Forecasting Models for Coronavirus (COVID-19): A Survey of the State-of-the-Art

Parikshit Mahalle [1], Asmita B. Kalamkar [2], Nilanjan Dey [2], Jyotismita Chaki [2], Aboul ella Hassanien [2], and Gitanjali R. Shinde [2]

[1]Sinhgad Institutess Smt. Kashibai Navale College of Engineering
[2]Affiliation not available

October 30, 2023

Abstract

When new virus and its respective disease cause more infections, it is very important to decide the strategies to control the spread and determine its impact. Considering the recent example of Coronavirus initially identified in Wuhan China has now targeted Italy badly, It is very important to study different forecasting models to control this pandemic. In the view of this, this study present the comparative analysis of various forecasting models, their classification and the techniques used. The detail analysis of forecasting models with respect to the parameters like data source, techniques, algorithms, mathematical parameters is also presented in this study. In the sequel, this study also presents useful recommendations to help government and healthcare community in designing better strategies and in taking productive decisions to control this outbreak.

Forecasting Models for Coronavirus (COVID-19): A Survey of the State-of-the-Art

Gitanjali R. Shinde[1*], Asmita B. Kalamkar[2], Parikshit N. Mahalle[3,4], Nilanjan Dey[5], Jyotismita Chaki[6], Aboul Ella Hassanien[7]

[1,2]Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Pune,Maharashtra, India.
[3]Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Pune,Maharashtra, India.
[4]Post Doc Researcher at Department of Communication, Media and Information Technologies, Aalborg University, Copenhagen, Denmark.
[5]Dept. of Information Technology, Techno International New Town, Kolkata, India.
[6]School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India.
[7]Faculty of Computers & Information, Information Technology Department, Cairo University, Egypt.

{[1]gr83gita@gmail.com,[2]asmitakalamkar@gmail.com,
[3,4]aalborg.pnm@gmail.com, [5]neelanjan.dey @gmail.com,
[6]jyotismita.c@gmail.com, [7]aboitcairo@gmail.com}

Communicating Author: *Gitanjali R. Shinde

Abstract:

Coronavirus (COVID-19) is a pandemic that has affected over 170 countries around the world. The number of infected and deceased patients has been increasing at an alarming rate in almost all the affected nations. Governments all over the world have been forced to take critical and tough decisions as step to contain the spread of the disease. Forecasting techniques can be inculcated thereby assisting the Government in designing better strategies and in taking productive decisions. These techniques assess the situations of the past thereby enabling better predictions about the situation to occur in the future. These predictions will help governments all over the world to prepare for the forthcoming situations. Forecasting techniques play a very important role in yielding accurate predictions. This study categorizes forecasting techniques into two types, namely, stochastic theory/mathematical models and data science/machine learning techniques. Data collected from various platforms

also play a vital role in forecasting. In this study two categories of datasets have been discussed i.e. big data accessed from WHO/National databases and data from a social media communication. Forecasting of a pandemic can be done based on various parameters such as the impact of environmental factors, incubation period, the impact of quarantine, age, gender and many more. These techniques and parameters used for forecasting are extensively studied in this work. However, forecasting techniques come with their own set of challenges (technical and generic). This study discusses these challenges and also provides a set of recommendations for the people and the governments that are currently fighting the global COVID-19 pandemic.

Keywords: COVID-19, Forecasting models, Machine Learning Method, Prediction, Big data, Epidemic, Pandemic

1. Introduction

The world has been facing threats in the form of pandemics periodically over the centuries. The aftermath of these pandemics have always had a huge impact on the world and have also turned the tables over. COVID-19, the current devastating pandemic is also running its course currently in the world. Not only economies are crashing but the overall strengths and morals of the heavily impacted nations are being compromised. The key to understanding the pandemic starts with an understanding of the disease itself, and the progression of the natural course of the disease. The disease is often defined as the state that negatively affects the body of a living person, plant or animal. The disease affects the body because of a pathogenic infection. The natural course of the disease starts before the onset of the infection. After which it progresses through the pre-symptomatic stage. The last stage is the clinical phase. In the clinical phase, a patient receives the prognosis of the disease. After a successful treatment of the disease, the patient enters into the remission stage. Remission refers to a decrease in the symptoms or a complete disappearance of the disease. The patient needs to follow instructions given by the doctor very strictly in the remission stage. This will ensure that the disease does not recur. If treatment is not successful a patient can be dead or chronically disabled. There are some important terms used to represent the statistics:

(i) Case-fatality rate: It is defined as the ratio of the number of patients who die due to the disease to the number of people who have it.

(ii) Observed survival rate: it is the prediction of the probability of survival.

(iii) Relative survival rate: It is defined as the ratio of the observed survival to the expected survival.

Without medical help, a person can die. A disease generally progresses because of the exposure to the infection. Because of this exposure to infection hosts are formed. Hosts refer to the group of people who are more susceptible to get affected. When an infected host comes in contact with more people then disease starts to spread. Figure1 depicts the host formation and progression [1].
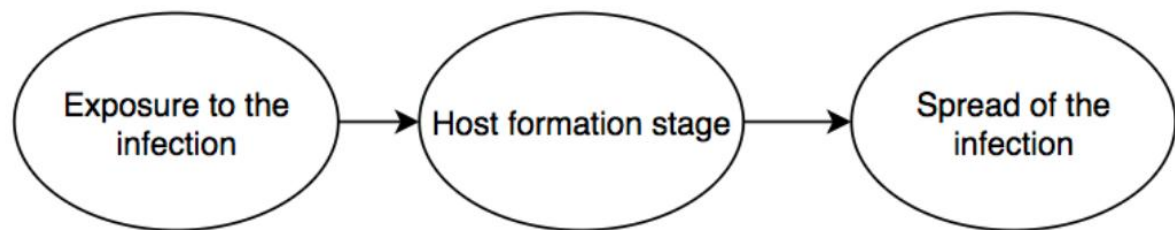

Figure 1: Host formation and progression

The diseases are mainly categorized into 2 parts:
  (i) Congenital diseases
  (ii) Acquired diseases

Congenital diseases exist in the body right from the birth. These diseases are generally activated through genetic disorders, environmental factors or a combination of these factors. These diseases are generally hereditary in nature i.e. passed through generations. For example hear conditions, down syndrome etc. In contrast to the former, acquired diseases spread through the living organisms. These are not hereditary in nature. These types of diseases are further classified into infectious and non-infectious diseases. As the name suggests infectious diseases spread through pathogenic agents (virus or bacteria or any microorganism). For example, SARS, SARS COVID-19 etc. Non-infectious diseases do not occur due to a pathogenic agent. For example, cancer, auto-immune disorder etc. A traditional model for the cause of the infectious disease is defined. It is called as an Epidemiologic Triad. It is depicted in figure 2.
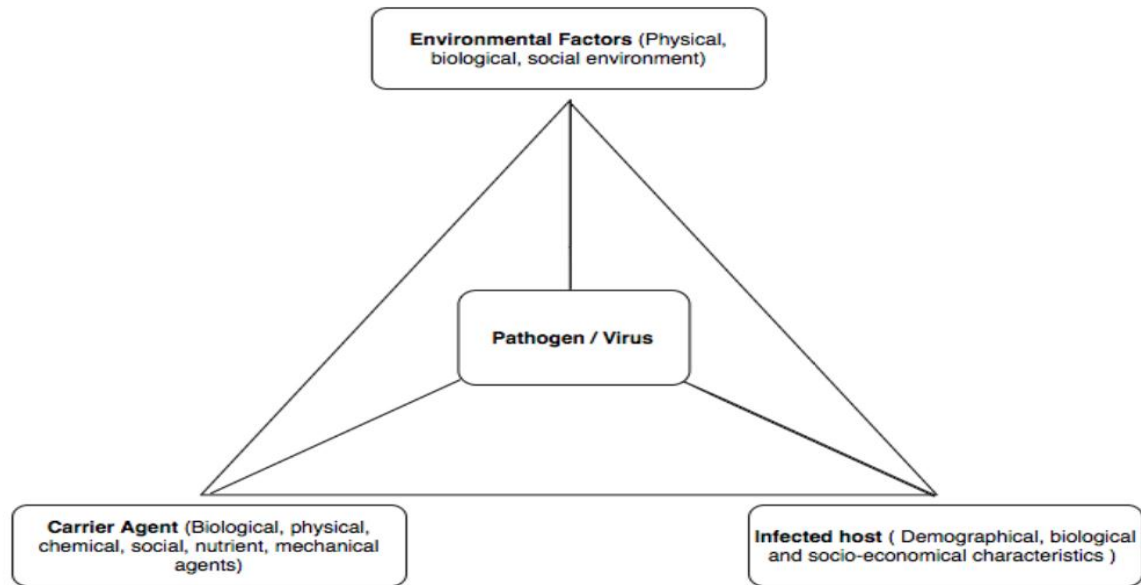
Figure 2: Epidemiologic Triad

