ORIGINAL ARTICLE

Modeling the spread of dangerous pandemics with the utilization of a hybrid-statistical–Advanced-Fuzzy-Cognitive-Map algorithm: the example of COVID-19

Peter P. Groumpos[1] · Ioannis D. Apostolopoulos[2]

Abstract

Purpose The novel Coronavirus SARS-coV-2 outbreak late in 2019 and early 2020, known today as the COVID-19 pandemic, has spread fast throughout the world. It has considerably affected the lives of all people around the globe while the number of deaths related to the pandemic keeps increasing worldwide. Being able to predict the spread of the pandemic has been very helpful to governments to decide on actions. Statistical prediction models are capable of modeling a single snapshot but have several well-known weaknesses, such as linear assumptions between pandemic variables, while they cannot confirm the actual causality between studied factors. In the present work, the authors propose a state space Advanced Fuzzy Cognitive Maps (AFCM) approach model to predict the spread of the pandemic, using dynamic cause and effect relationships between pre-defined factors.
Methods State-Space Advanced Fuzzy Cognitive Maps are proposed for modeling the spread of the pandemic, utilizing several social, policy, and healthcare factors. Statistical data from Greece, South Korea, and Germany are gathered to evaluate the performance of the proposed model.
Results The proposed methodology was able to predict the pandemic trend in the studied countries, in terms of the total number of confirmed patient cases, yielding a coefficient of determination of 0.99, 0.94, and 0.97 respectively. The Pearson's correlation coefficient was found to be 0.99, 0.97, and 0.98 respectively.
Conclusion The results demonstrate the effectiveness and the advantages of the proposed methodology when modeling uncertain and dynamic situations, like novel pandemics.

Keywords COVID-19 · Fuzzy cognitive maps · Pandemic forecast · Advanced Fuzzy Cognitive Maps

Introduction

A pandemic is a disease that spreads across a large area including multiple countries, continents, or worldwide. The novel Coronavirus outbreak late in 2019 and early 2020, known today as COVID-19 or SARS–CoV-2 (which refers to the virus causing the disease), is with us. The World Health Organization (WHO), on 11 March 2020, declared COVID-19 disease a pandemic (Stang et al. 2020). It has horribly spread across the stretches of the world, and too many people are suffering. COVID-19 is a disease that people are not immune to, spreading throughout the world beyond expectation. Almost 211 countries are suffering (Xu et al. 2020). Countries all over the world are declaring Health emergencies and taking whatever possible steps to keep the citizens safe (Xu et al. 2020). By June 2021, worldwide deaths are approaching 3.85 million, and more than 178 million cases of COVID-19 are confirmed. That is why the main objective of the present investigation is to identify the significant risk factors in transmitted COVID-19.

Common symptoms include fever, cough, headache, fatigue, shortness of breath, persistent pain or pressure in the chest, bluish lips or face, and loss of smell and taste (Lu et al. 2020). The majority of patient instances show mild

✉ Ioannis D. Apostolopoulos
  ece7216@upnet.gr

  Peter P. Groumpos
  groumpos@ece.upatras.gr

[1] Department of Electrical and Computer Engineering, University of Patras, 26504 Patras, PC, Greece

[2] School of Medicine, Department of Medical Physics, University of Patras, PC 26504 Patras, Greece

symptoms most of the times; nevertheless, some progress to Acute Respiratory Distress syndrome (ARDS) possibly precipitated by multi-organ failure and blood clots (Murthy et al. 2020; Huang et al. 2020). It takes four-five days after the exposure to the virus for some symptoms to be revealed, but this time may also range from two to fourteen days (Lu et al. 2020; Murthy et al. 2020).

The virus is primarily spread between people during close contact, most often via small droplets produced by coughing, sneezing, and talking. The droplets usually fall to the ground or onto surfaces rather than traveling through air over long distances (Galbadage et al. 2020). It is most contagious during the first three days after the onset of symptoms, although the spread is possible before symptoms appear, and from people who do not show symptoms. Estimates of the number of people infected by one person with COVID-19 are referred to as the basic productive number, R0 have varied widely (Battiston and Gamba 2021). Elderly human beings with underlying diseases and affected by COVID-19 had a higher mortality rate. Younger ages and especially children are not severely affected by the disease. For some reason, males are more susceptible to the disease (Bwire 2020; Huang et al. 2020).

At a time when humanity is facing the alarming spread of the new pandemic, one feels that despite the predictions and measures, the tools that the evolution of human thought has acquired are not yet capable of arming us with both knowledge and know-how to deal with pandemics in a timely and effective manner. Statistical prediction models, although studied in great depth, are capable of modeling a single snapshot and, moreover, have the well-known weaknesses about their flexibility, their ability to handle situations involving uncertainty, and manage a multitude of interdependent factors that affect the rate of disease spread (Groumpos 2021; Ala'raj et al. 2021). Moreover, the statistical models and the correlation between variables of the pandemic are commonly based on linear assumptions. However, most, if not all, medical problems, especially new and not known pandemics, are nonlinear in their nature (Groumpos 2021).

Since the pandemic has caught the whole world by surprise, no one was ready to deal with it. Very few studies have been conducted and all research articles are either on SARS and MERS (Huang et al. 2020; De Wit et al. 2016) or on COVID-19 the last 3–4 months (Toğaçar et al. 2020; Hijikata et al. 2020; Gralinski and Menachery 2020; Wrapp et al. 2020; Li and De Clercq 2019; Ling 2020). Again, most of them use statistical methods and probability theories. New theories are needed to address the many challenging questions for COVID-19.

In our study, we analyzed and conducted research on the factors that affect the extent and capacity of the spread of the pandemic, isolating the most important ones. Our goal is to develop a decision-making system that can predict the potential development of pandemics in a given region and/or country, based not only on correlation factors, but also on causality. Besides, this system, in addition to forecasting, will be able to be for the supervision or the effectiveness of measures.

Methods

For this study, the modeling methodology provided by Fuzzy Cognitive Maps, is proposed as a solution. In this section, we provide a detailed description of the research methodology, as well as the parameters of the proposed model.

Classical theory of Fuzzy Cognitive Maps

Cognitive Maps (CM) were firstly introduced by R. Axelrod in 1976 (Axelrod 1976) as a formal way of representing social scientific knowledge and modeling decision in social and political systems. However, CM, as had been developed by Axelrod, were not sufficient to address the dynamic behaviors of complex systems, especially in the presence of uncertainty and ambiguity. They were not able to capture the true knowledge of the complex dynamic systems, especially as the time was evolving. Thus, an extension of CM was needed and was introduced ten years later 1986 by Kosko (Kosko 1986), bringing in the computational fuzzy logic. It was a new soft computing method had been named the Fuzzy Cognitive Map (FCM) approach. FCMs have become popular and used extensively with excellent results since: (1) complex dynamic systems have many parameters and they are difficult to be modeled with precise mathematical methods, (2) they are nonlinear, and time-varying, (3) depend on many real data and on the evolution of time. On the other hand, FCMs are very simple and are close to human reasoning. In addition, they do not need to have a precise mathematical model of the process. They can address any complex dynamic system without using complex mathematics. FCM methods depend heavily on accumulated human knowledge. This is due to the use of experts' opinion and their human Decision-Making Support Systems (DMSS).

Fuzzy Cognitive Maps came as a combination of methods of fuzzy logic and neural networks. They constitute a computational method that is able to examine situations during which the human thinking process involves fuzzy or uncertain descriptions. An FCM presents a graphical representation utilized to describe the cause and effect relations between nodes, thus giving us the opportunity to describe the behavior of a system in a simple and symbolic way.

The FCM theories embody the accumulated knowledge and experience from experts who know how the medical system behaves in different circumstances. This is one of the strongest points of the FCM approach that rely mainly on the experience of the experts. This knowledge is extracted

using linguistic variables, which then are transformed to numeric values using a defuzzification method. Figure 1 shows a representative diagram of an FCM. During the simulation, the value of each concept is calculated using the following rule:

$$A_i(k+1) = f\left(k_2 \times A_i(k) + k_1 \times \sum_{j=1,\ j\neq i}^{N} A_j(k) \times W_{ji}\right) \quad (1)$$

Where N is the number of concepts, $A_i\ (k+1)$ is the value of the concept C_i at the iteration step $k+1$, $A_j(k)$ is the value of the concept C_j at the iteration step k, W_{ji} is the weight of interconnection from concept C_j to concept C_i. The constant "k_1" expresses the influence of all the other interconnected concepts on the configuration of the new value of the concept A_i, and "k_2" represents the proportion of the contribution of the previous value of the concept in computing the new value. Their values usually are set equal to 1, unless the experts of a particular dynamic system can determine these influences and contributions to the behavior of the system. Meanwhile, function f is the sigmoid function:

$$f = \frac{1}{1 + e^{(-\lambda \times x)}} \quad (2)$$

where $\lambda > 0$ determines the steepness of function f. The FCM's concepts are given some initial values, which are then changed depending on the weights W_{ij} and the way the concepts affect each other. The calculations stop when a steady state is achieved, i.e., when the concepts' values become stable.

