Chaos, Solitons and Fractals 152 (2021) 111340

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Leveraging weather data for forecasting cases-to-mortality rates due to COVID-19

Ogechukwu Iloanusi[a,*], Arun Ross[b]

[a] Department of Electronic Engineering, University of Nigeria, Nsukka 410001, Enugu State, Nigeria
[b] Michigan State University, East Lansing, MI 48824 USA

ARTICLE INFO

Article history:
Received 21 January 2021
Revised 10 August 2021
Accepted 11 August 2021
Available online 18 August 2021

Keywords:
COVID-19
COVID-19 cases-to-mortality ratios
Regression analysis
Forecasting
Weather conditions
Temperature
Solar irradiation
Rainfall
Relative humidity
Deep learning
Random forest

ABSTRACT

There are several recent publications criticizing the failure of COVID-19 forecasting models, with swinging over predictions and underpredictions, which have made it difficult for decision and policy making. Observing the failures of several COVID-19 forecasting models and the alarming spread of the virus, we seek to use some stable response for forecasting COVID-19, viz., ratios of COVID-19 cases to mortalities, rather than COVID-19 cases or fatalities. A trend of low COVID-19 cases-to-mortality ratios calls for urgent attention: the need for vaccines, for instance. Studies have shown that there are influences of weather parameters on COVID-19; and COVID-19 may have come to stay and could manifest a seasonal outbreak profile similar to other infectious respiratory diseases. In this paper, the influences of some weather, geographical, economic and demographic covariates were evaluated on COVID-19 response based on a series of Granger-causality tests. The effect of four weather parameters, viz., temperature, rainfall, solar irradiation and relative humidity, on daily COVID-19 cases-to-mortality ratios of 36 countries from 5 continents of the world were determined through regression analysis. Regression studies show that these four weather factors impact ratios of COVID-19 cases-to-mortality differently. The most impactful factor is temperature which is positively correlated with COVID-19 cases-to-mortality responses in 24 out of 36 countries. Temperature minimally affects COVID-19 cases-to-mortality ratios in the tropical countries. The most influential weather factor – temperature – was incorporated in training random forest and deep learning models for forecasting the cases-to-mortality rate of COVID-19 in clusters of countries in the world with similar weather conditions. Evaluation of trained forecasting models incorporating temperature features show better performance compared to a similar set of models trained without temperature features. This implies that COVID-19 forecasting models will predict more accurately if temperature features are factored in, especially for temperate countries.

1. Introduction

COVID-19 is a virus that initiated a challenging pandemic throughout the whole world in 2020. The impact of COVID-19 is known to be devastating and can lead to several other illnesses. Though COVID-19 is believed to have originated in China and by the first week of February 2020 was mainly confined to Wuhan, in Hubei, China, it however led to the death of over 390,000 individuals, globally, by the end of May 2020 [1]. Even though cities like Milan in Italy, initiated schemes for total lockdown by the third week of February 2020, most countries had to institute a nationwide lockdown to contain this virus by March 2020.

The global confirmed cases and deaths were 9,927 and 213, respectively, by January 31, 2020, however, it is alarming to note that over 103 Million individuals had contracted this virus and over 2.3 million persons died from it by January, 31 2021 [1]. From all indications, the current spread and death toll were never expected, and the situation could have been better. The fear about the virus is the high mortality versus confirmed cases. With respect to a total of 6,188,259 cases, over 391,271 had died from the virus by May 31, 2020 [1]. This is roughly a situation of 1 death per 16 cases. Absolutely, everyone fears a repeat of the same trend of mortality rate due to COVID-19, as experienced within the three months (March to May 2020).

One of the challenges of this virus is that it took people and nations unawares. If it were possible to forecast the spread of the virus, the impact probably would not have been this bad. Nations and cities could have intensified lockdowns; individuals could have had access to more medical aid; nations would have known

* Corresponding author.
E-mail addresses: Ogechukwu.iloanusi@unn.edu.ng (O. Iloanusi), rossarun@cse.msu.edu (A. Ross).

https://doi.org/10.1016/j.chaos.2021.111340
0960-0779/

Nomenclature

C:M	cases-to-mortality ratios/response

the strengths and limitations of the virus, and most importantly, focused on affected locations that require immediate attention. Hence, individuals and nations can be alerted, and they can take measures to minimize the calamity due to COVID-19.