The four important factors involved in the epidemiologic triad are environmental factors, carrier agent, infected hosts and the pathogens. The agent is usually the carrier of the infection. The infection is transmitted to the host when an agent comes in contact with the host under a certain environment. A pathogen is also known as a vector. A vector is an organism that transmits the infection via virus or bacteria from one host to another [2]. A disease takes the form of an epidemic when the following two conditions are met. First is when several people become affected by a similar nature of illness/disease that has the same root cause and second is when the number of infected people rapidly increases over the period. When this epidemic crosses the local boundaries and covers a wide geological area at the same time then it becomes a pandemic. A Pandemic does not give any information about the severity and impact of the disease. It merely states the fact that people across a wide geographical area are being infected with the disease. Pandemics are often referred to as outbreaks because of their spread pattern. The type of the outbreak determines the mortality rate of the disease. Over the last few years, it has been seen that because of the change in lifestyle, increased global travel and urbanization, infectious diseases quickly escalate into a pandemic. To prevent these epidemics, strong policies need to be administered. Otherwise, the situation can take a drastic turn rapidly. Since the beginning, mankind has faced epidemics and pandemics. The first epidemic faced by mankind was in the early 1300's called black death. It was one of the worst pandemics seen by humankind. This epidemic took millions of lives. It has been observed that this disease targeted most of the elderly people and people who are exposed to psychological

stressors [3,4]. The next pandemic faced by people was in the early 1500's called smallpox. 50% of the mortality rate was observed [5]. After which mankind had to face one of the deadliest pandemics called the fifth cholera pandemic which took more 1.5 million lives [6]. Following this, in 1918 one of the devastating Spanish flu influenza pandemics was observed. This pandemic took 20 million -110 million lives. In 1957 the Asian flu influenza pandemic occurred which took nearly 0.7-1.5 million lives [6,7]. In 1981 the world witnessed a new pandemic: HIV/AIDS. It was observed that more than 70 million patients were infected with the virus. According to WHO, Global health observatory data 36.7 million deaths occurred due to this pandemic [8,9]. After the HIV/AIDS pandemic, the world witnessed a new wave of different pandemics starting with SARS in 2003. This pandemic affected 4 continents and 37 countries across the globe [10,11]. In 2009 swine flu pandemic took place in which about 151,700–575,500 deaths were reported [12,13]. SARS pandemic was followed by the MERS pandemic in 2012. It affected 22 countries across the globe [14]. Two pandemics then followed the MERS. First was the Ebola pandemic in 2013 followed by the zika pandemic in 2015. Both the pandemics reported deaths in thousands [15,16]. Currently, the whole world is witnessing the coronavirus 19 (COVID-19) pandemic. More than 100 plus countries till date are majorly affected by COVID-19. This count is increasing as each passing day. Throughout the history of these epidemics, one thing was observed, that is, with the progress in time, these epidemics escalated into pandemics or many times referred to as the outbreak of the virus/disease. An epidemic escalates into a pandemic when the situation gets out of control at the local source where the outbreak was first observed to spread. The novelty of the disease and the uncertainty that prevails regarding the disease has lead to a lot of rumours regarding its whereabouts. People are unclear about the pre-clinical symptoms and the ways to handle it. Yet another important factor to consider is that lots of people who have preclinical symptoms do not reach the hospitals on time due to negligence or fear of testing positive for the disease. If somebody has the symptoms they have to act on it as soon as possible. This can help to save a lot of lives. If an early outbreak in any nation is successfully controlled then the situation can be prevented from escalating into a pandemic. Whenever these pandemic occur, world economies are majorly hit. Billions of dollars need to be invested in controlling an outbreak as well as in the development of a vaccine for the new disease [17]. While studying the outbreak or spread of any disease it is imminent to take all related factors into the account. J. Gaudart et al. [18] have taken extensions of the classical Ross-McKendrick-Mac Donald approaches. These approaches are combined with demographic and spatial dependencies of the virus on the host as well as the spread of disease. This

research discusses the retro prediction model to study the spread of the COVID-19. To predict the spread of the HIV/AIDS pandemic Kaplan's model was used in [19]. But the prediction focussed on drug addicts using injector/syringe. Hence the study was focused on the spread pattern pertaining to the specific group of people. MERS was another pandemic faced by the world. In order to analyze the transmission route of the MERS, decision tree and apriori algorithms were used in [20]. In [21] a maximum likelihood method was used to assess the spread of the SARS epidemic using the construction of phylogenetic tree. In [22] SVM was used to address the same issue. The neural forecasting model was used in [23] for obtaining a forecast for swine flu.

COVID-19 is a novel disease that has evolved into a pandemic. This novel disease has been reported by the World Health Organization (WHO) on 31st December 2019 in Wuhan, China. Soon after the outbreak in China lots of countries were in the grasp of COVID-19. According to WHO globally 634835 confirmed cases have been registered, 29891 deaths have been recorded till date [24]. Following is a list of the majorly hit countries grouped according to the region. In west pacific region along with China (Infected patient count-82482, Death count-3311), Republic of Korea (Infected patient count-9688, Death count-160), Australia (Infected patient count-4297, Death count-18), Malaysia (Infected patient count-2479, Death count-28) and Japan (Infected patient count-1887, Death count-55) were majorly hit. In the European region, Italy has the highest confirmed case count which is 98446. The number of deaths recorded in Italy is 10910. In the same region Spain (Infected patient count-80437, Death count-6441), Germany (Infected patient count-56512, Death count-453), France (Infected patient count-41748, Death count-2630), The United Kingdom (Infected patient count-19639, Death count-1279), Switzerland (Infected patient count-14200, Death count-273), etc. are the countries that have reported an alarming increase in the number of infected patients as well as the deceased. In South-East Asia region along with Thailand (Infected patient count-1640, Death count-9), Indonesia (Infected patient count-1264, Death count-117) and India (Infected patient count-1234, Death count-33) have reported multiple cases. In Eastern Mediterranean region Iran (Infected patient count- 38484, Death count-2656), Pakistan (Infected patient count-1817, Death count-17) and Saudi Arabia (Infected patient count-1302, Death count-5), were majorly affected. In the United States of America, 121414 people are currently infected with the number of deaths recorded as 2093. In the African region, South Africa (Infected patient count-1204, Death count-1) and Algeria (Infected patient count-451, Death count-27) have reported multiple cases. These statistics have been taken from WHO dated 29 March 2020 [24]. There are

numerous techniques from the field of statistics, data science, ML and AI that can be used for the purpose of prediction. The detailed study is presented in section three.

This study is organized into 4 main sections. The paper starts with the natural course of the disease; categorization of the diseases, along with the global history of pandemics where the COVID-19 outbreak is also mentioned. Section 2 provides an overview of the COVID-19 virus the different measures exercised by the government in confining the outbreak. Section 3 provides a survey of the multiple forecasting techniques and their categories. Section 4 deals with analysis, policies/recommendations for the control of the outbreak and the challenges that exists in the forecasting models.