However, despite the many theoretical developments of FCMs, based on the material presented in the previous section and their success to many medical applications
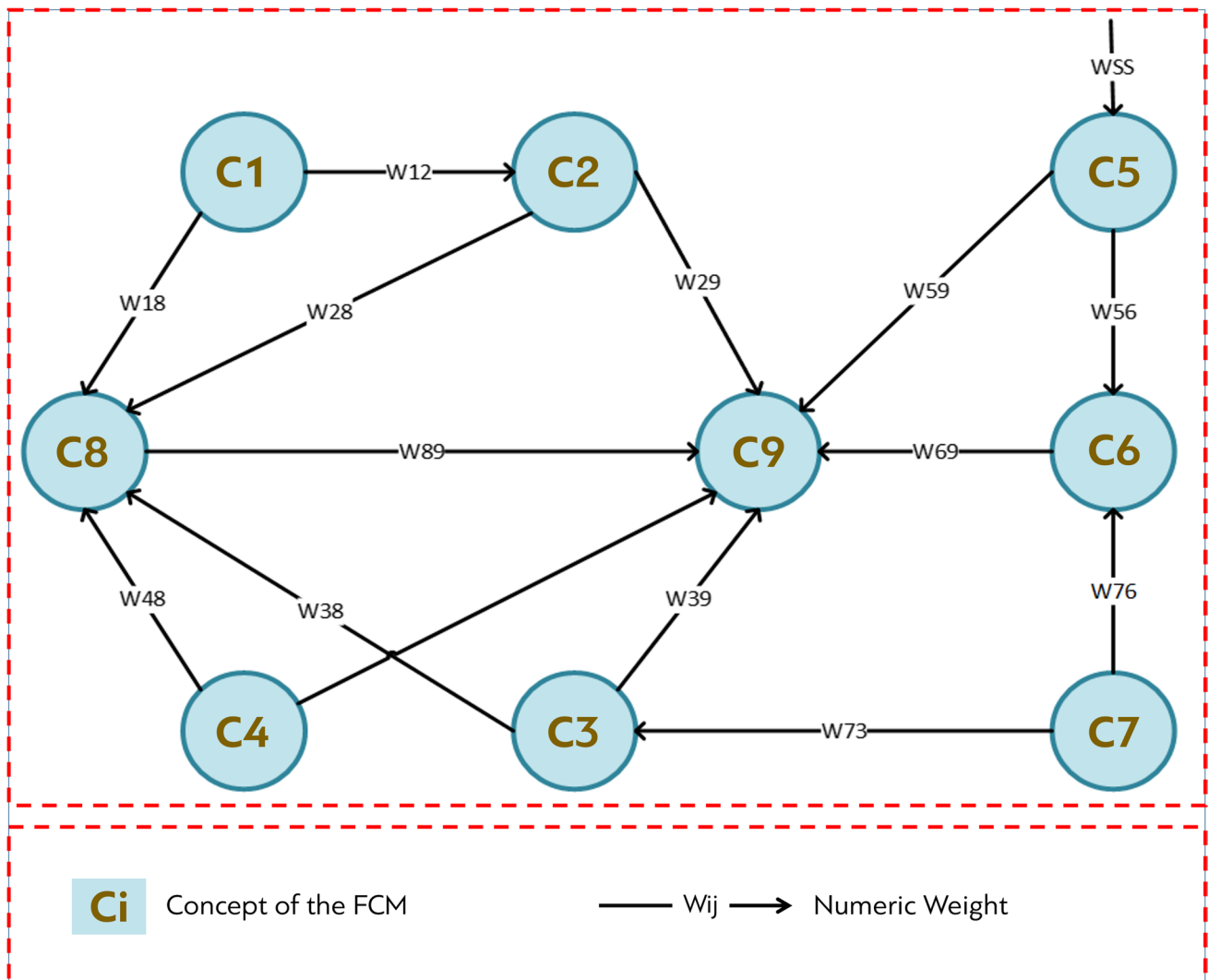

Fig. 1 A simple Fuzzy Cognitive Map (FCM)

(and not only), FCMs still have a number of limitations and drawbacks (Groumpos 2018; Mpelogianni et al. 2018). All theories are essentially based on the material of the "**Methods"** section, which will be referred from now on as the **"classical FCM theories"**. They do not go into the depth of the dynamic behavior of complex systems, especially the medical ones. In addition, the initial system structure described by experts and the learning principles-algorithms cannot follow the evolution of the medical complex dynamic system. The natural-human world evolves and the human-made systems progress by applying knowledge derived from observations of and familiarity with repeatable human events and phenomena of nature. Our perceptions, understanding, and ability to model medical problems enable us to develop policies, processes, and products that are invented and required in solving them. This requires new advanced interdisciplinary theories and technologies (Groumpos 2018; Mpelogianni et al. 2018).

The main problem stems from the existing approach that all concepts (say N) are grouped as one vector A. The values A of all N concepts at all instant of times, k, are calculated using Eqs. (1)–(2). Similarly, the learning-training algorithms for upgrading the causality coefficients W_{ij}, are performed for all N concepts and for all instant of times, k (Anninou et al. 2018; Groumpos 2010). The equations utilized for Hebbian learning are using all N concepts. This is not mathematically correct since, from the total number N of concepts, there is a number of concepts whose values remain constant for a long period of time when other concepts are changing slowly or fast. There are also exogenous variables that are not easily identified in the first place and thus cannot be part of the classical learning algorithms. Another issue is that, in the classical FCM theories, the algorithm was always converging to the same final values, regardless of the initial conditions. This, in return, forces us to increase the number of simulations consulting the behavior of the changes of the weights Wij through their changes based learning algorithms (Anninou et al. 2018).

Another drawback has to do with the Nonlinear Hebbian Learning (NHL) learning method (Groumpos 2010; Papageorgiou et al. 2002, 2003a, b). While running several simulations, we have observed that, due to the way weights are being calculated, if the number of iterations of the algorithm is increased, in order to reach a steady-state, the causality sometimes reverses all or some of the Wij values. This is a *very serious drawback as it changes the causality between concepts* (Mpelogianni et al. 2018) and in several occasions, instead of having a lower, we are going to have a larger result causing serious problems not only in the interpretation of the obtained results but also on stability issues to a number of real-life systems. This is very crucial to medical problems and especially for COVID-19.

All the above drawbacks and some other ones were addressed by the research team of the Laboratory for Automation and Robotics (LAR) of the University of Patras under the supervision of the author. The new approach addressing these drawbacks and deficiencies of FCMs is called the state space Advanced Fuzzy Cognitive Maps (AFCM) approach. The basic mathematical approach is to separate the total number of concepts N to (1) N1-State Concepts, (2) N2-Input Concepts, and (3) N3-Output Concepts.

The Advanced Fuzzy Cognitive Maps (AFCM) approach is an evolution of the classic methodology, which promises more accurate results for a large variety of complex medical systems. This methodology, which is thoroughly analyzed in (Groumpos 2010), overcomes major issues of the classic FCMs. Those issues are: (a) the presence of concepts of different nature is ignored in the traditional approach and, (b) the utilization of the classic sigmoid normalization function was fuzzing the system, especially in cases of the existence of several concepts. The AFCM approach was used in two medical applications with excellent results (Antigoni 2013; Anninou et al. 2018).

The proposed State–Space AFCM. Basic aspects and input concepts

The new proposed AFCM approach uses the basic theory of FCM with the difference that not all variables-concepts of the system are treated in the same way as in the "**classical FCM theories**".

The procedure of constructing the AFCM is remaining the same (Mpelogianni et al. 2018; Apostolopoulos et al. 2021). First, experts should define all the basic system's variables-concepts. Then, the main change and innovation are that after this procedure, variables-concepts should be divided into three categories:

- Fuzzy concept states
- Fuzzy concept inputs
- Fuzzy concept outputs

The fuzzy concept states of a dynamic system refer to a minimum set of variables, known as state variables, which fully describe the system and its response to any given set of inputs. The fuzzy concept inputs concern signals that stimulate the system. The fuzzy output variables constitute those that we should examine their behavior. In that way, we take into consideration what exactly each concept does. We no more treat outputs in the same way as inputs, and inputs are separated from states. The mathematical description of the system and the combination of initial states and inputs are sufficient to provide information about both the future states and outputs. Mathematically, the standard form of the new

model is described by the system general equations (Mpelogianni et al. 2018),

$$X[k+1] = Ax[k] + Bu[k] \qquad (3)$$

$$Y[k+1] = Cx[k] + Du[k] \qquad (4)$$

where k = 0,1,2,3……..until convergence

$$Y(Z) = Ex(Z) \qquad (5)$$

where Z = 1,2,3…. until the final desired time. In addition, $x(k) \in R_{nxn}$ is a concept state vector, $u(k) \in R_{rxr}$ is an exogenous known concept input vector and is in the discrete-time unit "k" and $y(k) \in R_{mxm}$ the output concept vector. In addition, $y(Z) \in R_{pxp}$ is the concept output vector, and "Z" is the discrete instant of time(s) Z that we are interested to know the concept outputs. For example, the discrete times can be every 12, 18 or 24 h. The discrete times k and Z can change according to each individual case. The matrices A, B, C, and D are constant matrices having been determined by the engineers in close consultation with the physician staff. A similar approached for the FCM application for the knee case.