It appears the benefits of COVID-19 forecasting models have not yet outweighed the costs. There are several recent publications hammering on the failure of COVID-19 forecasting models [2–5] which have been misleading in decision and policy making. There are cases of drastic over predictions and underpredictions. A first premise for this research is that over a year's experience in COVID-19 pandemic reveals that the COVID-19 spread rate is influenced by more physical interaction of individuals and ease of lockdowns, irrespective of the weather conditions [6–9]. Secondly, it has also been reported that COVID-19 virus might be taken in our stride and could manifest a seasonal outbreak profile similar to other infectious respiratory diseases, therefore, with increased morbidity, virulence and mortality with weather variations [10]. Hence, in this paper, we do not solely study the influence of weather conditions on the COVID-19 cases or fatalities, rather, we study their influence on COVID-19 cases-to-mortality ratios.

Our objectives are, firstly, to study the influence of some demographic, geographical, economic and four weather parameters on COVID-19 responses, and secondly, incorporate the most influential factor(s) in training models for forecasting the cases-to-mortality rate of COVID-19 in clusters of countries in the world with similar weather conditions.

The strengths of our study include the following:

1) Data taking all seasons into consideration, from January 22, 2020 to June 15, 2021, as opposed to data from just a season, were used. The reliance of AI on data for proper prediction or forecasting of COVID-19 cannot be over emphasized [11].
2) Lessons learnt from earlier publications on predictions and forecasting of COVID-19 were considered.
3) The influence (if any) of climatic parameters, some demographic, geographical, and economic factors on the COVID-19 responses were studied across 36 countries with varying geographical and economic conditions based on Granger-causality tests.
4) The impact of COVID-19 cases-to-mortality responses on predictors, viz. temperature, humidity, rainfall and solar irradiance, were analyzed and established using relationship equations via regression analysis.
5) The influential weather factor(s) were incorporated as features in training random forest and deep learning models for forecasting COVID-19 cases-to-mortality rate in clusters of countries with similar climatic conditions.

2. Review of related work

Several traditional approaches based on switching Kalman filters, Gaussian regression, Susceptible-Infected-Removed (SIR) models, have been proposed for forecasting the spread, confirmed and recovery cases due to COVID-19 virus in [12–14]. A susceptible–exposed–infected–removed (SEIR) model was trained for forecasting the spread of COVID-19 in China using data acquired between January 10 to February 15, 2020 for training, and February 16 to March 3, 2020 for testing, in [15]. In [16], polynomial representation techniques were applied on USA and UK COVID-19 data for forecasting the spread of COVID-19 cases between March 10 and April 3, 2020. COVID-19 spread in Saudi-Arabia was forecasted using an autoregressive integrated moving average (ARIMA) model in [17]. Fractal theory and Fuzzy logic were combined for a 10-day forecasting of COVID-19 spread (April 16 – April 25, 2020) in 10 countries from time-series data of January 22 2020 to April 15, 2020 in [18]. In [19], a SIR model was proposed for forecasting the spread of COVID-19 in four countries - South Korea, Iran, Italy and India. The spread of COVID-19 in six African countries, namely, Nigeria, Algeria, South Africa, Egypt, Kenya and Senegal, was studied in [20] using a modified SEIR model, on data acquired between February 14 and April 14, 2020. The forecast was done for March 21 to May 21 2020. A logistic growth model was used for a 6-day forecasting of the daily cumulative spread of COVID-19 in the US using data from January 22 to April 6, 2020 in [21].

Machine and deep learning approaches have been utilized in forecasting the spread of COVID-19. Several support vector regression methods were used to forecast the spread of COVID-19 in Brazil in [22]. In [23], a hybrid Artificial Intelligence model comprising a long short-term memory (LSTM) module and a Natural Language Processing (NLP) module, was proposed for predicting the infection rate and spread of COVID-19 in all cities in China. A forecast for the number of persons at risk of contracting COVID-19 in England and Wales was conducted in [24] based on census and hospital capacity data in order to estimate the hospitalization rates at the Ceremonial County and Clinical Commissioning group levels. A comparative review of deep learning models for forecasting the spread of COVID-19 time-series data was done in [25]. In [26] a neural network model was used for predicting the total cumulative spread of COVID-19 globally. In [27], the ecological niche model (ENM) was used for forecasting zones in the risk of spreading COVID-19 virus in three megacities in China, namely, Beijing, Guangzhou, and Shenzhen, based on data acquired between January 21 to February 27, 2020. A deep learning model was proposed for predicting the spread of COVID-19 in the USA in [28]. A time-series forecasting based on deep learning was done in [29]. Linear regression and support vector machine models were trained for a 10-day prediction of total new cases, recoveries, and deaths due to COVID-19 worldwide in [30]. A scenario-driven forecasting of COVID-19 was done for Belgium in [31].