2. Coronavirus overview

COVID-19 affects the respiratory system of the human body which is caused due to coronavirus-2. This virus is highly contagious. It is spreading through the bodily droplets in the air. Common symptoms include fever, tiredness, and dry cough. Along with these symptoms, a patient also experiences shortness of breath, aches and pains and sore throat. Very few people have experienced diarrhoea, nausea or a runny nose. People having high fever, cough or difficulty in breathing should call their doctor and seek medical help immediately. Human to human transmission is exponentially increasing the count of the infected people. The incubation period of this disease is 1- 14 days or even longer [24].When the COVID-19 started to spread at an unprecedented rate, preventive measures were exercised by the Chinese government. These measures included a complete lock-down of the heavily infected areas, ban on international travels, suspending schools and other non-essential daily activities. The main aims of these measures were to limit interpersonal contact, considering the contagious nature if the disease. The curfew imposed by the government was strictly observed. As the incubation period of the virus is longer than other viruses it is very difficult to analyze the optimal time required to observe a curfew. If the curfew is lifted too soon the situation can become dangerous. The people who get infected fall under three categories. First in the category are the elderly, who are highly susceptible to the virus. Statistics show that because of the weak immune system the elderly succumb to the disease easily. The second category is that of the children. As the immune systems of young children are still under development, the children are at higher risk. The third category is that of the people who have diseases like diabetes, high BP, asthma, cancer, cardiovascular disease, etc. As their immune systems have been compromised already due to a

prevailing medical condition, these people become easy targets. Infections experienced by the third category of people can be fatal [17].

3. Forecasting Techniques

In the literature, forecasting has been done based on various forecasting techniques and different data sources. To understand and improve the forecasting this section categorizes these techniques into multiple types for better analysis. This categorization is done based on the data sources used i.e. big data accessed from WHO/National databases and data from social media. Categorization is also done based on techniques that are used for forecasting i.e. data science/ machine learning techniques. However, there are also a few other categories that are used in the literature for forecasting. In nutshell, these categories are broadly divided into the following four sets:

    a. Big data
    b. Social media/other communication media data
    c. Stochastic theory/mathematical models
    d. Data science/Machine learning techniques

Various statistical, analytical, mathematical and medical (symptomatic and asymptomatic) parameters are taken into consideration for analysis. However, major significant parameters are listed below:

    a. Daily death count
    b. Number of carriers
    c. Incubation period
    d. Environmental parameters i.e. temperature, humidity, wind speed
    e. Awareness about COVID-19
    f. Medical facilities available
    g. Social distancing, quarantine, isolation
    h. Transmission rate
    i. Mobility
    j. Geographical location
    k. Age and Gender
    l. Highly and least vulnerable population
    m. Underlying disease
    n. Report time
    o. Government policies and many more

Apart from these above-mentioned parameters, there can be many influential factors which need to be further investigated.

3.1 Big Data

Effectiveness of forecasting is based upon the quality of data source used for forecasting. Forecasting results may vary based on the impurities in the data sources. Data mining and big data techniques always play a vital role in healthcare systems [25-28]. In the literature, researchers have done forecasting based upon data sources received from authenticated national and international sources. Here, analysis of big dataset is done by using various techniques like mathematical equations or machine learning techniques. Soumyabrata Bhattacharjee [29] has presented the impact of environmental factors like temperature, wind speed and humidity on the spread rate. This analysis is done based on the data accessed from WHO and the local weather database. Alexis Akira Toda [30] has presented decision-making schemes by analyzing the COVID-19 data of countries like China, Japan, Korea, European countries, and North America obtained from Johns Hopkins University. Diego Caccavo [31], Marlena M. Siwiaket al. [32], Bushra Zareieet al. [33],Pedro Teles [34] and Lucia Russo[35] have analyzed COVID-19 databases accessed from WHO, Italy national data and Johns Hopkins to predict the mortality rate. Pai Liu et al. [36] presented the impact of disease control interventions and traffic restrictions on the spread rate. The analysis has been done on the dataset retrieved from US Centers for Disease Control (CDC). S. Nadim et al. [37], Pear Hossain et al. [38], Tarcísio M. et al.[39], Marco Claudio Train et al. [40] have presented the importance of quarantine in order to reduce the spread rate of COVID-19.Giulia Giordano et al. [41] have presented the data analysis of Italy based on Italy's national data. As per Italy's official release, there are a total 27980 infected cases and 2158 deaths of people who were positive of Coronavirus. Looking at the effect of the Pandemic in Italy, Giulia Giordano has proposed the SIDARTHE Model that helps in redefining the reproduction number. This epidemic prediction model compares the infected density with the level of symptoms. Jia Wangping [42] has presented a study in which, COVID-19 data from Jan 22, 2020, to Mar 16, 2020, has been used in time series form for analysis. Extended susceptible-infected-removed (eSIR) model. The prediction has been estimated using the Markov Chain Monte Carlo method and results show that the reproductive number in Italy is 4.10 and 3.15 in Hunan. The anticipated endpoint in Italy would be April 25. Details of the literature evaluation are summarized in table 1.

Table 1. Evaluation of COVID-19 forecasting on Big Data

Sr. No	Title of the work	Studied City/ Country	Data Source	Parameters	Remark
1	Statistical investigation of the relationship between the spread of coronavirus disease (COVID-19) and environmental factors based on the study of four most affected places of China and five most affected places of Italy [29]	In China, Beijing, Chongqing, Shanghai, Wuhan	WHO, Local Weather Underground	Maximum relative humidity, maximum temperature, and highest wind speed	The relation between spread rate and environmental factors
2	Susceptible-Infected-Recovered (SIR) Dynamics of COVID-19 and Economic Impact [30]	China, Japan, Korea, European countries, and North America	Johns Hopkins University, GitHub repository	Transmission rate, Infection rate, and recovery rate	Decision making
3	Chinese and Italian COVID-19 outbreaks can be correctly described by a modified SIRD model [31]	China, Italy	WHO, open-source repository operated by Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE)	the average number of contacts per person per time, kinetic of recovery, kinetic of death, number of Susceptible, Infected, Recovered and the Death count	Forecasting numbers of COVID-19 patients
4	From a single host to global spread. The global mobility based modeling of the COVID-19	Wuhan, China	Johns Hopkins University of Medicine Coronavirus Resource	number of susceptible individuals; the number of infected	Mortality rate

	pandemic implies higher infection and lower detection rates than current estimates. [32]		center GitHub repository	individuals	
5	A model for COVID-19 prediction in Iran based on China parameters [33]	Iran	WHO	Time-dependent transmission rate, time-dependent recovery rate, and time-dependent mortality rate	prediction of the number of COVID patients in next month
6	Predicting the evolution of sars-COVID-2 in Portugal using an adapted sir model previously used in South Korea for the MERs outbreak [34]	Italy, Portugal	Italy national data	Number of susceptible, exposed, asymptomatic infected, mild-to-severe infected patients, hospitalized cases, recovered cases, and total population	Forecasting numbers of COVID-19 patients
7	Tracing DAY-ZERO and Forecasting the Fade out of the COVID-19 Outbreak in Lombardy, Italy: A Compartmental Modelling and Numerical Optimization Approach [35]	Italy	Italy national data	Total population, number of confirmed COVID patient, number of unreported infected persons, recovered persons, deaths counts	Forecasting numbers of COVID-19 patients
8	COVID-19 Progression Timeline and Effectiveness of Response-to-Spread Interventions across the United States	US	US Centers for Disease Control (CDC)	Disease control interventions and traffic restrictions	Impact of disease control interventions and traffic restrictions on spread rate

	[36]				
9	Short-term predictions and prevention strategies for COVID-2019: A model-based study [37]	China	WHO	Number of quarantined people, number of isolated people	To reduce the spread rate effective management of quarantined individuals is more effective than management of isolated individuals
10	The effects of border control and quarantine measures on the global spread of COVID-19 [38]	China	International Air Transport Association (IATA) database, Chinese Center for Disease Control and Prevention	Mobility, quarantine	Impact of border control and quarantine on the death count
11	Expected impact of COVID-19 outbreak in a major metropolitan area in Brazil [39]	Brazil	WHO	Number of Susceptible, Exposed, Infectious, Recovered	Policy-making for avoiding outbreak in metropolitan cities
12	Modeling the epidemic 2019-nCoV event in Italy: a preliminary note [40]	Italy	Italy national data	Contact rate, the transmission rate	Impact of quarantine on spread rate
13	A SIDARTHE Model of COVID-19 Epidemic in Italy [41]	Italy	Italy national data	Transmission rate, probability rate of detection, mortality rate, e rate of recovery,	Forecasting numbers of COVID-19 patients

14	Extended SIR prediction of the trend of the epidemic of COVID-19 in Italy and compared with Hunan, China [42]	Italy	Johns Hopkins University	Early detection and isolation of individuals with symptoms, traffic restrictions, medical tracking, and entry or exit screening	Impact of policies on spread rate

