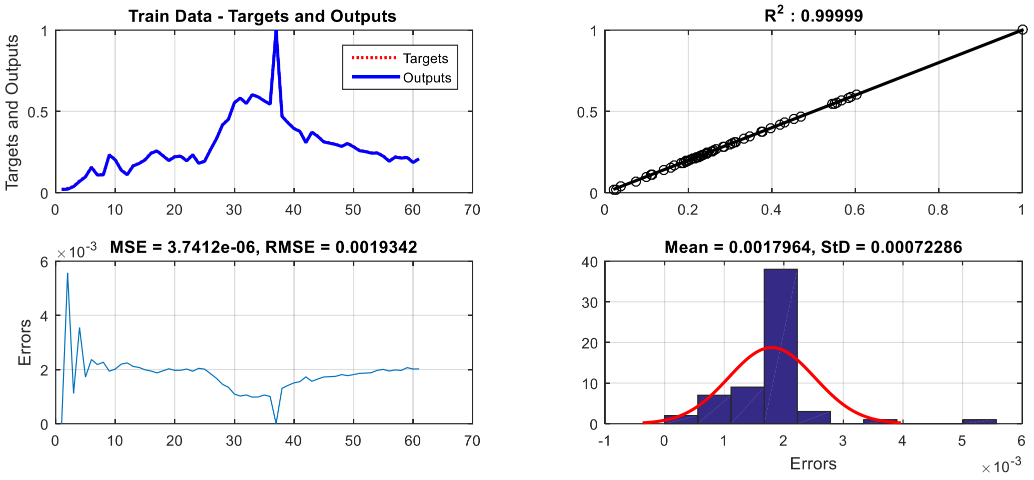

Fig 4. Comparison of actual and predicted value by ANFIS for train data

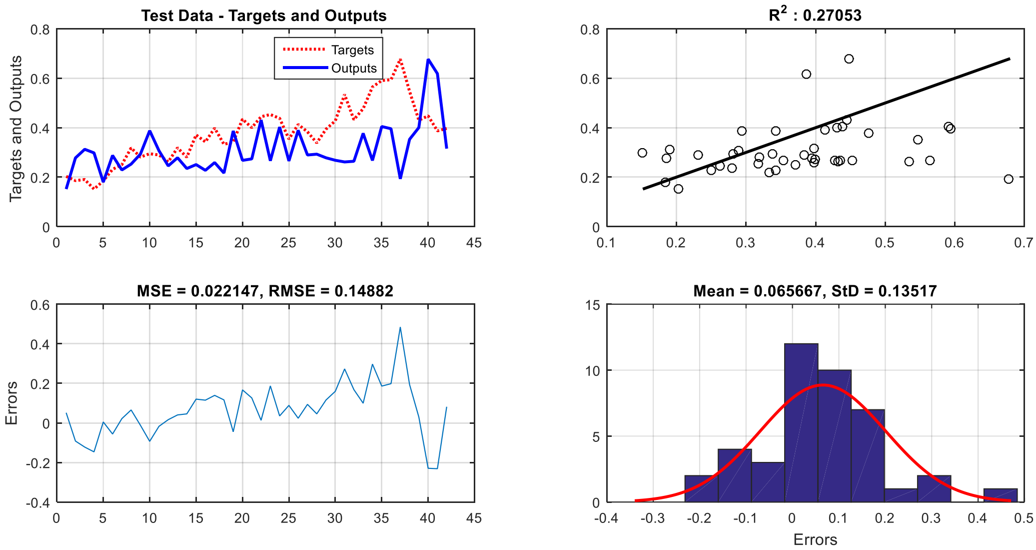

Fig 5. Comparison of actual and predicted value by ANFIS for test data

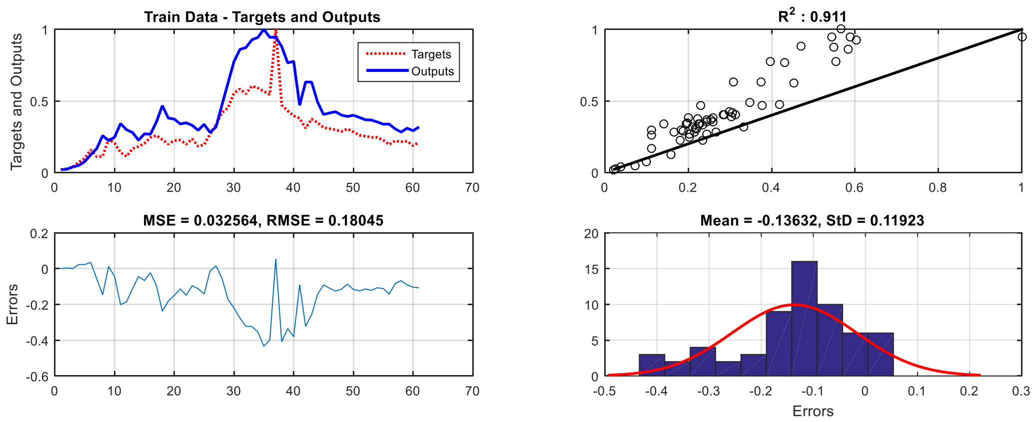

Fig 6. Comparison of actual and predicted value by GA- ANFIS for train data

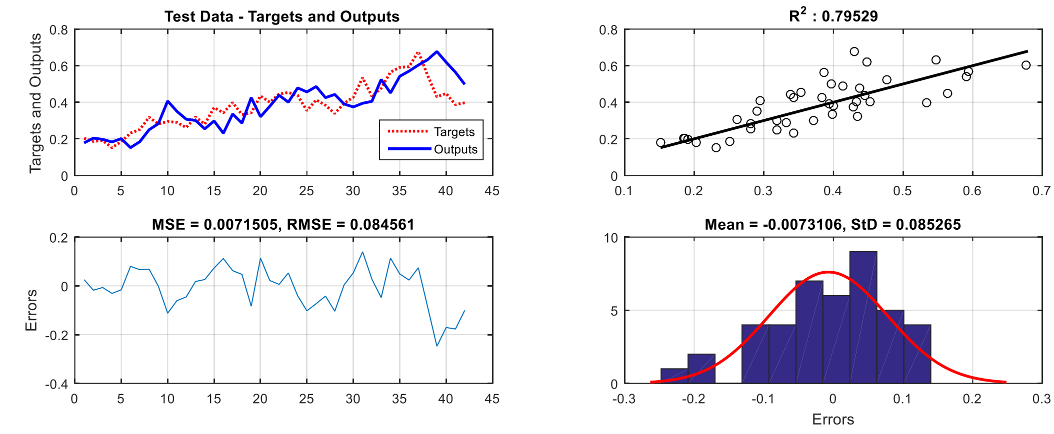

Fig 7. Comparison of actual and predicted value by GA- ANFIS for test data

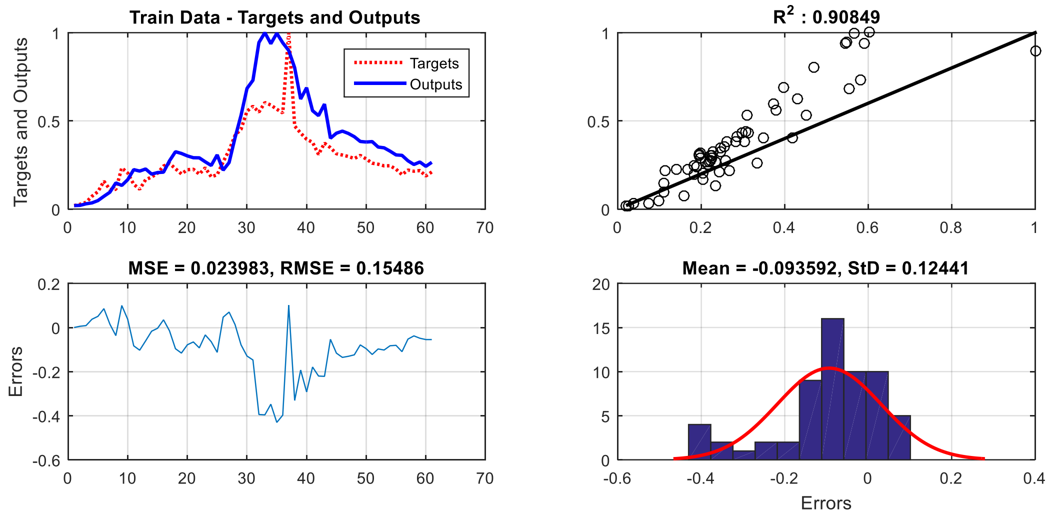

Fig 8. Comparison of actual and predicted value by PSO- ANFIS for train data

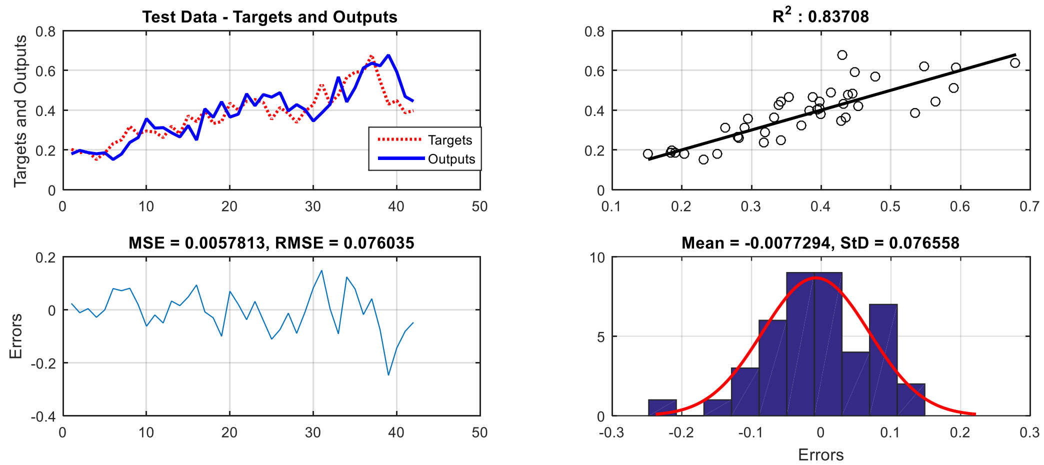

Fig 9. Comparison of actual and predicted value by PSO- ANFIS for test data

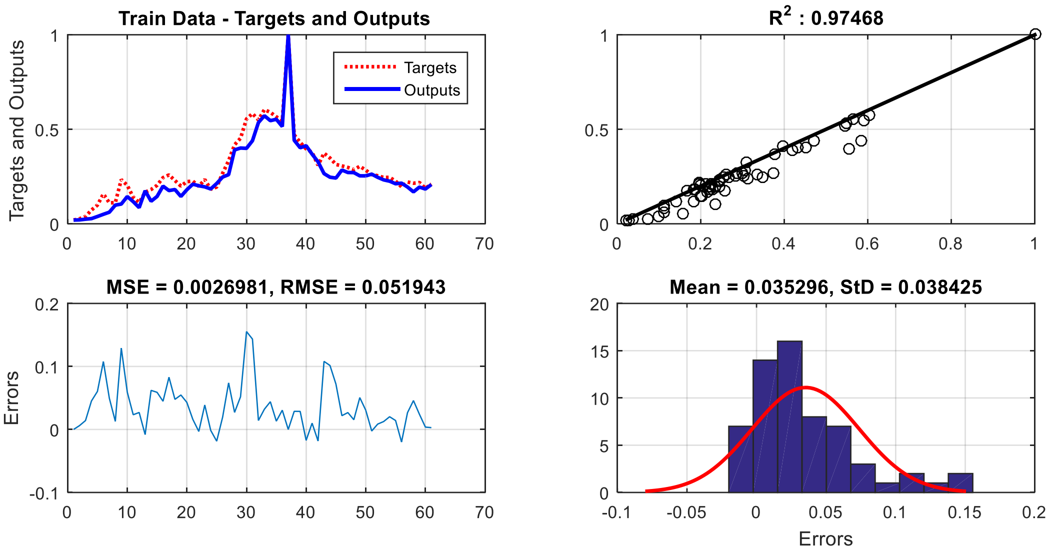

Fig 10. Comparison of actual and predicted value by PSO-GA- ANFIS for train data

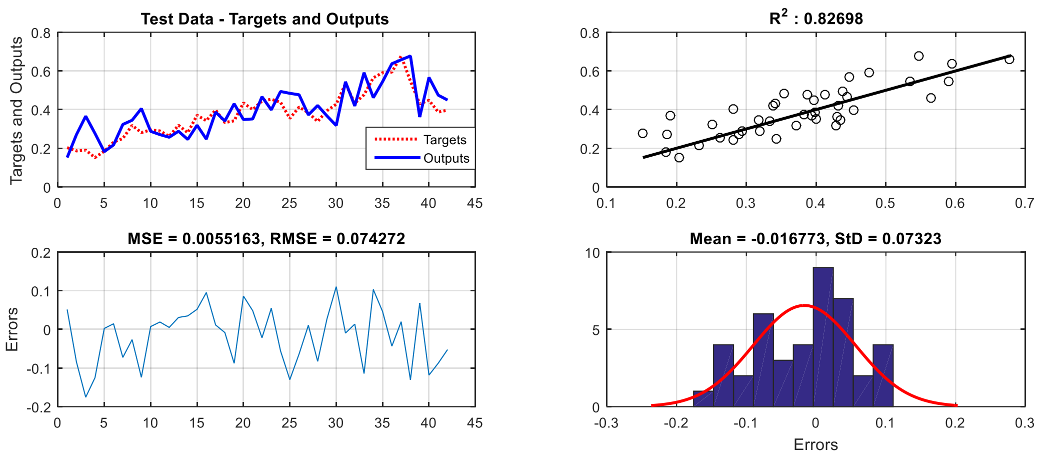

Fig 11. Comparison of actual and predicted value by PSO-GA- ANFIS for test data

6- Conclusions