Please note that the discrete-time Z of Eq. (5) is different from the discrete times of k = k_1, k_2, k_3……not only of Eqs. (3) and (4) but also of Eq. (1). The discrete-time k of Eqs. (3) and (4) can also be different from the k of Eq. (1). A lot of discussions are required for this topic and need to be addressed in future research.

The discrete-time Z is set by the medical doctors. This is in stark contrast with the "k" variable in Eq. (1). The latter expresses the iteration steps until convergence. In the new model, the number of iteration steps depends on the samples that have been available at all times, and the final time Z when the outcome result is desired. In the new AFCM method, Eqs. (3) and (4), the problem of convergence is complicated and needs to be addressed in future research.

The algorithm that describes the new model in detail, adjusted in the prediction of the pandemic spread 19, is presented as follows: (with the term experts are the medical personnel).

Step 1: Experts define the basic concepts (known until today).
Step 2: Experts separate them into the Fuzzy States, Fuzzy Inputs, and Fuzzy Outputs.
Step 3: Experts construct an AFCM schema-diagram drawing the appropriate interconnections.
Step 4: Experts construct weight matrices between States-States, States-Inputs, States-Outputs, Inputs-Outputs using fuzzy variables, The matrices A, B, C, and D of Eqs. (3) and (4) are determined by the engineers in close consultation with the medical doctors.

Step 5: Convert weight matrices into numeric values using a defuzzification method.
Step 6: Determine Initial states in time unit "0" (Patients Initial Symptoms)-initial conditions.
Step 7: Determine also all Inputs for discrete times "k".
Step 8: Run simulations for the AFCM till convergence.
Step 9: Obtain the output Y(Z). The concept that is associated with the spread prediction.

For creating the proposed model, the advanced methodology, as proposed in Mpelogianni and Groumpos (2019); Apostolopoulos et al. (2021), was followed. Based on the recent literature for the pandemic of COVID-19, we propose some fundamental concepts to describe the spread of the disease. The proposed concepts are illustrated in Table 1.

Concept #1: This factor may be utilized to estimate the current situation (Kucharski et al. 2020). For example, if the days since the first death was recorded are below 30, we can assume that the spread has not generally reached its peak yet and we can expect an ascending behavior in the future. This concept will only affect a particular state concept, as described later.

Concept #2: This factor can be used as an indicator of the pandemic trend, although its one-sided assessment does not provide much information. Similar to the first concept, this input element affects a particular state concept described later.

Concept #3: This concept behaves as an estimator of the significance of the pandemic at a particular period of time (week) and could affect (Giuliani et al. 2020).

Concept #4: The number of intubated in an intensive care unit may both affect the quality of the healthcare system, considering a possible insufficiency in available units.

Concept #5: The restriction policy applied to the habitants plays a vital role in handling the outbreak, as several real-life facts demonstrated during the recent months. More specifically, the stricter the measures, the more controllable the spread of the disease will be, given that the measures have been taken in time (Sattenspiel and Herring 2003).

Concept #6: The restriction policy referring to the industrial zones of a state is perhaps the most decisive factor to control the spread of the disease, due to the high congestion observed in large industries.

Concept #7: The state concept referred to as the "impact of restrictions" aims to handle both the individual and the industrial restrictions to produce a combined factor affecting the spread of the disease.

Concept #8: The number of intensive care units per thousand populations is a factor influencing the sufficiency and the quality of the healthcare system and may constitute an essential parameter, as a possible insufficiency in available units may result in a growth of the outbreak.

Table 1 The concepts of the system

#	Concept definition	Type	Possible values	Relation with
1	Days since the first death occurred	Input Concept	0–180	13
2	Rate of increase in recorded deaths over the last week	Input Concept	− 1000% to + 1000%	11,13
3	The exact number of patients at a given time per 10,000 population	Input Concept	Decimal number	7,13
4	Intubated in an intensive care unit	Input Concept	Integer	9,13
5	Restriction policy for habitants	Input Concept	Herd immunity (no restrictions), light restriction policy, (medium austere police), austere restriction, state of emergency	7
6	Restriction policy for industry	Input Concept	Herd immunity (no restrictions), light restriction policy, (medium austere police), austere restriction, state of emergency	7
7	Impact of restrictions	State Concept	Consistent, relaxed, violation of rules	13
8	Intensive Care Units per 1000 population	Input Concept	Integer	11
9	Available medical staff per 1000 population	Input Concept	Integer	11
10	Quality of Healthcare System according to The Healthcare Access and Quality (HAQ) Index	Input Concept	0–100	11
11	Healthcare system effectiveness	State Concept	Very bad, bad, adequate, average, good, excellent, state of the art	13
12	Current situation estimate	Input Concept	Descending, steady descending, Steady, steady — ascending, ascending, gallop	13
13	Prediction	Output Concept	Descending, steady descending, Steady, steady – ascending, ascending, gallop	-

Concept #9: Highly related to the previous concept, the number of medical experts per thousand populations is correlated to the mortality rate (Gulliford 2002).

Concept #10: The ranking of the specific country, according to World Health Organization, is utilized as a factor affecting state concept #11 (i.e., the effectiveness of the healthcare system).

Concept #11: Concepts #8–10 are contributing to the state concept "Effectiveness of the Healthcare System," which is related to the spread of the virus to the society.

Concept #12: The estimation of the current situation is an optional parameter supplied to the model, aiming to increase its perception of the current situation.

Concept #13: The output concept, which attempts to predict the spread of the disease the next week, based on the inputs and the states of the last week.

Revised sigmoid function

In order to apply the FCM methodology, the values of all the input concepts must be between the interval [0,1], where 0 (zero) denotes that the value of the concept is very small and 1 (one) that the value is very big. This is the reason why the sigmoid function (Eq. (2)) is needed; to suppress the result to the interval [0,1]. But due to the shape of the curve any concept which, because of the additive calculation method (Eq. (1)) is assigned a value

beyond 3 leads the sigmoid function to correspond it to the value one which is greatly problematic as the final output is interpreted as "high" even if this is not always the expected or correct result. Continuing on the subject of the sigmoid function, there is another drawback that leads to high output values. This is the fact that the center of the curve, instead of being on the (0,0) point on the xy axis, it is on the (0.5,0) point. This means that each concept's lowest value can be 0.5. This problem, combined with the first one, makes it difficult to interpret the result even with the use of the experts' interpretation criterion (Eq. (3)).

$$R(x) = \begin{cases} 0, x < 0.5 \\ \frac{x-0.5}{0.5}, x \geq 0.5 \end{cases} \tag{6}$$

However, this is not efficient for medical problems. For this reason, we use the new approach proposed in Mpelogianni and Groumpos (2018). The sigmoid function (Eq. (2)) is changed to Eq. (7); so to fit the values to the desired interval, the slope of the curve changes depending on the variable we want to fit.

$$f(x) = m + \frac{M - m}{1 + e^{-r*(x-to)}} \tag{7}$$

where $x \in R$ and $f(R) = (m,M)$.

In the above-mentioned equation, we have.

- m is the lower limit of the curve
- M is the upper limit of the curve
- r is the slope of the curve and
- to is the symmetry to the y-axis

Based on the simulations, the slope of the sigmoid function was set to 1.5.

Weight suggestions

Ideally, the weight tables are defined by the experts according to their knowledge and experience. For the needs of this proposal and to make a first assessment of the model, the weight tables are arbitrarily determined by the authors, based on their own assessment, but also in direct proportion to the response of the algorithm.

The weight definition field was defined for this task as follows: (a) very weak impact (VW), (b) weak impact (W), (c) medium impact (M), (d) strong impact (S), and (e) very strong impact (VS). The weights may take negative values according to the mentioned categories.

The overall weight table referring to the proposed AFCM system is presented in Table 2. As the values of the weights reflect real-time relations, a dynamic approach is considered to alter the weight values over time, according to a specific set of rules, which is explained in the "The proposed State – Space AFCM. Basic aspects and input concepts" section. In Table 2, information regarding which weight is also affected by a certain rule is given. For example, the impact of the first concept on the output is by default set to "Very Strong", however, this impact may alter later when the pandemic is heading towards the end of its cycle.

Rules

There are several approaches to achieve the appropriate dynamism to transform the model from a static one to a robust one, able to adapt to changing environments. In machine learning, this is achieved through extensive training on large sets of data. In this way, the connections between the concepts are adjusted according to the situation, aiming to reflect the behavior of the training data. The AFCMs can be benefited from a similar approach, Hebian learning (Papageorgiou et al. 2003a, b, c); however, since the pandemic is not over, no complete data are available to explore this strategy.