3.2 Social media data/other communication media data

In this digital era, social media communication and internet searches are the most easily accessible platforms that provide more information about COVID-19. The social media and web search correlate with the number of daily COVID cases. Keeping this in mind few researchers has taken datasets from Google, Baidu search engines [43,44] , mobile phones [45,46], newspapers [51] and various websites [47-49] like Github[50] over a particular duration of time. Analysis of these datasets is done by various techniques as discussed before i.e. machine learning techniques or mathematical equations/stochastic theory based on the parameters which were discussed earlier. Xiaolin Zhu et al. [46] have presented a spatially pandemic model for predicting the death count. This study aims to build a prediction model that will analyze the growth of the virus for the next one month considering the current dynamics of COVIO-19. Three different scenarios have been taken into consideration for the study which includes residents, residents with Wuhan travel history and residents affected as a result of local outbreak. The decay rate has also been introduced in the study to appreciate the efforts of different cities to alleviate the spread of the disease. Phone data has been used to collect the statistics of city-wise residents who had travelled back from Wuhan and the city-based model has been trained using the prevailing statistics and validated against the new cases as on February 11. The same model has been used to predict cases up to March 12, 2020, under the aforementioned three scenarios. The study predicted that the number of infections would be around 72172, 54348 and 149774 by March 12, 2020. The potential outcome of the study is a spatial model and its predictions will certainly help the government in optimizing the allocation of resources in each city during the next one month when the epidemic reaches a serious state of concern. Details of this analysis are summarized in table 2 as follows:

Table 2. Evaluation of COVID forecasting on social media Databases

Sr. No	Title of the work	Studied City/ Country	Data Source	Parameters	Remark
1	Propagation analysis and prediction of the COVID-19 [43]	China	Search engine and Tab Baidu	Number of basic regenerations, the incubation period and the average number of days of cure	Impact of future prediction and backward on the spread of COVID-19
2	A Cybernetics-based Dynamic Infection Model for Analyzing SARS-COV-2 Infection Stability and Predicting Uncontrollable Risks [44]	China	Baidu map big data	The incubation period, number of carriers, contact rate	Impact of medical facilities, Social responsibility, administrative responsibility on death count and spread rate
3	Assessing the spread risk of Wuhan novel coronavirus within and beyond China, January-April 2020: a travel network-based modeling study [45]	China	Mobile phone data, air passenger itinerary data, and case reports	Domestic and international travel	Impact of domestic and international travel on COVID global spread
4	Spatially Explicit Modeling of 2019-nCoV Epidemic Trend based on Mobile Phone Data in Mainland China [46]	China	Mobile phone	Spatially pandemic model, the Decay rate	Prediction of the death count
5	On a quarantine model of coronavirus infection and data analysis	China and Korea	https://www.worldometers.info/coronavirus/	Sub-population of latently infected individuals, the incubation	Quarantine is not sufficient and stricter measures are needed

	[47]			period	
6	Data-Based Analysis, Modelling and Forecasting of the COVID-19 outbreak [48]	China	https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6	Reproduction number	Forecasting of death numbers
7	Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020 [49]	China	Internet searches and social media data	Confirmed cases and suspected cases of COVID-19	Prediction of COVID-19 outbreak
8	The Impact of School Closure for COVID-19 on the US Healthcare Workforce and the Net Mortality Effects [50]	US	Github page, https://github.com/jbayham/us_childcare_obligations	Social Distancing (school closure)	Impact of social distancing on the death count
9	Modeling and predicting the Spatio-temporal spread of Coronavirus disease 2019 (COVID-19) in Italy [51]	Italy	websites of the main Italian newspapers	Time and space	Impact of space and time on decision-makers to intervene on the local policies

3.3 Stochastic theory/ Mathematical models

In a few past pandemics, the traditional approach of the mathematical and stochastic theory was used to estimate the loss of human and also to predict the total death count until a particular period or end of the pandemic. This traditional approach is

very effective and shows better predictions. Hence in the current pandemic situation of COVID-19 researchers [52-58] have used the same traditional approach for estimating the death count and the spread rate of COVID-19. The approach is also used to predict the total death count till the end of the pandemic. The analysis is done on databases accessed from authorized sources or search engines, mobile phone data and newspaper reports. Reza Sameni [59] has proposed a pattern of the virus with the help of mathematical modeling. This study uses a model from the family of the well-known compartmental models known as susceptible infected-recovered (SIR) model. A study has shown that the measures taken by the countries are positively affecting the mortality rate. Along with that, the facilities that are created to house the infected people, has contributed greatly in stopping the spread of the disease. However, this mathematical model has limitations in terms of accuracy because it is developed for the underlined dataset. Yuan et al. [60] presented the Boltzmann's function-based analysis. It has been observed that the prediction accuracy is better and it can also help governments to assess the severity of the situation and take appropriate actions. Jennifer Beam Dowd et al. [61] proposed the impact of age and gender on the death count using mathematical modelling. It has been observed that this virus is largely affecting the elderly. Now, in this case, the age structure of a particular country plays a vital role. In Italy, 23% of the population is above 65 years of age and hence the threat is maximized for the countries having similar age structure as that of Italy. The same situation can be faced by South Korea. Hence the policies like social distancing and quarantine can help to slow down and stop the spread of the virus. Xi He et al. [62] has presented the impact of pre-symptomatic transmission on the death count using mathematical modeling in this infector-infections and the transmission rate is studied. From the observation it was inferred that the rate of transmission was at its peak on or before the symptom onset. 44% of transmission can be seen even before the first symptoms become physically visible. Hence the disease control authorities should take the pre-symptomatic transmission into account while implementing the measure to curb the spread. Vasily Giannakeas et al. [63] presented an online tool for healthcare management using stochastic theory. Amitava Banerjee et al. [64] presented the impact of underlying conditions like heart disease and diabetes on the death rate. The impact of mobility on the spread rate of COVID19 is presented by Alexander F et al. [65]. Biqing Chen et al. [66], Yueling Maaet al.[67] and Peng Shi et al.[68] presented the impact of environmental factors on death count and spread rate of COVID-19. This analysis is based on the parameters listed as earlier and the details of this analysis are summarized in table 3.

Table 3. Evaluation of COVID forecasting based on the mathematical and stochastic theory

Sr. No	Title of the work	Studied City/ Country	Data Source	Parameters	Remark
1	A New, Simple Projection Model for COVID-19 Pandemic [52]	China, Italy, Iran, Germany, France, USA, South of Korea	WHO	Effectiveness of government intervention, public response, and healthcare system	Forecasting numbers of COVID-19 patients
2	Predicting the number of reported and unreported cases for the COVID-19 epidemic in South Korea, Italy, France, and Germany [53]	South Korea, Italy, France, and Germany	Korean Center for Disease Control, the Italian Ministry of Health, the French Public Agency of Health, and the Robert Koch Institute of Germany	Reported and unreported numbers	Forecasting number of COVID-19 cases
3	Mathematical predictions for COVID-19 as a global pandemic [54]	Global death count,129 countries	John Hopkins Hospital (JHH),WHO	Exposed and infected population	Prediction infection rate
4	Tracking and forecasting milepost moments of the epidemic in the early-outbreak: framework and applications to the COVID-19 [55]	China	Centers for Disease Control, China	Stages of spread	Prediction spread rate
5	A simple iterative map forecast of the COVID-19 pandemic [56]	China	WHO	Fitting parameter calculated from the total number of cases and new	Prediction spread rate

				cases each day	
6	Assessing the potential impact of COVID-19 in Brazil: Mobility, Morbidity and the burden on the Health Care System [57]	Brazil	Official Airline Guide (OAG)Data SUS	Outbreak probability, effective distance, Social vulnerability	Prediction of spread rate
7	Trend analysis of the COVID-19 pandemic in China and the rest of the world [58]	China, Italy, Japan and Germany	Johns Hopkins University	Duration, number of deaths	Epidemic spread pattern
8	Mathematical Modelling of Epidemic Diseases; A Case Study of the COVID-19 Coronavirus [59]	Global death count	WHO	Mortality rate, infection rate, re-infection rate, the recovery rate	Disease control measures: Lockdown, Social distancing
9	Forecasting the Cumulative Number of COVID-19 Deaths in China: a Boltzmann Function-based Modeling Study [60]	China	WHO	Number of deaths	Forecasting the Cumulative Number of COVID-19 Deaths
10	Demographic science aids in understanding the spread and fatality rates of COVID-19 [61]	Italy, Republic of Korea, Nigeria Brazil, USA, UK	population data from (https://population.un.org/wpp), death data taken from the most recent Italian data	Age and gender	Impact of age and gender
11	Temporal dynamics in viral shedding and transmissibility of COVID-19 [62]	China	(LOCAL) Guangzhou Eighth People's Hospital in	The incubation period, pre-symptomatic transmission,	Impact of incubation period on spread rate