Most of these models utilized data obtained from January to April 2020 to train models for forecasting COVID-19 spread. Going by the predictions of these approaches and comparing them against the real outcomes of COVID-19 from May 2020 till June 2021, the actual values are far from the predicted values in many cases; therefore, we learn that there are other factors that affect the spread that should really be taken into consideration when predicting the spread of COVID-19. A key factor is the effect of temperature on COVID-19, whose effect has been considerably researched though not adequately utilized in time series forecast of COVID-19.

The effects of temperature, evaporation, precipitation, and regional climate on COVID-19 in 31 States in Mexico were studied in [32]. The statistical analysis shows that the number of cases decreases with increase in temperature, leading to a delay in COVID-19 spread in tropical climates, for the data studied between February 29 to March 31 2020. A study of the effect of temperature on the spread of COVID-19 was carried out on data collected within January 22 to March 31 2020, on 10 most affected cities in China. The study involved a statistical analysis using Sim & Zhou' quantile-on-quantile (QQ) approach. In a nut-shell, a correlation coefficient of values within −0.3038 to 0.6612 were obtained in the ten cities studied. In [33], a study of the effect of temperature and humidity on the 50 States of the USA was conducted using data acquired within January 1 and April 9, 2020. There was more focus on humidity and the effects were measured by comparing the standard deviations from the monthly mean absolute humidifies in

the 50 USA States. The results were used in predicting COVID-19 spread in India from April to December 2020. In [34], the effect of temperature, rainfall, sunlight, humidity and windspeed on COVID-19 were evaluated in Rio de Janeiro, Brazil, for COVID-19 data acquired between March 12 and April 28, 2020. The results relating these weather factors to COVID-19 based on correlation coefficient, yielded a negative correlation for solar radiation (-0.609), followed by windspeed (-0.440), temperature (-0.406) and interestingly, a positive correlation coefficient of 0.265 for humidity and 0.127 for rainfall. The study in [35] showed that high temperatures stifle the spread of COVID-19 in newly infected individuals. The study in China [36] using data acquired within January and February 2020 showed that there is a relationship between COVID-19 and temperature and solar irradiation. The study in [37] showed a positive correlation of 0.347 between COVID-19 and temperature for data acquired within February 27 and May 2, 2020. In [10], studies on the influence of weather parameters on COVID-19 were reviewed. It was observed that COVID-19 may have come to stay and could manifest a seasonal outbreak profile similar to other infectious respiratory diseases.

From the review done in this work and the close observance of COVID-19 data for sixteen (16) months, it is evident that the effects of weather parameters - temperature or humidity - on COVID-19 *cases* could differ from the effects of temperature and humidity on *mortalities* due to COVID-19. A positive correlation coefficient between temperature or humidity and COVID-19 cases resulted from the studies in [37,38]. In some instances, this positive correlation is attributed to no lockdown states, frequent movements in the summer, etc. However, a negative correlation has mostly resulted in studies relating COVID-19 mortalities to temperature or humidity in [32,35,39], probably due to the devastating nature of the virus in the cold season.

3. Data preparation

Data used for the experiment include COVID-19 cases and deaths data for 36 countries acquired from John Hopkins University's Center for Systems Science and Engineering (JHU CSSE) data repository provided in [1]. Weather data used for the experiments include temperature, relative humidity, rainfall and solar irradiance. Data on world demographics and COVID-19 global reports were acquired from [40]. Average daily temperature ($^{\circ}C$), rainfall data (mm), solar irradiance in ($kw\text{-}hr/m^2/day$) and relative humidity data (%) were acquired for 36 countries from the National Aeronautics and Space Administration's (NASA) solar and meteorological data sets viewer [41]. The four weather variables from [41] comprise daily readings taken from January 1, 2020 to June 8, 2021. COVID-19 cases and deaths data for 36 countries contain daily readings from January 22, 2020 to June 15, 2021.