To overcome this issue, it is proposed to apply a set of rules which are designed with respect to the epidemic curve, which is a graph in which the number of new cases of an outbreak disease is plotted against time. Epidemic curves are ordinarily constructed after the disease outbreak is over. However, in the case of Coronavirus disease, good estimates have been extensively utilized by the stated and the epidemiologists to assess the situation and plan possible measures. We propose to fit the model under consideration to the epidemic curve, i.e., to let the model know that the disease will approximately follow this distribution. The reader should note that only the general trend of the pandemic curve is adopted here as a guidance for the model. In essence, the model is developed to seek and identify a pandemic peak, before it alternated its behavior. Depending on the current situation, which is defined by the values of the concepts, this general trend, curve, slope, standard deviation may be altered.

The set of rules proposed for a more precise approach to reality is presented in Table 3.

Table 2 Table of weights

#	Concept definition	Impact on concept	Impact on output	Affected by rule
1	Days since the first death occurred	-	Very Strong	Yes
2	Rate of increase in recorded deaths over the last week	Very Strong	Strong if Positive, Medium if Negative	No
3	The exact number of patients at a given time per 10,000 population	Negative Strong	Medium	No
4	Intubated in an intensive care unit	Negative Weak	Weak	No
5	Restriction policy for habitants	Medium	-	No
6	Restriction policy for industry	Strong	-	No
7	Impact of restrictions	-	Very Strong	Yes
8	Intensive Care Units per 1000 population	Strong	-	No
9	Available medical staff per 1000 population	Very Strong	-	No
10	Quality of Healthcare System according to The Healthcare Access and Quality (HAQ) Index	Medium	-	No
11	Healthcare system effectiveness	-	Strong if Positive, Medium if Negative	No
12	Current situation estimate	-	Medium	No
13	Prediction	-	-	No

Table 3 Set of rules

Condition	Outcome
While Concept 1 < peak	Impact on output = positive descending
While Concept 1 > peak	Impact on output = negative ascending
If the change in restriction happens after the current situation estimate equals spreading or galloping	Impact on output = Weak
If number of patients in an intensive care unit equals the capacity of the healthcare system	Healthcare system effectiveness impact = Very Weak

Predicting the peak

If the total number of patients is increasing, but at a declining rate, then there is a serious indication that the outbreak is reaching its peak value. To record this trend in our model, a global variable denoted as reach_peak is introduced, with values zero or one. To define the actual peak, the following conditions must be true: (a) the rate of increase in confirmed disease patients is below 1%, and (b) the value of reach_peak is one.

Experimental cases

To evaluate the conception, specific simulations were performed, utilizing the available statistics related to the spread of COVID-19 for Greece, South Korea, and Germany. The simulations were performed over one month, separated into four weeks. The inputs of the model are the corresponding weakly averages, while the prediction (i.e., the output) refers to the severity of the spread for the following week. We selected to perform the simulations at different times for each country to investigate the performance over a larger period.

Greece

For Greece, the period of simulation is from 12 March 2020 until 12 April 2020. The statistics were collected from the National Health, while the information related to the quality of the healthcare system was collected from the website of the Ministry of Health. The rate of confirmed new cases is illustrated in Fig. 2.

South Korea

For South Korea, the period of simulation is from 1 March 2020 until 1 April 2020. The statistics were collected from the Centers for Disease Control & Prevention, while the

information related to the quality of the healthcare system was collected from the website of the Ministry of Health and Welfare. The rate of confirmed new cases is illustrated in Fig. 3.

Germany

For Germany, the period of simulation is from 15 March 2020 until 15 April 2020. The statistics were collected from the Robert Koch Institute, while the information related to the quality of the healthcare system was collected from the website of the Federal Ministry of Health. The rate of confirmed new cases is illustrated in Fig. 4.

Results

In the following Table, the weekly predictions of the AFCM model are presented for each country. Besides, the numeric predictions for each week are provided. Due to interconnections between the input concepts and the state concepts, some iterations were necessary for the system to reach a steady-state for each week. Those iterations are also presented in Table 4. The column "Actual Situation" is computed by the statistics and corresponds to the real progress of the pandemic for the week under consideration. The numeric output is the model's actual response after a stable equilibrium is achieved. This equilibrium is reached after the number of iterations provided in the relevant column. The numeric output may be interpreted for comparison reasons to a verbal statement describing the prediction of the spread for the following week.

The response of the system demonstrates its effectiveness in two situations: (a) the correct prediction of the alteration of the situation from steady-ascending to steady in the example of Greece, (b) the correct prediction of the alteration of the situation from steady-ascending to ascending and then back to steady ascending in the example of the outbreak in Germany. Due to the absence of available information related to a descending behavior of the spread, the model could not be evaluated further. In Table 5, the input values for the first two weeks in Greece are given.

The system was able to capture the disturbance caused by the predictor concepts 1–6 and concept 7 to the response concept 13 and turn this disturbance to an interpretable outcome connected to the risk of the spread of COVID-19 the next week. For the system to reach a steady-state and to respond in the desired way, the translation of the input values to desired intervals was mandatory. For example, the value

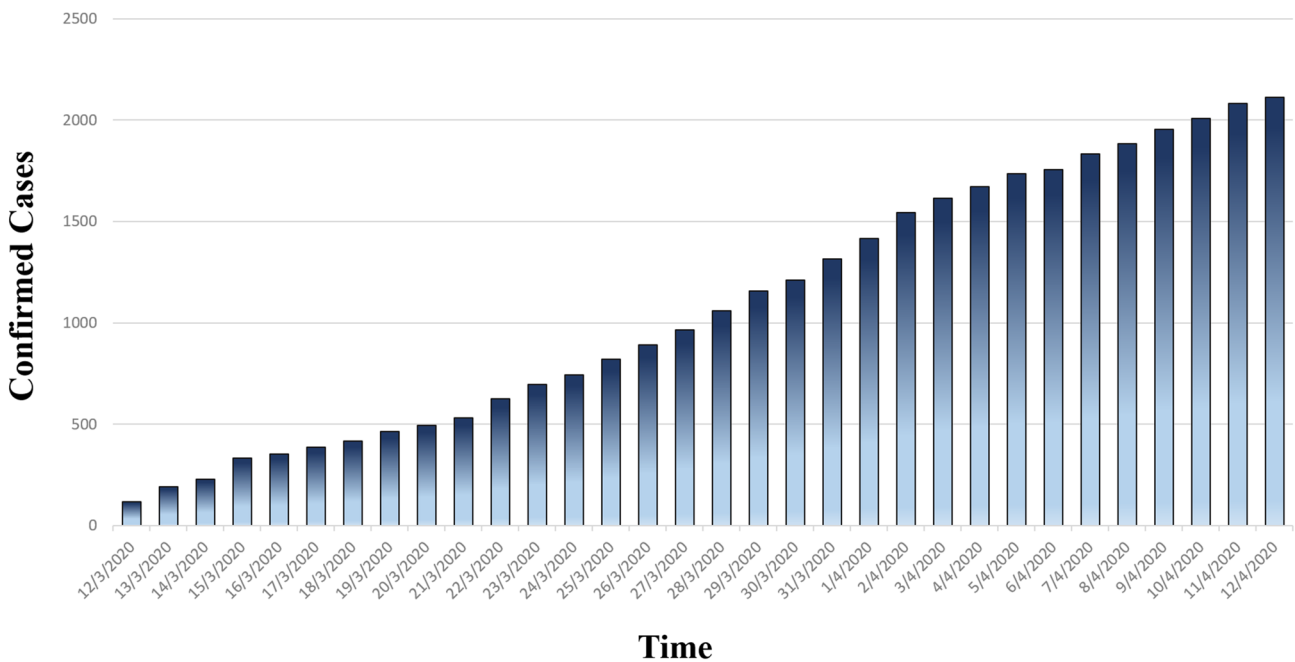# Rate of confirmed instances for Greece
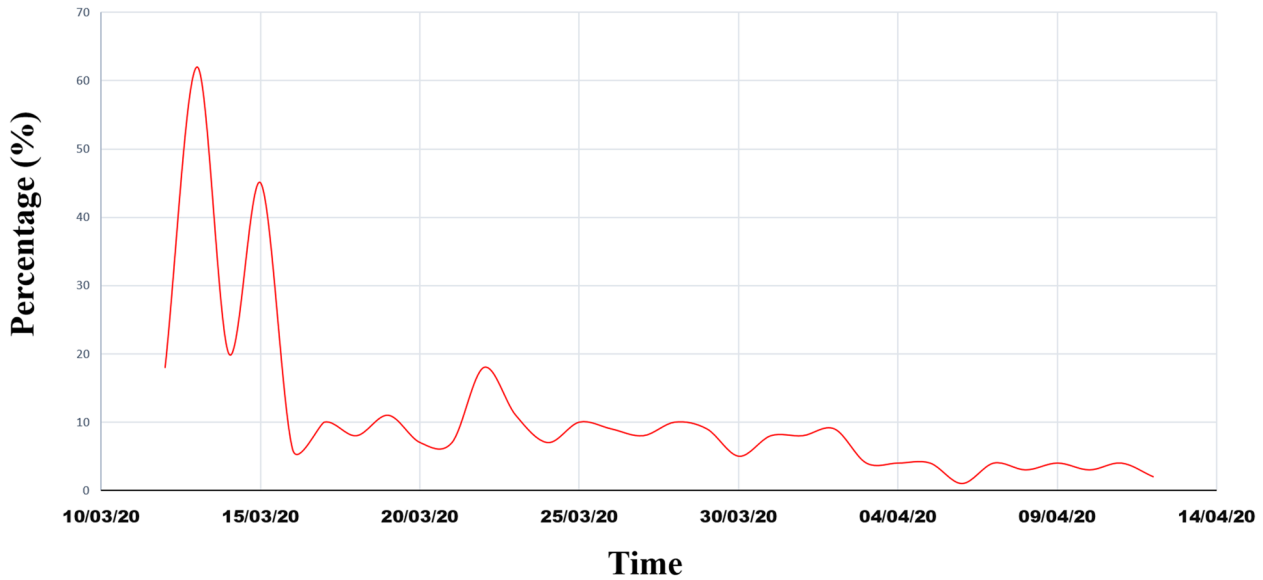
Confirmed cases on a daily basis