			Guangdong, China	post-symptomatic transmission	
12	Estimating the maximum daily number of incident COVID-19 cases manageable by a healthcare system [63]	US	American Hospital Association, https://drive.google.com/drive/folders/1T5I2VKuvVFD0FmFGItcH_ZX1XioH4vYU	Type of bed acute bed, critical care bed, and ventilator	Online Tool for No of patients can be admitted in hospital healthcare system management
13	Estimating excess 1- year mortality from COVID-19 according to underlying conditions and age in England: a rapid analysis using NHS health records in 3.8 million adults [64]	UK	primary and secondary care electronic health records in England	Impact of Underlying conditions like heart disease, diabetic on mortality	Estimation of excess 1- year mortality from COVID-19 with underlying conditions
14	Eliminating COVID-19: A Community-based Analysis [65]	China	Chinese Center for Disease Control and Prevention	Mobility rate	Impact of travel/mobility of the infected community
15	Roles of meteorological conditions in COVID-19 transmission on a worldwide scale [66]	China, Italy, Japan	World Health Organization (WHO)	Air temperature, relative humidity, wind speed, and visibility	Single environment factor doesn't make any major impact, however, multi-factors can impact on spread rate
16	Effects of temperature variation and humidity on the mortality of COVID-19 in Wuhan [67]	China	Health Commission of Hubei Province, people's Republic of China,	Number of death, temperature, humidity	Environment factors like temperature and humidity may impact COVID-19 mortality rate

			Shanghai Meteorological Bureau and Data Center of Ministry of Ecology and The environment of the People's Republic of China		
17	The impact of temperature and absolute humidity on the coronavirus disease 2019 (COVID-19) outbreak - evidence from China [68]	China	China National Health Commission, meteorological authority in mainland China	Impact of temperature and absolute humidity on the coronavirus disease 2019	Lower and higher temperatures may be positive to decrease the COVID-19, there is no major impact of Humidity

3.4 Data Science/Machine Learning Techniques

Nowadays machine learning techniques are used worldwide for predictions due to its accuracy. However, to use machine learning (ML) techniques, there are a few challenges as very little data is available. For instance, the challenges involved in training a model are appropriate selection of parameters and the selection of the best ML model for prediction. Researchers have done predictions based on datasets that are available and used the best ML model as per the dataset [17, 69-72]. Jagadish Kumar and Hembram [73] presented a model based on the Logistic equation, Weibull equation, and the Hill equation to find infection rates in China and Italy. In this research work, data analysis is done to understand the effect of environmental factors on the spread of coronavirus disease. Data analysis is done on 4 cities in China namely Beijing, Chongqing, Shanghai and Wuhan and 5 cities of Italy namely Bergamo, Cremona, Lodi and Milano. The number of infected people is greater in the above mentioned cities. Three environmental factors are mainly focused in this study i.e. maximum environmental temperature, relative humidity, and wind speed. For data analysis, data is collected from a report published by WHO for China and Italy. Data is taken from the official GitHub repository of the Department of Civil Protection, Italy. The results show that there is a negligible relation between humidity and wind

speed with the spread of coronavirus. Similarly it has been observed that higher/maximum temperatures have a negligible to a moderate impact on spread of the virus. The result shows that there is no sign of any major effect of temperature on the virus. However results may vary depending on the dataset. Dave DeCapprio et al. [74] proposed a models using logistic regression, gradient boosted trees, and a hybrid model using Medicare data. The outcome of these models will help to initiate control strategies and to initiate corrective measures in time to control the spread. The details of this analysis are explained in table 4.

Table 4. Evaluation of COVID forecasting based on Data Science/Machine Learning Techniques

Sr. No	Title of the work	Studied City/ Country	Data Source	Parameters	Remark
1	Finding an Accurate Early Forecasting Model from Small Dataset: A Case of 2019-nCoV Novel Coronavirus Outbreak [17]	China	Small dataset, Chinese health authorities	Corrective feedbacks of model	Forecasting Suspected numbers of COVID-19
2	Composite Monte Carlo Decision Making under High Uncertainty of Novel Coronavirus Epidemic Using Hybridized Deep Learning and Fuzzy Rule Induction [69]	China	Chinese Center for Disease Control and Prevention	Cost of isolation, cost of treatment, no of suspects, no of confirmed COVID patients	Decision making
3	Estimation of the final size of the second phase of the coronavirus COVID-19 epidemic by the logistic model [70]	China	WHO	Daily death count	Forecasting of Short duration death count and the final total death count

4	Evaluating the effect of a public health intervention on the global-wide spread trajectory of COVID-19 [71]	Global death count,102 countries	WHO	Degree of intervention and starting intervention time	Impact of a public health intervention on the global-wide spread
5	Prediction and analysis of Coronavirus Disease 2019 [72]	China	2003 SARS data	Death count	Forecasting of death numbers
6	Epidemiological study of novel coronavirus (COVID-19) [73]	China, Italy, Iran, Spain, Germany, France, USA, Republic of Korea, Switzerland, and the UK	WHO	Infection rate	Prediction of infection rate
7	Building a COVID-19 Vulnerability Index [74]	Global Data	International Classification of Diseases	Pre-existing medical conditions	Identify individuals who are at the greatest risk

Tremendous work has been going on the COVID-19 apart from the above discussed work. Researchers are working to investigate efficient and accurate models in order to predict the death count. Researchers are also working to provide a list of guidelines that can be followed by the government and people to reduce the spread rate of the COVID-19.

3.5 Discussion

As stated earlier, the literature survey presented above is based on broadly four categories like the size of the dataset, source of the dataset, and techniques applied for forecasting like mathematical /analytical or machine learning/data science. This survey is carried out on various medical and non-medical parameters and it is very clear that the basic purpose of all these studies is to estimate the final size of this COVID -19 pandemic. However, it is very interesting to note that, all the studies have referred the China epidemic as the basis and all forecasts have been done based on the early statistics which are

available from the outbreak in China. Outcomes of these studies are very much useful for multiple purposes like controlling the spread of COVID-19 globally, controlling the spread of COVID-19 for a specific country, deciding its impact, building vulnerability index of COVID-19, establishing a correlation between environmental conditions (metrological conditions) and the spread rate, deciding reproduction number, establishing the correlation between quarantine and isolation with the spread of COVID-19, trend analysis of COVID-19 pandemic and tracking the spread of coronavirus locally and globally. The COVID-19 pandemic having been in existence for a very short period now, it is very important to analyze the trend of its spread and infected cases. All affected nations are looking towards mitigation plans to control the spread of the disease with the help of some modelling techniques. In the sequel, the outcomes of these forecasts are multi-fold. Every forecast is carried out with some perspective irrespective of which category it may represent. From these studies and the forecasts made, it is very clear that the major outcome is to support the government and healthcare communities to initiate critical action, decisions, control measures and public restrictions in time. Another outcome is to support the government in establishing mechanisms that provide control measures to be considered internationally for the global control of this pandemic as well as restrictions to the public in terms of quarantine, isolation, contact tracing, recommendation in terms of metrological conditions (mainly Air, Temperature, relative humidity, wind speed and visibility) and its impact on the spread. However, despite these useful outcomes, there are still many issues and challenges which are still unaddressed. The first and most important issue is whether the modelling and predictions based on China's dataset would suffice to address the issues of all countries. There is a need for reassessment to ensure that the control measures initiated by China to regulate the outbreak are enough to control this global pandemic. Many researchers have presented models for disease predictors to decide the reproductive number, but all of them have relied on similar datasets. It is also a crucial factor to rethink whether the same mathematical or prediction model is also suitable to predict the spread and reproduction number for all the countries across the globe. Literature shows that all the models presented are tested based on the numbers of the China epidemic. It is also equally important to ensure that the same model tested for the China's dataset can also be applicable to control the outbreak of COVID-19 globally. Another issue is that in the literature very limited details regarding the key characteristics of Coronavirus and the symptoms of COVID-19 are available. In this sequel, the challenge is to identify a vulnerable group of people with these limited details regarding viruses and disease. There is also a need to consider multiple peaks in the model not only for short term prediction but also to predict

the outbreak later in the year. Before confirming the forecasting mechanism, there is a need to reconsider these issues and challenges for better accuracy. There are few and important medical and non-medical parameters which still need to be investigated as evident from the literature. Few of which are, the genetic relations pertaining to the geographical locations need to be studied in to order to confirm the forecast. The Ethnicity (civilization, society, culture) of the infected people is another important parameter which needs to be reviewed. Correlation between the spread and its impact on a specific patient considering the underlined pre-existing medical complications is also another important parameter to be considered for more effective and accurate forecasting. It should be noted that not a single study or model available in the literature has considered the existing treatment options and has assumed that no vaccination option will be available for the next one year [75]. However, these are also some important parameters which need attention to fine-tune the model further.