4. A stable response parameter for COVID-19 spread

One of the goals of this paper is to train models that can forecast the rate of COVID-19 pandemic in clusters of countries in the world with similar weather and pandemic conditions. Many countries experienced an increase in the COVID-19 cases with the lifting of lockdowns in the summer of 2020. Plots of COVID-19 daily cases and daily mortalities with time (for the period from January 22, 2020 till June 15, 2021) for the five (5) major continents of the world based on COVID-19 data available from [40] are presented in Fig. 1(a) and (b).

A 15-term moving averager was convolved with each time-series data to obtain the smooth curves in Fig. 1(a) and (b). Let $\boldsymbol{y}$ be the number of days in the COVID-19 time-series set. Averaged or convoluted responses for each country were obtained by convolving each country's set of cases or mortalities data, respectively, with a 15-term, 1 stride, moving averager. The convoluted response, $r[n]$, for a given input $\boldsymbol{x[n]}$, can be expressed as:

$$r[n] = \frac{1}{15} \sum_{\substack{k=0 \\ n=0}}^{\substack{y-15 \\ 14}} x[n-k] \tag{1}$$

here, $\boldsymbol{k}$ is the number of delay terms. Responses in the form of: $\{r[n] \mid n = 0, \ldots (y-15)\}$, were obtained for the cases and mortality series, defined as, $\{c_1, c_2, \ldots c_n\}$, and, $\{m_1, m_2, \ldots m_n\}$, respectively.

Plots of daily cases with time in the five (5) continents - Africa, Asia, Europe, North America and South America are shown in Fig. 1(a). There are pronounced peaks for Asia, Europe and North America in the months November 2020 to January 2021 and March to April 2021. In 2021, cases peaked for Asia, within April to May 2021, as a result of the delta COVID-19 variant that is believed to have originated in India.

In Fig. 1(b) however, there are pronounced peaks within April to May 2020 and November 2020 to January 2021 for North America. Europe peaked within April to May 2020 and November 2020 to April 2021. It is noteworthy that the daily mortality curves dropped in the summer with the lifting of lockdowns in 2020 in Europe and North America, though there were rises in the daily cases. Africa peaked within July to August 2020, and January 2021. South America peaked from July to September 2020; and within April to May in 2021. Evidently, the mortality curves are not proportionate to the cases' curves for the five continents as the year progresses, meaning that there are significant factors affecting mortality due to COVID-19. Mortalities followed a bathtub curve for North America and Europe, especially, in 2020, while daily cases kept on rising with time. This makes the use of cases or mortalities as responses difficult for COVID-19 forecast.

Series of box plots for the daily cases and mortalities per country are shown for 36 countries in Fig. 1(c) and (d), respectively, for a period from January 22, 2020 through June 15, 2021. The box plots for the daily cases and mortalities show that the magnitude, range and average of the daily cases and mortalities vary per country.

Using number of COVID-19 daily cases or daily mortalities as responses for prediction for a cluster of countries with varying population pose a problem as mortality is a function of the cases which are in turn function of a country's population, thereby making them unstable responses. Also, as the mortalities are dependent on the number of cases, it is ideal to take both the cases and mortalities into consideration, hence we do this by dividing the cases by mortalities. This is termed the case-to-mortality ($\boldsymbol{C:M}$) ratios. This serves as a common metric for comparing the impact of COVID-19 amongst nations with varying demographics. We define the COVID-19 cases-to-mortality $\boldsymbol{C:M}$ response as the ratio of the daily number of COVID-19 cases $\boldsymbol{c}$ to number of daily mortalities $\boldsymbol{m}$ delayed by 7 days. $\boldsymbol{C:M}$ is similar to case fatality rate used in epidemiology and has been studied in [42,43]. In determining the $\boldsymbol{C:M}$ ratios, the mortalities series were offset by 7 days from the cases series, to factor in a delayed effect of cases on mortalities. A 7-days delay was used because mortality peaks around 7 days later than the cases as observed from the acquired data. However, we evaluated the best delayed effect in this Section by varying delays of (a) 7 (b) 14 (c) 21 and (d) 28 days cases on mortality ratios. We demonstrate this by showing how these four different delays manifest in $\boldsymbol{C:M}$ plots for two major continents – Europe and North America. The results are shown in Fig. 2(a) for Europe and in Fig. 2(b) for North America. A glance at the curves in Fig. 2(a)