In this paper, a co-evolutionary algorithm (i.e., PSO-GA) was proposed to enhance the performance of the ANFIS model by defining optimal values for its parameters. The PSO-GA-ANFIS model was used to predict COVID-19 (a previously unknown coronavirus) detected in Wuhan, Hubei Province, China. The proposed model (i.e., PSO-GA-ANFIS) could substantially forecast total confirmed cases in 2 days. Likewise, PSO-GA-ANFIS showed a better performance compared to other forecast models regarding $MSE, RMSE$, and R^2. Iran's confirmed COVID-19 cases dataset was employed to evaluate the performance of the proposed method, suggesting its excellent performance. Since PSO-GA-ANFIS yielded promising results, it can be used for various applications. This research is still in its infancy as there are scant historical data to guarantee the accuracy of the model. COVID-19 broke out in Iran nearly three months ago. Thus, the performance of the forecasting model can surely be enhanced by employing further epidemiological data from the outbreak of COVID-19 in the existing models.

References

Ahmed, K., Ewees, A. A., El Aziz, M. A., Hassanien, A. E., Gaber, T., Tsai, P. W., & Pan, J. S. (2016, October). A hybrid krill-ANFIS model for wind speed forecasting. In International Conference on Advanced Intelligent Systems and Informatics (pp. 365-372). Springer, Cham.

Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H., & Jianhua, Z. (2019). Forecasting copper prices using hybrid adaptive neuro-fuzzy inference system and genetic algorithms. *Natural Resources Research*, 28(4), 1385-1401.

Al-qaness, M. A., Ewees, A. A., Fan, H., & Abd El Aziz, M. (2020). Optimization method for forecasting confirmed cases of COVID-19 in China. *Journal of Clinical Medicine*, 9(3), 674.

Al-Qaness, M. A., Elaziz, M. A., & Ewees, A. A. (2018). Oil consumption forecasting using optimized adaptive neuro-fuzzy inference system based on sine cosine algorithm. *IEEE Access*, 6, 68394-68402.

Bagheri, A., Peyhani, H. M., & Akbari, M. (2014). Financial forecasting using ANFIS networks with quantum-behaved particle swarm optimization. *Expert Systems with Applications*, 41(14), 6235-6250.

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in brief, 105340.

Catalão, J. P. D. S., Pousinho, H. M. I., & Mendes, V. M. F. (2010). Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting. *IEEE Transactions on Power Systems*, 26(1), 137-144.

Cauchemez, S., Van Kerkhove, M. D., Riley, S., Donnelly, C. A., Fraser, C., & Ferguson, N. M. (2013). Transmission scenarios for Middle East Respiratory Syndrome Coronavirus (MERS-CoV) and how to tell them apart. Euro surveillance: bulletin European sur les maladies transmissible= European communicable disease bulletin, 18(24).

Chen, Y., Liu, Q., & Guo, D. (2020). Emerging coronaviruses: genome structure, replication, and pathogenesis. *Journal of medical virology*, 92(4), 418-423.

Cheng, C. H., Wei, L. Y., Liu, J. W., & Chen, T. L. (2013). OWA-based ANFIS model for TAIEX forecasting. *Economic Modelling*, 30, 442-448.

Cheng, C. H., & Wei, L. Y. (2010). One step-ahead ANFIS time series model for forecasting electricity loads. *Optimization and Engineering*, 11(2), 303-317.