Fig. 2 The rate of new confirmed diseased cases for Greece between 12/3/2020 and 12/4/2020

of concept 1 (days) is expected to grow every week by 7, which must be interpreted as a small change in the value of the concept handled by the algorithm; therefore, the actual number of days were divided by 180 to fit the interval (0,1). The same strategy was followed for all the input concepts, as it is shown in Table 4.

Rate of confirmed instances for South Korea
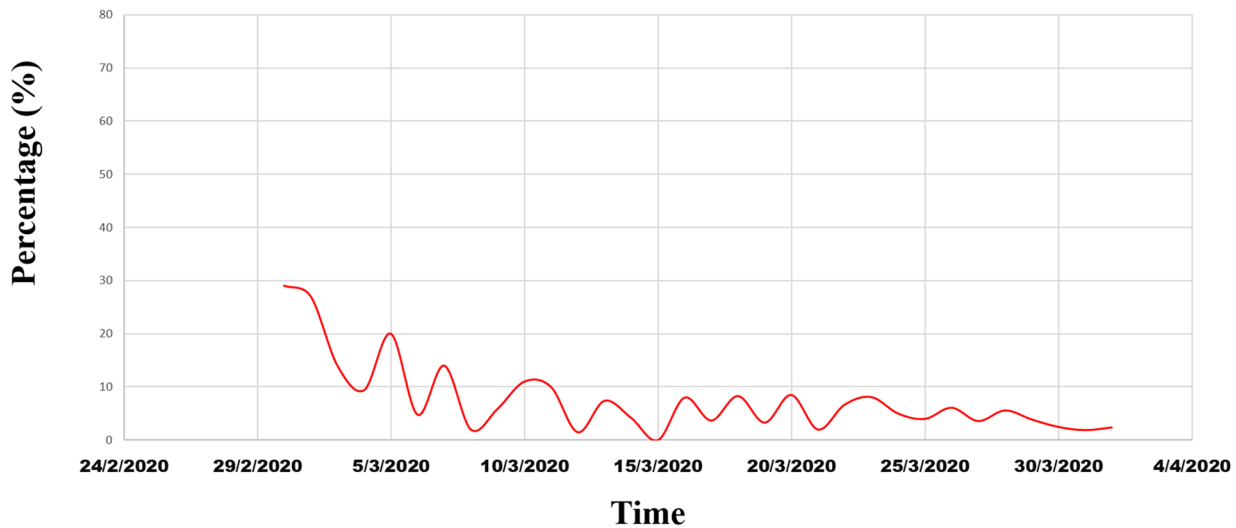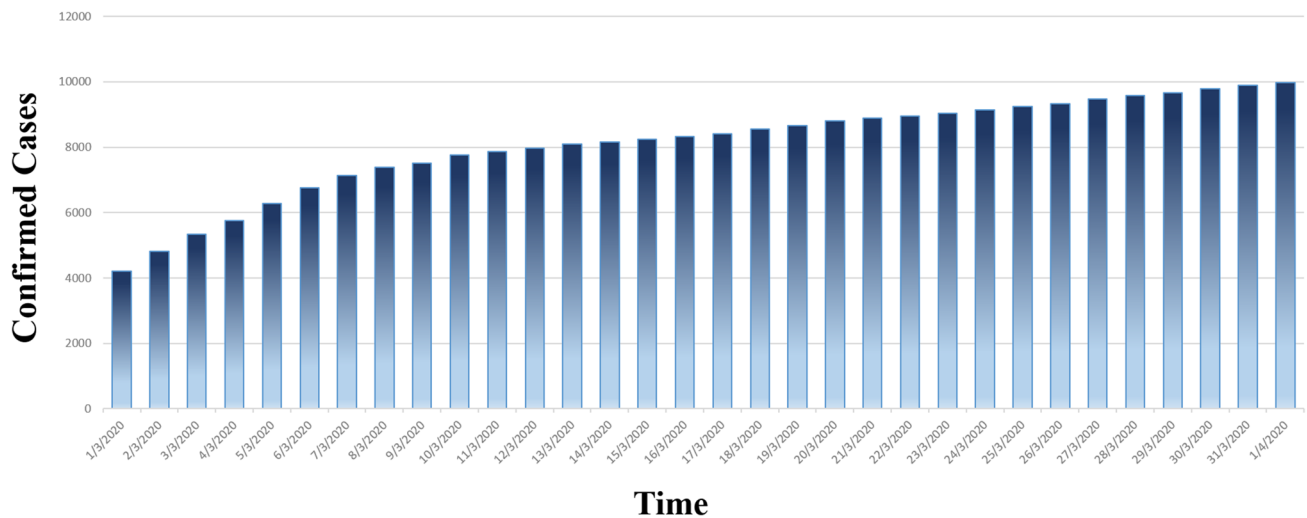

Confirmed cases on a daily basis

Fig. 3 The rate of new confirmed diseased cases for South Korea between 1/3/2020 and 1/4/2020

### Results in terms of predicted patient cases

To further evaluate the system, we simulate the COVID-19 outbreak in Greece, South Korea, and Germany. Tables 6 and 7 summarize the results.

For Greece, the proposed methodology is slightly underestimating the potential spread of the pandemic. With a mean absolute error of 210 out of the 2114 total confirmed cases of the period under consideration, an estimation error is calculated at 9.9%. The time series of confirmed versus predicted patients for Greece is illustrated in Fig. 5.

For South Korea, the model is underestimating the potential spread of the pandemic, as in Greece. With a mean absolute error of 1752 out of the 9976 total confirmed cases of the period under consideration, an estimation error is calculated at 17.9%%. The time series of