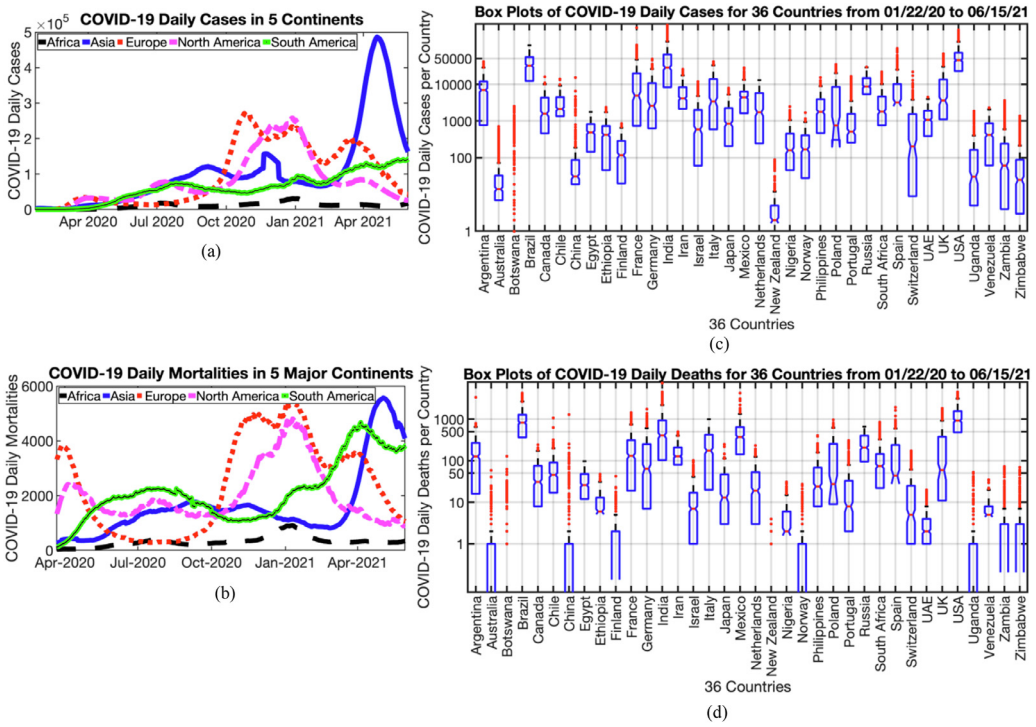

Fig. 1. Plots of COVID-19 (a) Daily cases with time (b) Daily mortalities with time in 5 Continents of the world. Box plots of (c) Daily cases (d) Daily deaths for 36 countries from January 22, 2020 to June 15, 2021.

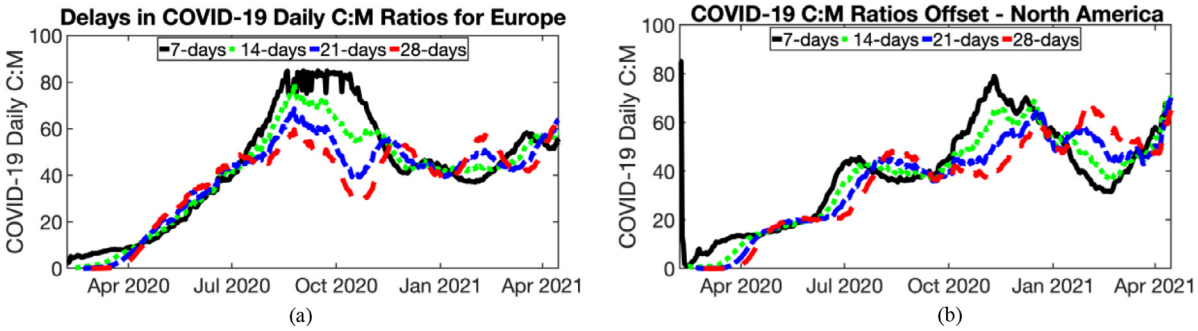

Fig. 2. Effect of varying delays of (a) 7 (b) 14 (c) 21 and (d) 28 days cases-to-mortality ratios.