4. Challenges of forecasting models

Forecasting plays an important role in every domain [76-78] due to its benefits to save resources or to improve the economy. However it comes with its challenges. In the case of COVID-19, there are also many challenges for forecasting the death count and spread rate as COIVD 19 incubation period is very much longer and very fewer datasets are available for the purpose. Few such challenges of forecasting models are listed as follows:

1) **Tracking of the people:** The tracking of infected personnel and other people who came in contact with them is truly one of the difficult tasks.
2) **No definitive treatment:** No definite treatment has been defined for the COVID-19 infection until now.
3) **The high fatality rate in susceptible people:** People with prior history of diseases are highly susceptible to the infection.
4) **Longer incubation period:** As COVID-19 has an incubation period of 14 days, it is impossible to identify patients beforehand. During the defined incubation period patients can infect all the people who come in contact with him/her.
5) **Lack of proper data:** Data accuracy is an important factor in achieving effective forecasting methods.

Along with these challenges, some more challenges are important to make a note of:

- **Proper lockdown:** It is very difficult for any country to implement a lockdown. To decide the proper conditions of a lockdown is a very complicated task.
- **The optimal period for lockdown:** The optimal period for lockdown is not only crucial but also a critical task.
- **Aware but do not cause panic:** It is important to educate people but in the process, it is important to remember not to create panic.
- **Essential services identification and delivery:** It is imminent for any country to identify essential services before lockdown. Even amongst lockdown lack of these services can cause a massive panic.
- **Crystal clear communication of messages:** The decisions taken by the government should be neatly explained to the public with proper justification. Without proper communication, the situation can become messy.

5. Recommendations

If an epidemic is controlled properly in the initial stage and if proper measures are taken to prevent it from crossing the geological boundaries then it could save a lot of lives with lesser impact. Forecasting and proper study of the pattern of disease spread could be very helpful in the planning of control strategies. At this stage, a complete lockdown imposed in the affected area is a good solution to prevent and hopefully stop the spread. But when the epidemic turns into pandemic it covers a larger geological area. This means tremendous growth in the number of infected people. This stage greatly impacts the nations that have been infected with the virus. When the virus is contagious, immediate lockdown in that particular area might save a lot of people from being infected. Awareness about the pandemic among the people is a very crucial part. If people know what the symptoms are and how to act if they have them, then it can help governments and doctors in the process. If people are afraid then they might not come forward and it can lead to a very disastrous situation. Furnishing proper information about the symptoms and treatments to the public might be helpful. Governments also need to keep tabs on the rumours that prevail around the disease. One rumour can turn the entire situation into chaos. As previously mentioned, the incubation period of this virus is too long as compared to other viruses. It is very important to keep track of the infected victims. Along with that, the people who came in contact with an infected person must be moved into the quarantine facilities. These quarantine facilities should be equipped with isolation units. The people who came in contact with

the infected person can be put into isolation until the incubation period is over. Following the incubation period testing should be performed on these patients. Additionally, the government needs to check the travel history and daily routine of the patient before he/she encountered the infection. This will help them in identifying the patient zero or the carrier agent. All the people who came into immediate contact with the infected patient needs to quarantine himself/herself. Home quarantine is also a good option if an isolated room with different latrine facilities are available. Whenever in-home quarantine care should be taken that the person remains within the confined limits of access until the test results turn out to be negative. If home quarantine facilities are not available, the governments should build such facilities. An early forecast of the upcoming situation may help governments work in a better way. If accurate predictions about the growth in several patients are done, then the government will be in an able position to handle the situation at the hand. Even if the prediction of forecasts suggests the worst-case scenario is about to take place the government can be well prepared. COVID-19 is highly contagious. Doctors, nursing staff and supporting staff should wear masks when treating the patients. Alcohol-based sanitizers should be used for sanitization. Also, if possible hazmat suites can be used by the doctors to treat patients who are showing severe symptoms. Governments around the world should provide masks and all necessary stock to the hospitals so that they can work effectively. Social distancing is one of the measures that can be implemented. It means people should maintain at least two meters distance between themselves. This can potentially stop the spreading of the disease.

In the face of this pandemic, most of the nations are in complete lockdown state. But in this lockdown governments should find alternative measures to deliver food, medications and other essential services to the people. A detailed area wise timetable for delivery as per request from people can potentially stop people from rushing for supplies and creating havoc. Also until the cities are in complete lockdown complete sanitization of the cities can be done by the governments. Sanitization process can start with public places and then other parts of the cities can be covered. It has been observed that the people who are having diseases like high BP, diabetes, asthma are more susceptible to the infection. Also, children and elderly people are at a higher risk. Governments should identify such people and keep track of their count. Finally, people should take really good care of personal hygiene. Frequent washing of hands, avoiding touching of face and eyes frequently, covering the mouth whenever sneezing or coughing, avoiding physical contact and drinking at least 3 litters of water daily are a few activities that can help maintain personal hygiene. People should

strictly follow the lockdown conditions imposed by the country or city. People should avoid stepping out of the house until it is very necessary. Avoiding air conditioners is a good practice as a controlled temperature can affect the health very easily. Lockdown potentially alters the lifestyle and routine of people. A complete lockdown can cause a massive panic. Entertainment via TV or other mediums like Netflix, Amazon prime, Hot star, etc can provide a little relief. The complete lockdown severely affects the economy of the world. Work from home policy can come in handy in unforeseen situations as these. Universities can provide students with online classes. Hence the academic loss can be contained. Also, online assessments can help in the process. Following are a few recommendations to stop the spread of the disease as soon as possible:

- Governments should take strict action against the people or organizations that violate the lockdown conditions without any convincing reason. All public transport services should be suspended except for the transportation of essential services and goods.
- All the places of worship should be closed for prayers. No religious congregations should take place in this period. All types of social gatherings (political, academic, cultural, etc) should be banned.
- The government should identify vacant facilities that can be turned into quarantine facilities. People with recent travel history to any foreign land/infected countries should be kept in strict isolation for the 14 days.
- The government should identify the services that come under the essential category. The personnel who are responsible for making essential services available to the public should be provided with passes for the ease of transportation. The export and import of the goods should be restricted.
- The government should take strict action against people who black-market the essential commodities and services. State-wide borders should be sealed off until the situation gets under control.
- The offices that come under essential services should instruct employees to follow social distancing while working. Also, appropriate sanitization facilities like hand wash, sanitizers, etc should be made available for employees.
- In the case of death, there should be a restriction on the number of people (not more than 12-15) allowed to attend the funeral.

Governments around the world should be ready. These are really difficult times but better preparedness can help avoid a lot of panic.

6. Conclusion

The COVID-19 pandemic is spreading its wings across the globe at a surprisingly faster rate and has already resulted in thousands of deaths across

countries. Unfortunately, this number is sure to grow within a short period and governments and healthcare organizations would soon face scarcity of resources. In this sequel, it is important to analyze various forecasting models for COVID-19 in order to empower governments and allied organizations with more appropriate information possible. An overall comprehensive study on analysis of COVID-19, its forecasting, impacts, and control measures are presented in this study. The major contribution of this study is the analysis of several forecasting models available in the literature and their classification, challenges of these models and recommendations to control this pandemic. Based on the available forecasting methods, we studied various statistical, analytical, mathematical and medical (symptomatic and asymptomatic) parameters. Also, common yet significant parameters have been taken into consideration which includes death count, metrological parameters, quarantine period, medical resources, mobility, etc. In this study, we have done the categorization of various forecasting methods into four major sets which include big datasets accessed from WHO/National data sources, social media/other communication media data, stochastic theory/mathematical models and data science/Machine learning techniques. This classification will surely help researchers to consolidate the forecasting methods more crisply and concisely as presented in this study.