Rate of confirmed instances for Germany
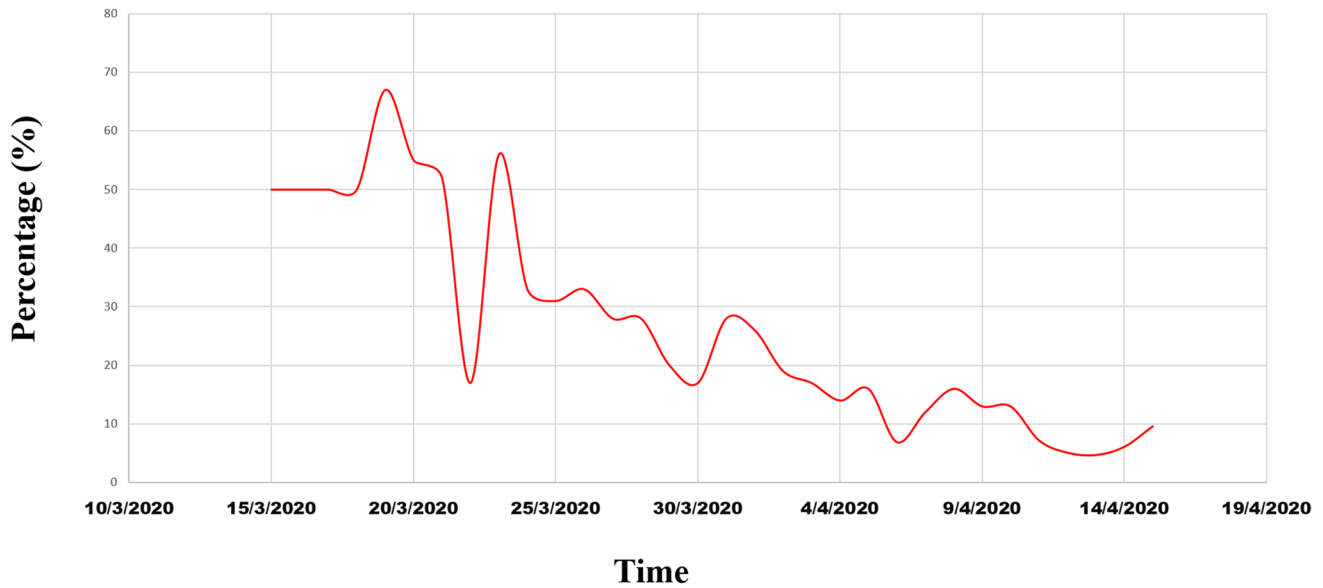
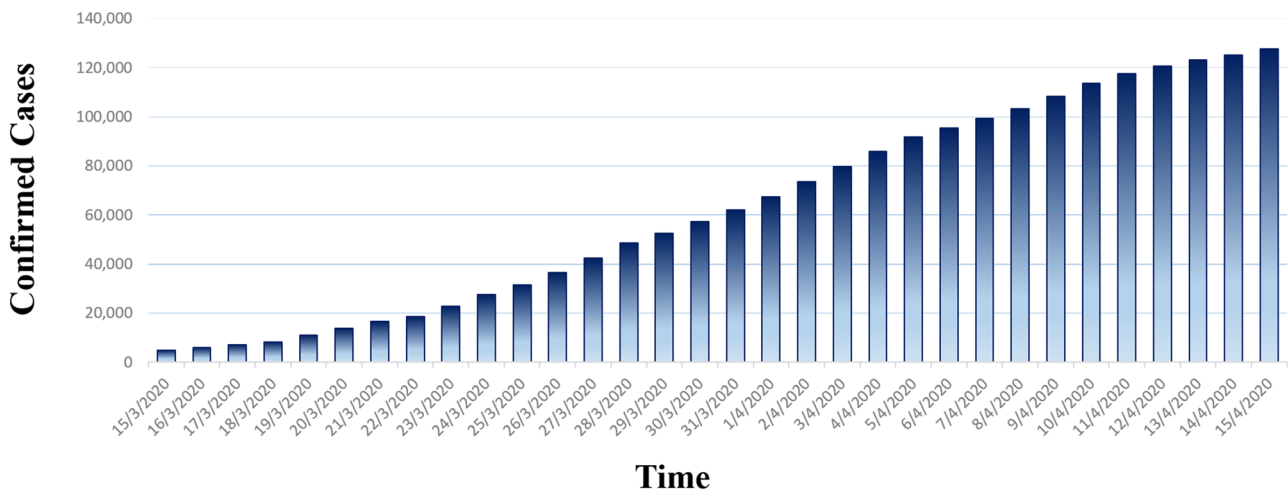
Confirmed cases on a daily basis

Fig. 4 The rate of new confirmed diseased cases for Germany between 1/3/2020 and 1/4/2020

confirmed versus predicted patients for South Korea is illustrated in Fig. 6.

For Germany, the model is underestimating the potential spread of the pandemic in the beginning and overestimating the spread later. With a mean absolute error of 7985 out of the 127,584 total confirmed cases of the period under consideration, an estimation error is calculated at 6%. The

time series of confirmed versus predicted patients for South Korea is illustrated in Fig. 7.

Despite the observed errors, the model is accurately predicting the trend of the spread, yielding remarkable Pearson's Correlation Coefficient (r) and Coefficient of Determination (R^2). For Greece, r and R^2 are 0.9976 and 0.9952, respectively. For South Korea, r and R^2 are estimated at

Table 4 Results of the simulation

Country	Week	Numeric output	Verbal output	Iterations	Actual situation
Greece	Week 1	0.58	Steady Ascending	7	Ascending
	Week 2	0.56	Steady Ascending	5	Steady Ascending
	Week 3	0.49	**Steady**	3	Steady
South Korea	Week 1	0.68	Steady—Ascending	6	Ascending
	Week 2	0.61	**Steady—Ascending**	8	Steady Ascending
	Week 3	0.57	Steady—Ascending	3	Steady Ascending
Germany	Week 1	0.63	Steady—Ascending	3	Steady Ascending
	Week 2	0.76	**Ascending**	4	Ascending
	Week 3	0.71	Ascending	6	Ascending

Table 5 Input values for Greece for the first two weeks

#	Concept definition	Week 1 actual values	Week 1 translated values	Week 2 actual values	Week 2 translated values
1	Days since the first death occurred	5	0.027	12	0.072
2	Rate of increase in recorded deaths over the last week	46%	0.046	29%	0.029
3	The exact number of patients at a given time per thousand population	0.035	0.035	0.0695	0.0695
4	Intubated in an intensive care unit	9	0.0009	34	0.0034
5	Restriction Policy for Habitants	Austere	0.8	Austere	0.8
6	Restriction Policy for Industry	Medium	0.5	Austere	0.8
7	Impact of Restrictions	Strong	0.75	Very Strong	0.85
8	Intensive Care Units per 1000 population	0.06	0.06	0.06	0.06
9	Available medical staff per 1000 population	6.3	0.25	6,3	6.3
10	Quality of Healthcare System according to The Healthcare Access and Quality (HAQ) Index	14	0.75	14	0.75
11	Healthcare System Effectiveness	Medium	0.55	Medium	0.54
12	Current Situation Estimate	No Input	-	No Input	-

Table 6 Statistical analysis of confirmed and predicted COVID-19 patient cases for each country

	Greece	South Korea	Germany
Confirmed cases			
Mean	1081	8094	62,459
Standard deviation	654	1497	42,342
Predicted cases			
Mean	870	9846	65,135
Standard deviation	450	2729	50,764

**Table 7** Errors and correlation between confirmed and predicted COVID-19 patient cases per country

Metrics (Confirmed vs Predicted)	Greece	South Korea	Germany
Mean Absolute Error (patients)	210	1752	7985
Correlation	0.9976	0.977	0.989
R^2	0.9952	0.941	0.9782

0.997 and 0.9941, respectively. Finally, for Germany, r and R^2 are 0.989 and 0.9782, respectively.

Discussion

A new method was proposed, the Advanced Fuzzy Cognitive Map (AFCM), for the simulation of the Coronavirus pandemic characteristics in various countries. The simulations were performed utilizing the statistical data from South Korea, Germany, and Greece. The model was able to capture the progress of the disease and to predict the change of the situation very accurately from an ascending spread to a steadier situation occurring when the spread is reaching its peak value. This behavior can be observed particularly in Figs. 6 and 7, where the two lines are simultaneously changing their trend. The reader should note that the particular modeling methodology is not depending its behavior on any pandemic curve. While compiled to expect a certain peak of the spread at some

Fig. 5 Confirmed and Predicted COVID-19 patients time series for Greece
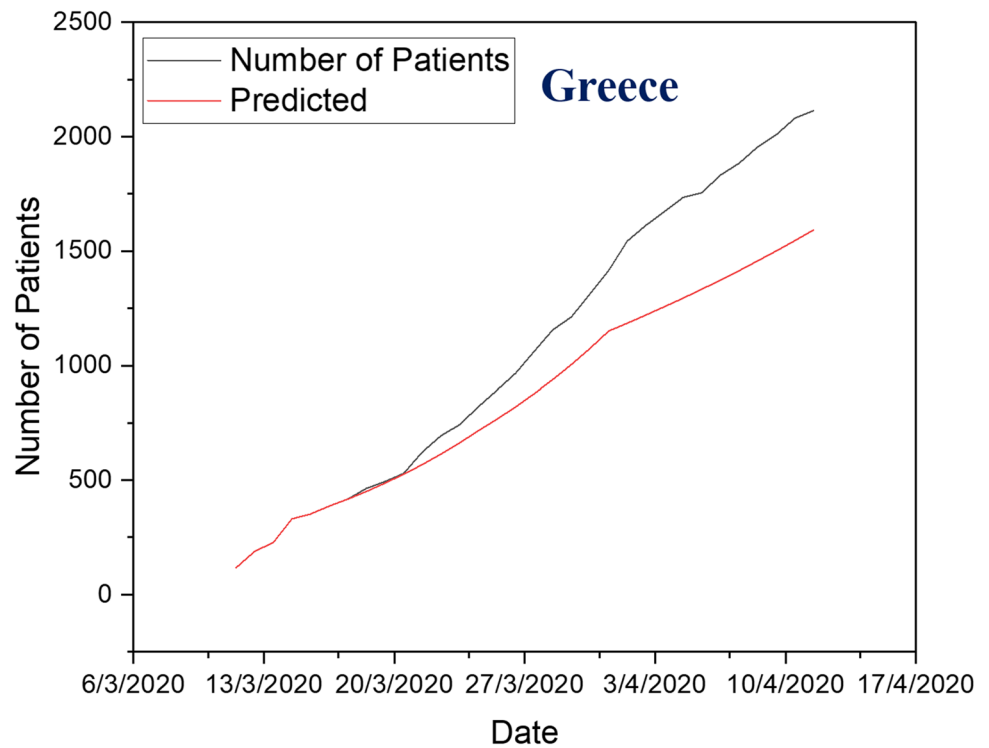
Fig. 6 Confirmed and Predicted COVID-19 patients time series for South Korea
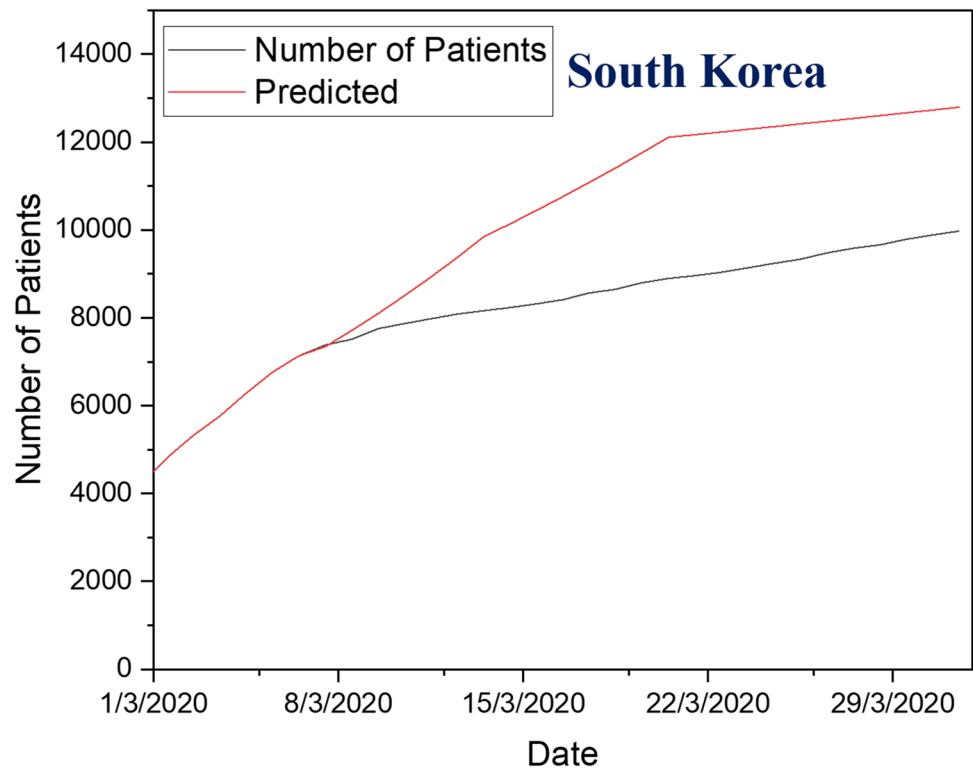