and (b) shows that the peaks and troughs increase in magnitude as the delays decrease from 28 to 7 days. Therefore, the magnitude of the curve with 7-days delay in both plots is more impactful than other delays, and is chosen for the rest of the experiments. It was also reported that mortality peaks a week after COVID-19 cases peak in [42], therefore this pre-analysis conforms with the results in [42]. Hence, for a convoluted cases' series, $\{c_1, c_2, \ldots c_k, \ldots c_n\}$ and mortality series, $\{m_1, m_2, \ldots m_k, \ldots m_n\}$, the ***C:M*** ratios are $\{{}^{c_1}/_{m_8}, \ldots {}^{c_k}/_{m_{k+7}}, \ldots {}^{c_{n-7}}/_{m_n}\}$. The resulting series were filtered of outliers by consistently decrementing values exceeding a ceiling value of 90 in steps of 5. Plots of the ***C:M*** ratios for the 5 continents are shown in Fig. 3(a). The ***C:M*** ratios peaked around the summer time for Europe and North America and plunged in the winter times.

A special case of India's COVID-19 situation is briefly examined in this Section by comparing India's cases, mortalities and ***C:M*** ratios with those of five other countries, Egypt, Mexico, Poland, South Africa and the USA in Fig. 3(b)–(d), respectively. COVID-19 cases and mortalities are seen to rise significantly from April 2021 and peak within the month of May 2021 in India in Fig. 3(b) and (c). Irrespective of this COVID-19 variant in India, the ***C:M*** ratios in Fig. 3(d) have been steady between 70 and 85 from months October 2020 till May 2021, showing that the mortalities have been increasing in proportion to the cases. Hence, despite the rising cases and mortalities, the cases have not been fatal, in the midst of the Indian COVID-19 variant. ***C:M*** ratios dropped afterward as a result of decreasing cases. The cases in Mexico have been fatal with low ***C:M*** ratios (below 20) throughout all the seasons.

It can be observed that rise and fall patterns in most of the curves in Figs. 1–3, appear to be seasonal. Furthermore, the ***C:M*** ratios plots in Figs. 2, 3(a) and (d), tend to peak after about a month's delayed effect during the hottest periods in summer and dip after about a month's delayed effect during the coldest periods in the winter.

Series of box plots for the daily ***C:M*** ratios for 36 countries are shown Fig. 3(e). There are clusters of countries that appear to have comparable or similar averages of daily ***C:M*** ratios, for instance, Argentina and Brazil; Canada and Portugal; thereby making this metric a stable response for predicting the rate of COVID-19 spread in some clusters of countries with similar ***C:M*** ratios.