Our study indicates that there is a need to reassess control measures initiated by China and other countries. Prediction of the spread and reproduction number should be analyzed on varied datasets. The models presented in the literature should be tested globally for more accurate global forecasting. On similar grounds, there is also a need to consider multiple peaks in the model not only for short term prediction but also to predict the outbreak later in the year. This study also indicates the challenges of various forecasting models and useful recommendations for the control of this pandemic.

We hope that by providing analysis of various forecasting models of COVID-19 to the government and healthcare community, it will be more helpful for adapting better intervention policies and explicitly, it will also help to alleviate the alarming effect of this pandemic. We agree that many of the papers referred to in this study for analysis are pre-print i.e. they do not peer review formally. However, due to the rapid growth of COVID-19 globally, there is a strong need for such a comprehensive survey as a contribution towards the society.

Conflict of interest statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1) Principles of Epidemiology in Public Health Practice, Third Edition
2) Text book of Epidemiology , Leon Gordis 4th Edition
3) Platt C. 2014. King Death: The Black Death and Its Aftermath in Late-Medieval England. Oxon U.K.: Routledge
4) DeWitte S N. 2014. "Mortality Risk and Survival in the Aftermath of the Medieval Black Death." PLoS One 9 (5): e96513.
5) Diamond J. 2009. Guns, Germs, and Steel: The Fates of Human Societies. New York: Norton.
6) Frieden N M. 1977. "The Russian Cholera Epidemic, 1892–93, and Medical Professionalization." Journal of Social History 10 (4): 538.
7) McKibbin W J, Sidorenko A A. 2006. "Global Macroeconomic Consequences of Pandemic Influenza." Analysis, Lowy Institute for International Policy, Sydney, Australia.
8) Dixon S, McDonald S, Roberts J. 2001. "AIDS and Economic Growth in Africa: A Panel Data Analysis." Journal of International Development 13 (4): 411–26.
9) Flahault A, Valleron A J. 1990. "HIV and Travel, No Rationale for Restrictions." The Lancet 336 (8724): 1197–98.
10) Keogh-Brown M R, Smith R D. 2008. "The Economic Impact of SARS: How Does the Reality Match the Predictions?" Health Policy 88 (1): 110–20.
11) Achonu C, Laporte A, Gardam M A. 2005. "The Financial Impact of Controlling a Respiratory Virus Outbreak in a Teaching Hospital: Lessons Learned from SARS." Canadian Journal of Public Health 96 (1): 52–54.
12) Tizzoni M, Bajardi P, Poletto C, Ramasco J J, Balcan D., and others. 2012. "Real-Time Numerical Forecast of Global Epidemic Spreading: Case Study of 2009 A/H1N1pdm." BMC Medicine 10 (1): 165.
13) Jain S, Kamimoto L, Bramley A M, Schmitz A M, Benoit S R., and others. 2009. "Hospitalized Patients with 2009 H1N1 Influenza in the United States, April–June 2015." New England Journal of Medicine 361 (20): 1935–44.
14) Park J, Kim J. 2015. "Hong Kong Sets 'Serious' Response to South Korea's MERS Outbreak." Reuters, June 8.

15) Jonas O B. 2013. "Pandemic Risk." Background paper for World Development Report 2014: Risk and Opportunity; Managing Risk for Development, World Bank, Washington, DC.

16) UNDP (United Nations Development Programme). 2017. "A Socio-Economic Impact Assessment of the Zika Virus in Latin America and the Caribbean: With a Focus on Brazil, Colombia, and Suriname." Synthesis report, UNDP, New York.

17) Fong, S. J., Li, G., Dey, N., Crespo, R. G., & Herrera-Viedma, E. (2020). Finding an accurate early forecasting model from small dataset: A case of 2019-ncov novel coronavirus outbreak. *arXiv preprint arXiv:2003.10776*.

18) Gaudart, J., Ghassani, M., Mintsa, J., Waku, J., Rachdi, M., Doumbo, O. K., &Demongeot, J. (2010, April). Demographic and spatial factors as causes of an epidemic spread, the copule approach: application to the retro-prediction of the black death epidemy of 1346. In *2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops* (pp. 751-758). IEEE.

19) Greenhalgh, D., & HAY, G. (1997). Mathematical modelling of the spread of HIV/AIDS amongst injecting drug users. *Mathematical Medicine and Biology: A Journal of the IMA*, *14*(1), 11-38.

20) Kim, D., Hong, S., Choi, S., & Yoon, T. (2016). Analysis of transmission route of MERS coronavirus using decision tree and Apriori algorithm. In *2016 18th International Conference on Advanced Communication Technology (ICACT)* (pp. 559-565). IEEE.

21) Amiroch, S., Pradana, M. S., Irawan, M. I., &Mukhlash, I. (2018, August). Maximum Likelihood Method on The Construction of Phylogenetic Tree for Identification the Spreading of SARS Epidemic. In *2018 International Symposium on Advanced Intelligent Informatics (SAIN)* (pp. 137-141). IEEE.

22) Hu, B., & Gong, J. (2010, August). Support vector machine based classification analysis of SARS spatial distribution. In *2010 Sixth International Conference on Natural Computation* (Vol. 2, pp. 924-927). IEEE.

23) Sultana, N., & Sharma, N. (2018, December). Statistical Models for Predicting Swine F1u Incidences in India. In *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)* (pp. 134-138). IEEE.

24) World Health Organization online available on https://www.who.int/emergencies/diseases/novel-coronavirus-2019/

25) Bhatt, C., Dey, N., &Ashour, A. S. (Eds.). (2017). Internet of things and big data technologies for next generation healthcare.

26) Hassanien, A. E., Dey, N., &Borra, S. (Eds.). (2018). Medical Big Data and internet of medical things: Advances, challenges and applications. CRC Press.

27) Lan, K., Wang, D. T., Fong, S., Liu, L. S., Wong, K. K., &Dey, N. (2018). A survey of data mining and deep learning in bioinformatics. Journal of medical systems, 42(8), 139.

28) Jain, A., &Bhatnagar, V. (2017). Concoction of ambient intelligence and big data for better patient ministration services. International Journal of Ambient Computing and Intelligence (IJACI), 8(4), 19-30.

29) SoumyabrataBhattacharjee," Statistical investigation of relationship between spread of coronavirus disease (COVID-19) and environmental factors based on study of four mostly affected places of China and five mostly affected places of Italy", 2020

30) Alexis Akira Toda," Susceptible-Infected-Recovered (SIR) Dynamics of COVID-19 and Economic Impact",2020#

31) Caccavo, D. (2020). Chinese and Italian COVID-19 outbreaks can be correctly described by a modified SIRD model. *medRxiv*.

32) Siwiak, M. M., Szczesny, P., &Siwiak, M. P. (2020). From a single host to global spread. The global mobility based modelling of the COVID-19 pandemic implies higher infection and lower detection rates than current estimates. *medRxiv*

33) Zareie, B., Roshani, A., Mansournia, M. A., Rasouli, M. A., &Moradi, G. (2020). A model for COVID-19 prediction in Iran based on China parameters. *medRxiv*.

34) Teles, P. (2020). Predicting the evolution Of SARS-Covid-2 in Portugal using an adapted SIR Model previously used in South Korea for the MERS outbreak. *arXiv preprint arXiv:2003.10047*

35) 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*

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

37) Nadim, S. S., Ghosh, I., &Chattopadhyay, J. (2020). Short-term predictions and prevention strategies for COVID-2019: A model based study. *arXiv preprint arXiv:2003.08150*

38) Hossain, M., Junus, A., Zhu, X., Jia, P., Wen, T. H., Pfeiffer, D., & Yuan, H. Y. (2020). The effects of border control and quarantine measures on global spread of COVID-19. *Alvin and Zhu, Xiaolin and Jia, Pengfei and Wen, Tzai-Hung and Pfeiffer, Dirk and Yuan, Hsiang-Yu, The Effects of Border Control and Quarantine Measures on Global Spread of COVID-19 (March 2, 2020).*

39) Rocha Filho, T. M., dos Santos, F. S. G., Gomes, V. B., Rocha, T. A., Croda, J. H., Ramalho, W. M., & Araujo, W. N. (2020). Expected impact of COVID-19 outbreak in a major metropolitan area in Brazil. *medRxiv*.