point of time, this would not always be the case. Altering specific concepts describing the current situation (e.g. impact of restrictions, restriction policy, and effectiveness of the healthcare system) would change the behavior of the model and could overcome the need of assigning a peak. More simulations are expected in the near future to investigate the response of the methodology in very uncertain conditions.

Fig. 7 Confirmed and Predicted
COVID-19 patients time series
for Germany
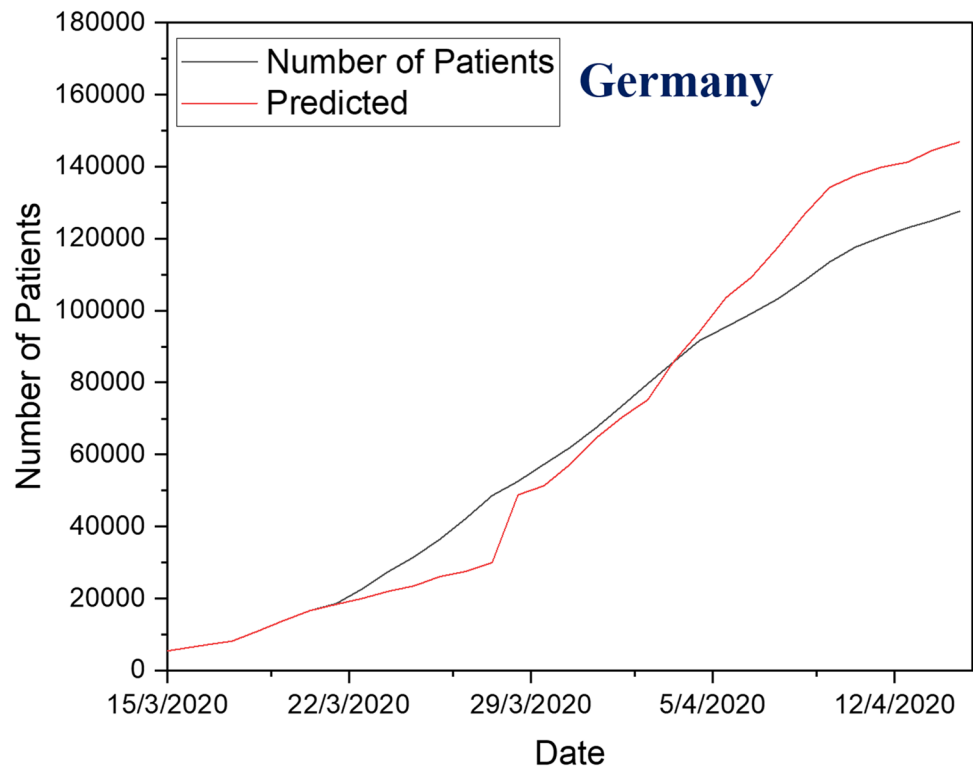
The obtained results demonstrate promising effectiveness of the proposed system despite the fact that the weights of the input and the state concepts were not analyzed or defined by experienced epidemiologists or data analysis experts. The approach of the current work exploits the benefits of the AFCMs, such as the convenient modeling of the causal relations between factors of different nature and their ability to perform in changing real-time environments by making full use of the disturbance caused by the one concept to the other. The idea of studying the known causality between pandemic factors and not predicting and discovering new ones is the core of AFCMs. The statement that "correlation does not imply causality, while causality always implies correlation" can be better understood with a specific example. Car accidents were reduced during the COVID-19 pandemic, due to restriction measures. One can notice that a negative correlation between the pandemic spread and car accidents is indeed present. However, this correlation does not imply any direct cause and effect relation between those two phenomena. For example, we do not expect that if the car accidents are increased, the pandemic will decline.

The study has some limitations. Firstly, the model was not fully evaluated regarding its performance of predicting the spread after the peak value is reached. This is due to the absence of real data and conformed data at the time this research was conducted. Secondly, the weights were defined by the authors and did not necessarily reflect the actual relations between the factors incorporated into the

AFCM. A more scientifically valid weight definition, which can be produced through cooperation between epidemiologists and committees, could enhance the robustness of the model in the future. The reader should note that the desired output depends heavily on the assigned weights and the pre-defined input concepts. In the present simulation, although arbitrarily defined, the selected concepts and weights yield very promising results. Thirdly, for a complete and accurate estimation, more factors have to be included. For example, a number of environmental factors may be incorporated into the model to describe the spread of communicable diseases that are prone to cause epidemics, such as the water supply facilities, the sanitation facilities, the quality of food, and the climate. Climate can affect disease transmission in a variety of ways. The distribution and population size of disease vectors can be heavily affected by the local climate. One more fact must be taken into consideration, which is many more real data are needed and form more countries.

Nevertheless, the proposed methodology may be appropriate to overcome the limitations of the statistical analysis, wherein several predictors are not taken into consideration (for example, the quality of the healthcare system). As it is demonstrated by the first simulation results, the AFCM modeling methodology succeeds in providing a reliable prediction for the spread of the disease both in early situations and in near-peak conditions. The decent handling of seemingly tough-to-handle attributes increases the dynamism of

the predictions and may be able to circumvent the relatively static approach of statistical methodologies.

This study gives strong evidence that the FCM theories are probably the only ones exploring the causality between the variables of medical problems in a sound mathematical and scientific foundation.

The response of the model in unexpected pandemic reactions and trends is to be investigated in the near future. More specifically, the effectiveness of the AFCMs in predicting situations where the pandemic trend is not following an expected pandemic curve, or the pandemic trend is showing heavy instability due to a number of reasons, is a future research opportunity.

Physicians are requested to handle information of different nature, e.g., patient's history, clinical diagnostic tests, medical images, personal health problems, and demographic characteristics (Apostolopoulos et al. 2020). The interpretation of these results involves ambiguity, fuzziness, and uncertainty, which plays a critical role in decision-making to a wide and diverse set of medical problems. The COVID-19 pandemic has all these characteristics. Although these approaches provide us with answers that are needed and used to study medical problems, they are still not sufficient and adequate to provide us with acceptable and convincing solutions. Therefore, new advanced scientific approaches and new mathematical models are urgently needed.

Conclusions

Accurate and reliable Medical Decision Systems predicting the spread of dangerous pandemics can be a critical tool for society. Modern public health emergency management methods do not exhaust the necessary medical measures, such as improving health systems, but extend to the whole spectrum of social and economic life. For this reason, it is necessary to develop powerful simulation tools that are able to take many variables of different nature, to be able to respond to uncertain and unpredictable situations by exploring the causation between them. Such tools can be irreplaceable and helpful for scientists, especially to medical professionals.

The proposed AFCMs appear to be a robust methodology for creating and monitoring such complex modeling systems, while, in this study, their advantages and effectiveness are proven in practice. It is the authors' opinion that these methods can be further developed and gain the trust and attention of scientists, especially the medical professionals in the near future.