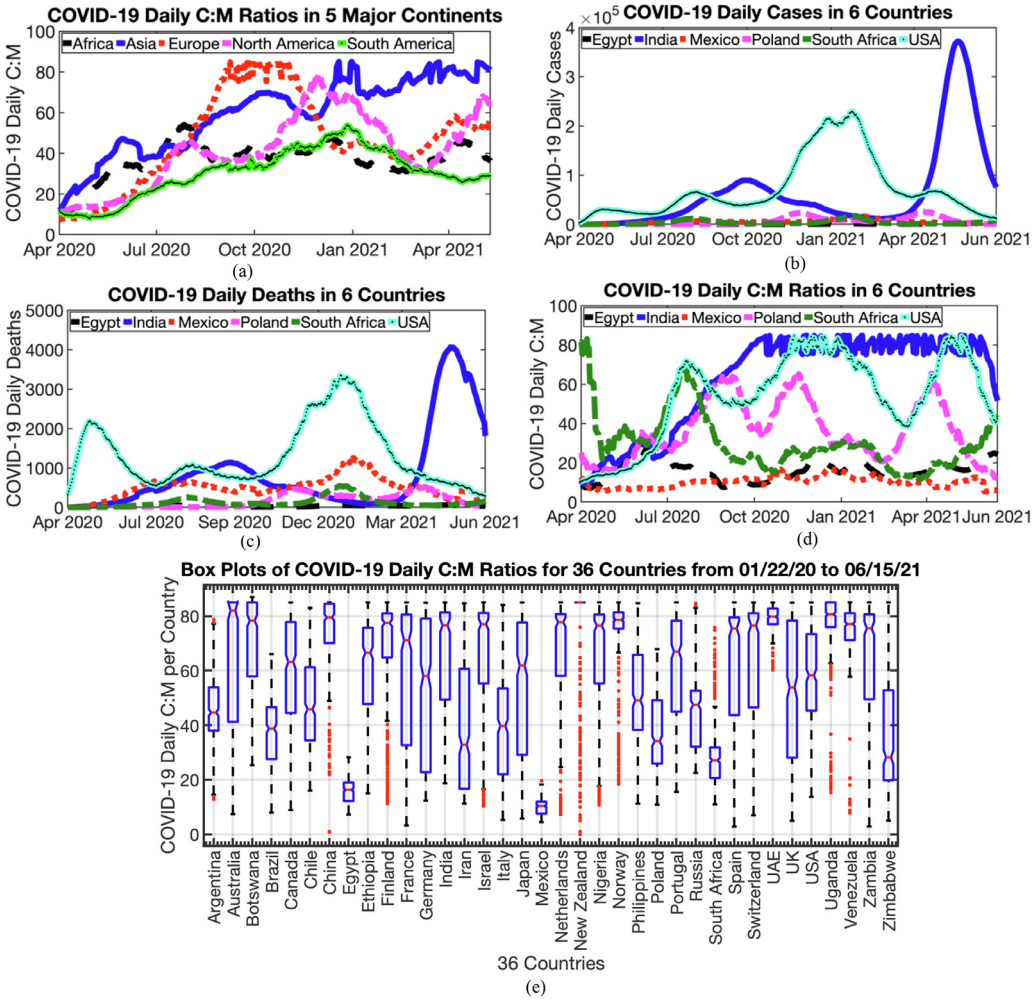

Fig. 3. Plots of COVID-19 (a) daily cases-to-mortality ratios (***C:M***) in 5 continents of the world and daily (b) cases (c) deaths and (d) ***C:M*** in six (6) countries. (e) Box plots of daily ***C:M*** for 36 countries from January 22, 2020 to June 15, 2021.

5. Problem statement and objectives of the paper

The goals of this paper are to evaluate if COVID-19 cases or mortalities are influenced by some demographic, geographical, economic and climatic factors in each country and to consider the most influential factor in forecasting new COVID-19 cases-to-mortality (***C:M***) rates per country. Factors that could affect the COVID-19 cases or mortality in a country include but are not limited to:

- Climatic and geographical parameters, viz., temperature, rainfall, windspeed, solar irradiation, relative humidity, approximate geographical location of a country
- Economic and demographic factors, viz., country's population and population density, median age per country, average age of deaths per country, total tests, number of tests per 1 million inhabitants, effectiveness of lockdown, health amenities available, country's GDP per capita.
- Social factors, viz., COVID-19 awareness and cooperation of inhabitants, health sector response.

It must be noted that this study does not take into account vaccination schedules for countries since the study was done when vaccines were not vastly prevalent and the delta strain was not in play.

Existing works have studied the effects of the weather parameters, viz., humidity, temperature, solar irradiation, rainfall, and wind speed, on COVID-19 pandemic. A key factor is the effect of temperature on COVID-19, whose effect has been considerably researched though not adequately utilized in time series forecast of COVID-19.

There are few publications that have measured the impact of the economic and social factors on the COVID-19 pandemic. In [44–46], the authors studied the adequate or inadequate health sector response to COVID-19 in specific countries or worldwide, In [47,48], the authors studied the successes, challenges and results of the effectiveness of lockdown. These factors are rarely quantified in existing works.

Also, there are recent publications highlighting the failure of COVID-19 forecasting models, with drastic underpredictions and over predictions, which have been misleading in policy making.

The objectives of the paper are to:

1) Determine if and how some covariates arising from the COVID-19 pandemic affect COVID-19 cases, mortality and cases-to-mortality (***C:M***) response by conducting a series of Granger-causality tests. Other covariates or factors not studied in this paper are possible areas for future work.
2) Evaluate the impact of the four weather parameters, viz., temperature, rainfall, solar irradiation and relative humidity, on COVID-19 cases, mortality and cases-to-mortality (***C:M***) response by conducting a series of Granger-causality tests on data from 36 countries with varying weather conditions.

Table 1
Deep learning and random forest architectures. Size of input layer depends on the number of features, ***f* = 1 or 2**.