Dehesh, T., Mardani-Fard, H. A., & Dehesh, P. (2020). Forecasting of covid-19 confirmed cases in different countries with arima models. medRxiv.

Eberhart, R., & Kennedy, J. (1999). A new optimizer using particle swarm optimization. In Proceedings of the 1995 Sixth International Symposium on Micro Machine and Human Science (pp. 39-43).

Ekici, B. B., & Aksoy, U. T. (2011). Prediction of building energy needs in early stage of design by using ANFIS. *Expert Systems with Applications*, 38(5), 5352-5358.

Elmousalami, H. H., & Hassanien, A. E. (2020). Day level forecasting for Coronavirus Disease (COVID-19) spread: analysis, modeling and recommendations. arXiv preprint arXiv:2003.07778.

Elaziz, M. A., Ewees, A. A., & Alameer, Z. (2019). Improving adaptive neuro-fuzzy inference system based on a modified salp swarm algorithm using genetic algorithm to forecast crude oil price. *Natural Resources Research*, 1-16.

El Aziz, M. A., Hemdan, A. M., Ewees, A. A., Elhoseny, M., Shehab, A., Hassanien, A. E., & Xiong, S. (2017, June). Prediction of biochar yield using adaptive neuro-fuzzy inference system with particle swarm optimization. In 2017 IEEE PES Power Africa (pp. 115-120). IEEE.

Ewees, A. A., El Aziz, M. A., & Elhoseny, M. (2017, July). Social-spider optimization algorithm for improving ANFIS to predict biochar yield. In 2017 8th international conference on computing, communication and networking technologies (ICCCNT) (pp. 1-6). IEEE.

Fanelli, D., & Piazza, F. (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons & Fractals*, 134, 109761.

Garg, H. (2016). A hybrid PSO-GA algorithm for constrained optimization problems. *Applied Mathematics and Computation*, 274, 292-305.

Ge, X. Y., Li, J. L., Yang, X. L., Chmura, A. A., Zhu, G., Epstein, J. H., ... & Zhang, Y. J. (2013). Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature*, 503(7477), 535-538.

Ghaffarzadegan, N., & Rahmandad, H. (2020). Simulation-based Estimation of the Spread of COVID-19 in Iran. medRxiv.

Goldberg D, E. (1989). Genetic Algorithm in Search, Optimization and Machine Learning. Reading: Addison2wes2 ley.

Grasselli, G., Pesenti, A., & Cecconi, M. (2020). Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response. *Jama*.

Guan, W. J., Ni, Z. Y., Hu, Y., Liang, W. H., Ou, C. Q., He, J. X., ... & Du, B. (2020). Clinical characteristics of 2019 novel coronavirus infection in China. MedRxiv.

Ho, Y. C., & Tsai, C. T. (2011). Comparing ANFIS and SEM in linear and nonlinear forecasting of new product development performance. *Expert Systems with Applications*, 38(6), 6498-6507.

Hu, Z., Ge, Q., Jin, L., & Xiong, M. (2020). Artificial intelligence forecasting of covid-19 in china. arXiv preprint arXiv:2002.07112.

Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3), 665-685.

Jung, S. M., Akhmetzhanov, A. R., Hayashi, K., Linton, N. M., Yang, Y., Yuan, B., ... & Nishiura, H. (2020). Real-time estimation of the risk of death from novel coronavirus (COVID-19) infection: inference using exported cases. *Journal of clinical medicine*, 9(2), 523.

Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In Proceedings of ICNN'95-International Conference on Neural Networks (Vol. 4, pp. 1942-1948). IEEE.

Kim, S. K. (2020). AAEDM: Theoretical Dynamic Epidemic Diffusion Model and Covid-19 Korea Pandemic Cases. medRxiv.

Kumar, D. T., Soleimani, H., & Kannan, G. (2014). Forecasting return products in an integrated forward/reverse supply chain utilizing an ANFIS. *International Journal of Applied Mathematics and Computer Science*, 24(3), 669-682.

Li, Q., Feng, W., & Quan, Y. H. (2020). Trend and forecasting of the COVID-19 outbreak in China. Journal of Infection, 80(4), 469- 496.

Liu, P., Beeler, P., & Chakrabarty, R. K. (2020). COVID-19 Progression Timeline and Effectiveness of Response-to-Spread Interventions across the United States. medRxiv.

Liu, Z., Magal, P., Seydi, O., & Webb, G. (2020). Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data. arXiv preprint arXiv:2002.12298.