40) Traini, M. C., Caponi, C., & De Socio, G. V. (2020). Modelling the epidemic 2019-nCoV event in Italy: a preliminary note. *medRxiv*

41) Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., &Colaneri, M. (2020). A SIDARTHE Model of COVID-19 Epidemic in Italy. *arXiv preprint arXiv:2003.09861*

42) Jia, W., Han, K., Song, Y., Cao, W., Wang, S., Yang, S., ... & Liu, M. (2020). Extended SIR prediction of the epidemics trend of COVID-19 in Italy and compared with Hunan, China. *medRxiv*.

43) Li, L., Yang, Z., Dang, Z., Meng, C., Huang, J., Meng, H. T., ... & Peng, H. (2020). Propagation analysis and prediction of the COVID-19. *arXiv preprint arXiv:2003.06846*.

44) Xiao, W., Liu, Q., Huan, J., Sun, P., Wang, L., Zang, C., ... & Gao, L. (2020). A Cybernetics-based Dynamic Infection Model for Analyzing SARS-COV-2 Infection Stability and Predicting Uncontrollable Risks. *medRxiv*

45) Lai, S., Bogoch, I. I., Ruktanonchai, N., Watts, A. G., Li, Y., Yu, J., ... & Li, Z. (2020). Assessing spread risk of Wuhan novel coronavirus within and beyond China, January-April 2020: a travel network-based modelling study

46) Zhu, X., Zhang, A., Xu, S., Jia, P., Tan, X., Tian, J., ... & Yu, J. (2020). Spatially Explicit Modeling of 2019-nCoV Epidemic Trend based on Mobile Phone Data in Mainland China. *medRxiv*.

47) Volpert, V., Banerjee, M., &Petrovskii, S. (2020). On a quarantine model of coronavirus infection and data analysis. *Mathematical Modelling of Natural Phenomena*, *15*, 24.

48) Anastassopoulou, C., Russo, L., Tsakris, A., &Siettos, C. (2020). Data-Based Analysis, Modelling and Forecasting of the novel Coronavirus (2019-nCoV) outbreak. *medRxiv*

49) Li, C., Chen, L. J., Chen, X., Zhang, M., Pang, C. P., & Chen, H. (2020). Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020. *Eurosurveillance*, *25*(10).

50) Bayham, J., &Fenichel, E. P. (2020). The Impact of School Closure for COVID-19 on the US Healthcare Workforce and the Net Mortality Effects. *medRxiv*.

51) Giuliani, D., Dickson, M. M., Espa, G., & Santi, F. (2020). Modelling and predicting the spread of Coronavirus (COVID-19) infection in NUTS-3 Italian regions. *arXiv preprint arXiv:2003.06664*.

52) Lu, J. (2020). A New, Simple Projection Model for COVID-19 Pandemic. *medRxiv*

53) Webb, G. (2020). Predicting the number of reported and unreported cases for the COVID-19 epidemic in South Korea, Italy, France and Germany. *medRxiv*

54) Victor, A. O. (2020). MATHEMATICAL PREDICTIONS FOR COVID-19 AS A GLOBAL PANDEMIC. *medRxiv*.

55))Wang, H., Zhang, Y., Lu, S., & Wang, S. (2020). Tracking and forecasting milepost moments of the epidemic in the early-outbreak: framework and applications to the COVID-19. *medRxiv*.

56) Botha, A. E., &Dednam, W. (2020). A simple iterative map forecast of the COVID-19 pandemic. *arXiv preprint arXiv:2003.10532*.

57) Coelho, F. C., Lana, R. M., Cruz, O. G., Codeco, C. T., Villela, D., Bastos, L. S., ... & Gomes, M. F. (2020). Assessing the potential impacts of COVID-19 in Brasil: Mobility, Morbidity and Impact to the Health System. *medRxiv*.

58) Weber, A., Ianelli, F., &Goncalves, S. (2020). Trend analysis of the COVID-19 pandemic in China and the rest of the world. *arXiv preprint arXiv:2003.09032*.

59) Reza Sameni, Mathematical Modeling of Epidemic Diseases; A Case Study of the COVID-19 Coronavirus

60) Long, C., Ying, Q., Fu, X., Li, Z., & Gao, Y. (2020). Forecasting the Cumulative Number of COVID-19 Deaths in China: a Boltzmann Function-based Modeling Study. *medRxiv*.

61) Dowd, J. B., Rotondi, V., Adriano, L., Brazel, D. M., Block, P., Ding, X., ... & Mills, M. C. (2020). Demographic science aids in understanding the spread and fatality rates of COVID-19. *medRxiv*.

62) He, X., Lau, E. H., Wu, P., Deng, X., Wang, J., Hao, X., ... & Mo, X. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *medRxiv*.

63) Vasily Giannakeas, Deepit Bhatia ; Matthew T. Warkentin ; Isaac I. Bogoch ; Nathan M. Stall, Estimating the maximum daily number of incident COVID-19 cases manageable by a healthcare system.

64) Amitava Banerjee et al., Estimating excess 1- year mortality from COVID-19 according to underlying conditions and age in England: a rapid analysis using NHS health records in 3.8 million adults

65) Siegenfeld, A. F., & Bar-Yam, Y. (2020). Eliminating COVID-19: A Community-based Analysis. *arXiv preprint arXiv:2003.10086*.

66) Chen, B., Liang, H., Yuan, X., Hu, Y., Xu, M., Zhao, Y., ... & Zhu, X. (2020). Roles of meteorological conditions in COVID-19 transmission on a worldwide scale. *medRxiv*.

67) Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., ... & Luo, B. (2020). Effects of temperature variation and humidity on the mortality of COVID-19 in Wuhan. *medRxiv*.

68) Shi, P., Dong, Y., Yan, H., Li, X., Zhao, C., Liu, W., ... & Xi, S. (2020). The impact of temperature and absolute humidity on the coronavirus disease 2019 (COVID-19) outbreak-evidence from China. *medRxiv*.

69) Fong, S. J., Li, G., Dey, N., Crespo, R. G., & Herrera-Viedma, E. (2020). Composite Monte Carlo Decision Making under High Uncertainty of Novel Coronavirus Epidemic Using Hybridized Deep Learning and Fuzzy Rule Induction. *arXiv preprint arXiv:2003.09868*.

70)   Batista, M. Estimation of the final size of the second phase of the coronavirus COVID-19 epidemic by the logistic model.

71)   Hu, Z., Ge, Q., Li, S., Jin, L., &Xiong, M. (2020). Evaluating the effect of public health intervention on the global-wide spread trajectory of Covid-19. *medRxiv*.

72)   Jia, L., Li, K., Jiang, Y., &Guo, X. (2020). Prediction and analysis of Coronavirus Disease 2019. *arXiv preprint arXiv:2003.05447*.

73)   Jagadish Kumar, K. P. S. S. Hembram, Epidemiological study of novel coronavirus (COVID-19)

74)   DeCaprio, D., Gartner, J., Burgess, T., Kothari, S., & Sayed, S. (2020). Building a COVID-19 Vulnerability Index. *arXiv preprint arXiv:2003.07347*

75)   COVID-19 in India: Guidance from the IndiaSIM Model- March 24, 2020. [https://cddep.org/covid-19/]

76)   Dey, N., Fong, S., Song, W., & Cho, K. (2017, August). Forecasting energy consumption from smart home sensor network by deep learning. In International Conference on Smart Trends for Information Technology and Computer Communications (pp. 255-265). Springer, Singapore.

77)   Hu, S., Liu, M., Fong, S., Song, W., Dey, N., & Wong, R. (2018). Forecasting China future MNP by deep learning. In Behavior engineering and applications (pp. 169-210). Springer, Cham.

78)   Singh, N., & Mohanty, S. R. (2018). Short Term Price Forecasting Using Adaptive Generalized Neuron Model. International Journal of Ambient Computing and Intelligence (IJACI), 9(3), 44-56.