Future research should address the aforementioned limitations and perhaps explore the possibilities of machine learning approaches, such as the backpropagation utilized to train the networks. One option is the development of a hybrid AFCM-Neural Network-Statistical approach,

wherein the advantages of each method are combined to constitute a robust predicting model. Methods of Artificial Intelligence (AI) such as machine learning, deep learning, and probabilistic methods for uncertain reasoning, should also be used to study COVID-19 (Apostolopoulos and Groumpos 2020; Apostolopoulos and Mpesiana 2020).

New concepts that the medical doctors can propose for the spread of the pandemic would develop new AFCM models. Close cooperation with medical doctors is needed. Models must be developed for different COVID-19 affected people and for different geographical regions.

Data and materials availability All data utilized in this study are publically available.

Declarations

Conflict of interest The authors declare no competing interests.

References

Ala'raj M, Majdalawieh M, Nizamuddin N. Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. Infect Dis Modell. 2021;6:98–111. https://doi.org/10.1016/j.idm.2020.11.007.

Anninou A, Poulios P, Groumpos P, Gliatis I. A novel software tool for detection of meniscus injury using dynamic fuzzy cognitive networks. Physiother Rehabil. 2018;3:2573–3312. https://doi.org/10.4172/2573-0312.1000155.

Antigoni AP, Peter GP. Non linear Hebbian learning techniques and fuzzy cognitive maps in modeling the Parkinson's disease. 21st Mediterranean Conference on Control and Automation, IEEE. 2013, p. 709–15. https://doi.org/10.1109/MED.2013.6608801.

Apostolopoulos ID, Groumpos PP. Non - invasive modelling methodology for the diagnosis of coronary artery disease using fuzzy cognitive maps. Comput Methods Biomech Biomed Eng. 2020;1–9. https://doi.org/10.1080/10255842.2020.1768534.

Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med. 2020;43:635–40. https://doi.org/10.1007/s13246-020-00865-4.

Apostolopoulos ID, Aznaouridis SI, Tzani MA. Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases. J Med Biol Eng. 2020;40:462–9. https://doi.org/10.1007/s40846-020-00529-4.

Apostolopoulos ID, Groumpos PP, Apostolopoulos DJ. Advanced fuzzy cognitive maps: state-space and rule-based methodology for coronary artery disease detection. Biomed Phys Eng Express. 2021;7:045007. https://doi.org/10.1088/2057-1976/abfd83.

Axelrod R. The analysis of cognitive maps. Struct Decis. 1976:55–73.

Battiston P, Gamba S. COVID-19: R 0 is lower where outbreak is larger. Health Policy. 2021;125:141–7. https://doi.org/10.1016/j.healthpol.2020.10.017.

Bwire GM. Coronavirus: why men are more vulnerable to Covid-19 than women? SN Compr Clin Med. 2020;2:874–6. https://doi.org/10.1007/s42399-020-00341-w.

De Wit E, Van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. Nat Rev Microbiol. 2016;14:523. https://doi.org/10.1038/nrmicro.2016.81.

Galbadage T, Peterson BM, Gunasekera RS. Does COVID-19 spread through droplets alone? Front Public Health. 2020;8:163. https://doi.org/10.3389/fpubh.2020.00163.

Giuliani D, Dickson MM, Espa G, Santi F. Modelling and predicting the spatio-temporal spread of coronavirus disease 2019 (COVID-19) in Italy. SSRN J. 2020. https://doi.org/10.2139/ssrn.3559569.

Gralinski LE, Menachery VD. Return of the Coronavirus: 2019-nCoV. Viruses. 2020;12:135. https://doi.org/10.3390/v12020135.

Groumpos PP. Fuzzy cognitive maps: basic theories and their application to complex systems. Fuzzy cognitive maps, Springer. 2010, p. 1–22. https://doi.org/10.1007/978-3-642-03220-2_1.

Groumpos PP. Intelligence and fuzzy cognitive maps: scientific issues, challenges and opportunities. Stud Inf Control. 2018;27:247–64. https://doi.org/10.24846/v27i3y201801.

Groumpos P. Modelling COVID-19 using Fuzzy Cognitive Maps (FCM). EAI Endorsed Transactions on Bioengineering and Bioinformatics. 2021;1:168728. https://doi.org/10.4108/eai.24-2-2021.168728.

Gulliford MC. Availability of primary care doctors and population health in England: is there an association? J Public Health. 2002;24:252–4. https://doi.org/10.1093/pubmed/24.4.252.

Hijikata A, Shionyu-Mitsuyama C, Nakae S, Shionyu M, Ota M, Kanaya S, Shirai T. Knowledge-based structural models of SARS-CoV-2 proteins and their complexes with potential drugs. FEBS Lett. 2020;594:1960–73. https://doi.org/10.1002/1873-3468.13806.

Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395:497–506. https://doi.org/10.1016/S0140-6736(20)30183-5.

Kosko B. Fuzzy cognitive maps. Int J Man Mach Stud. 1986;24:65–75. https://doi.org/10.1016/S0020-7373(86)80040-2.

Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, Eggo RM, Sun F, Jit M, Munday JD, others. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. The lancet infectious diseases. 2020. https://doi.org/10.1016/S1473-3099(20)30144-4.

Li G, De Clercq E. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). Nature Publishing Group. 2020. https://doi.org/10.1038/d41573-020-00016-0.

Ling C. Traditional Chinese medicine is a resource for drug discovery against 2019 novel coronavirus (SARS-CoV-2). J Integr Med. 2020;18:87. https://doi.org/10.1016/j.joim.2020.02.004.

Lu S, Lin J, Zhang Z, Xiao L, Jiang Z, Chen J, Hu C, Luo S. Alert for non-respiratory symptoms of Coronavirus Disease 2019 (COVID-19) patients in epidemic period: a case report of familial cluster with three asymptomatic COVID-19 patients. J Med Virol. 2020. https://doi.org/10.1002/jmv.25776.

Mpelogianni V, Groumpos PP. Re-approaching fuzzy cognitive maps to increase the knowledge of a system. Ai Soc. 2018;33:175–88. https://doi.org/10.1007/s00146-018-0813-0.

Mpelogianni V, Groumpos PP. Building energy management system modelling via State Fuzzy Cognitive Maps and learning algorithms. IFAC-PapersOnLine. 2019;52:513–8. https://doi.org/10.1016/j.ifacol.2019.12.596.

Mpelogianni V, Arvanitakis I, Groumpos PP. State feedback of complex systems using fuzzy cognitive maps. Int J Bus Technol. 2018;6:1–6. https://doi.org/10.33107/ijbte.2018.6.3.14.

Murthy S, Gomersall CD, Fowler RA. Care for critically ill patients with COVID-19. JAMA. 2020;323:1499. https://doi.org/10.1001/jama.2020.3633.

Papageorgiou E, Stylios C, Groumpos P. Decision making in external beam radiation therapy based on fuzzy cognitive maps. Proceedings First International IEEE Symposium Intelligent Systems, vol. 1, IEEE; 2002, p. 320–5. https://doi.org/10.1109/IS.2002.1044275.

Papageorgiou E, Spyridonos P, Stylios C, Ravazoula P, Groumpos P, Nikiforidis G. A Fuzzy Cognitive Map model for grading urinary bladder tumors. 5th Int. Conf. in Neural Networks & Experts Systems in Medicine & Healthcare 1st Int. Conf. in Computational Intelligence in Medicine & Healthcare, NNESMED/CIMED Conference, vol. 6, 2003a, p. 2003.

Papageorgiou E, Spyridonos P, Stylios C, Nikiforidis G, Groumpos P. Grading urinary bladder tumors using unsupervised Hebbian algorithm for fuzzy cognitive maps. Int J Biomed Soft Comput Human Sci. 2003b;9:33–9. https://doi.org/10.24466/ijbschs.9.2_33.

Papageorgiou EI, Stylios CD, Groumpos PP. An integrated two-level hierarchical system for decision making in radiation therapy based on fuzzy cognitive maps. IEEE Trans Biomed Eng. 2003c;50:1326–39. https://doi.org/10.1109/TBME.2003.819845.

Sattenspiel L, Herring DA. Simulating the effect of quarantine on the spread of the 1918–19 flu in central Canada. Bull Math Biol. 2003;65:1–26. https://doi.org/10.1006/bulm.2002.0317.

Stang A, Standl F, Jöckel K-H. Characteristics of COVID-19 pandemic and public health consequences. Herz. 2020;45:313–5. https://doi.org/10.1007/s00059-020-04932-0.

Toğaçar M, Ergen B, Cömert Z. COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches. Comput Biol Med. 2020;121:103805. https://doi.org/10.1016/j.compbiomed.2020.103805.

Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O, Graham BS, McLellan JS. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020;367:1260–3. https://doi.org/10.1126/science.abb2507.

Xu X, Chen P, Wang J, Feng J, Zhou H, Li X, Zhong W, Hao P. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. Sci China Life Sci. 2020;63:457–60. https://doi.org/10.1007/s11427-020-1637-5.