Lover, A. A., & McAndrew, T. (2020). Sentinel Event Surveillance to Estimate Total SARS-CoV-2 Infections, United States. medRxiv.

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., ... & Bi, Y. (2020). Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, 395(10224), 565-574.

Mahase, E. (2020). China coronavirus: what do we know so far? BMJ 2020;368:m308 doi: 10.1136/bmj.m308

Mandal, S., Bhatnagar, T., Arinaminpathy, N., Agarwal, A., Chowdhury, A., Murhekar, M., ... & Sarkar, S. (2020). Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical model-based approach. *Indian Journal of Medical Research*, 151(2), 190.

Massonnaud, C., Roux, J., & Crépey, P. (2020). COVID-19: Forecasting short term hospital needs in France. medRxiv.

Nishiura, H., Kobayashi, T., Yang, Y., Hayashi, K., Miyama, T., Kinoshita, R., ... & Akhmetzhanov, A. R. (2020). The rate of under ascertainment of novel coronavirus (2019-nCoV) infection: estimation using Japanese passengers' data on evacuation flights.

Peng, L., Yang, W., Zhang, D., Zhuge, C., & Hong, L. (2020). Epidemic analysis of COVID-19 in China by dynamical modeling. arXiv preprint arXiv:2002.06563.

Pousinho, H. M. I., Mendes, V. M. F., & Catalão, J. P. D. S. (2012). Short-term electricity prices forecasting in a competitive market by a hybrid PSO–ANFIS approach. *International Journal of Electrical Power & Energy Systems*, 39(1), 29-35.

Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J. M., ... & Chowell, G. (2020). Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infectious Disease Modelling*, 5, 256-263.

Russo, L., Anastassopoulou, C., Tsakris, A., Bifulco, G. N., Campana, E. F., Toraldo, G., & Siettos, C. (2020). Tracing DAY-ZERO and Forecasting the Fade out of the COVID-19 Outbreak in Lombardy, Italy: A Compartmental Modelling and Numerical Optimization Approach. medRxiv.

Shi, Z., & Fang, Y. (2020). Temporal relationship between outbound traffic from Wuhan and the 2019 coronavirus disease (COVID-19) incidence in China. medRxiv.

Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., ... & Agha, R. (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International Journal of Surgery*.

Svalina, I., Galzina, V., Lujić, R., & ŠImunović, G. (2013). An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices. *Expert systems with applications*, 40(15), 6055-6063.

Tang, B., Wang, X., Li, Q., Bragazzi, N. L., Tang, S., Xiao, Y., & Wu, J. (2020). Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *Journal of clinical medicine*, 9(2), 462.

Thompson, R. N. (2020). Novel coronavirus outbreak in Wuhan, China, 2020: Intense surveillance Is vital for preventing sustained transmission in new locations. *Journal of clinical medicine*, 9(2), 498.

Wang, L. F., Shi, Z., Zhang, S., Field, H., Daszak, P., & Eaton, B. T. (2006). Review of bats and SARS. *Emerging infectious diseases*, 12(12), 1834.

Wangping, J., Ke, H., Yang, S., Wenzhe, C., Shengshu, W., Shanshan, Y., ... & Miao, L. (2020). Extended SIR prediction of the epidemics trend of COVID-19 in Italy and compared with Hunan, China. *Frontiers in Medicine*, 7, 169.

Wei, L. Y. (2016). A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting. *Applied Soft Computing*, 42, 368-376.

Wise, T., Zbozinek, T. D., Michelini, G., & Hagan, C. C. (2020). Changes in risk perception and protective behavior during the first week of the COVID-19 pandemic in the United States.

Wu, T., Ge, X., Yu, G., & Hu, E. (2020). Open-source analytics tools for studying the COVID-19 coronavirus outbreak. medRxiv.

Zhang, G., & Hu, M. Y. (1998). Neural network forecasting of the British pound/US dollar exchange rate. *Omega*, 26(4), 495-506.

Zhao, S., Musa, S. S., Lin, Q., Ran, J., Yang, G., Wang, W., ... & Wang, M. H. (2020). Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the first half of January 2020: a data-driven modelling analysis of the early outbreak. *Journal of clinical medicine*, 9(2), 388.

Zheng, Z., Wu, K., Yao, Z., Zheng, J., & Chen, J. (2020). The prediction for development of COVID-19 in global major epidemic areas through empirical trends in China by utilizing state transition matrix model. Available at SSRN 3552835.