Deep Learning Model		**Random Forest Model**	
Layers	**Attributes**	**Layers**	**Attributes**
Input layer*	***f***-D feature input size	Input layer*	***f***-D feature input size
LSTM layer	100 hidden units	Regression Trees	200
Dropout layer	50% dropout layer		
LSTM layer	150 hidden units		
Dropout layer	50% dropout layer		
Fully connected layer	1D output size		
Regression output	Mean-squared-error	Regression output	Mean-squared-error

3) Establish the relationship of the four weather parameters (predictors) on COVID-19 cases-to-mortalities (***C:M***) ratios (response) in 36 countries via regression analysis.
4) Factor in the most impactful parameter(s) in the training of models for forecasting COVID-19 ***C:M*** response. Temperature data were used along with the ***C:M*** as input to some time-series based forecasting models.
5) Group countries according to their similar COVID-19 impact and climatic conditions in order to train a number of models, corresponding to the number of country groups, for the forecasting of COVID-19 ***C:M*** response.
6) Train models for COVID-19 forecasting using two approaches – Random forest and Deep learning. Test models for COVID-19 forecasting using new data for each country.

6. Proposed technique

6.1. Covariates' analysis

The influence of some demographic, geographical, economic, and climatic covariates or factors on COVID-19 cases or mortality in each country or across 36 countries were studied in this paper. The effects of these covariates on each country's COVID-19 response data and across 36 countries were examined using the Granger-causality tests for defined null hypothesis. The demographic covariates are country's population, population density, median age; geographical factors include each country's average longitude and latitude information; economic covariates are GDP per Capita, total COVID-19 tests carried out and tests per 1 million population. The four (4) weather factors are temperature, rainfall, solar irradiation and relative humidity.

The null hypothesis for the demographic H_0^d, geographical H_0^g, economic H_0^e, and weather H_0^w factors, are stated as follows:

H_0^d: A series of demographic data of 36 countries does not affect countries' yearly response of COVID-19 cases, mortalities, and cases-to-mortalities ratios.
H_0^g: A series of geographical location of 36 countries does not affect countries' yearly response of COVID-19 cases, mortalities, and cases-to-mortalities ratios.
H_0^e: A series of economic covariates of 36 countries does not affect countries' yearly response of COVID-19 cases, mortalities, and cases-to-mortalities ratios.
H_0^w: A country's weather data over time does not affect the country's daily response of COVID-19 cases c, mortalities m, and cases-to-mortalities (***C:M***) ratios with time.

The null hypothesis was tested using the Granger-causality test [49]. The p-value, at 5% significance level (0.05), is used to accept or reject the null hypothesis, H_0. When the p-value < 0.05, H_0 is rejected.

6.2. Regression analysis

Predictors for regression analysis constitute temperature, rainfall, solar irradiance and relative humidity. 144 sets ($36 \times 4 = 144$) of regression analysis of ***C:M*** responses on measurements of rainfall, temperature, solar irradiation and relative humidity from 36 countries were carried out.

Regression analysis was adopted in order to quantify the impact and relationship (if any) between the predictors and responses. Previous works focused mostly on correlation analysis, which determines to what extent weather conditions impact COVID-19 cases; however, it is desirable to additionally determine, the relationship between the weather conditions and COVID-19 cases or mortalities and establish the relationship. This is most effective using regression analysis.

6.3. Two models' architectures for COVID-19 forecasting

Two methods were explored for training models for forecasting in this paper: (1) Growing random forests and (2) Deep learning. The architectures of the two models used in this paper are outlined in Table 1. Each random forest model was grown with 200 regression trees. Random Forest employs bootstrap aggregation in combining learning outputs from all the trees. Long Short-Term Memory (LSTM) networks were used as building blocks in the deep learning architecture since the COVID-19 data is a time-series data. The LSTM networks have Sigmoid gate activation function. LSTM networks were used in forecasting of COVID-19 in [50,51]. The input layer has an input size of 1 or 2 depending on the feature size, ***f***.

6.4. Considerations on weather and COVID-19 data for regression analysis and forecasting

A causal and effect approach was followed in determining the impact of these four weather factors on ***C:M*** ratios. Hence, in this analysis, the following were factored-in:

1) Most countries first instance of confirming their first case of COVID-19 (between the last week of January and early February 2020).
2) Initial stage of actively containing the virus in various countries (in February and March 2020).
3) Inability to determine or confirm how many persons contracted the virus in the early stages of the COVID-19 infection, resulting in erroneous recording of COVID-19 cases and mortalities [52,53], which varies by country (in February and March 2020).
4) The fact that if weather conditions, such as temperature, impact and furthers the spread of COVID-19, the infection will actually take root in a person after 2 to 14 days (2 weeks) of exposure to the virus, i.e., the incubation period [54–56].
5) The fact that some patients in critical conditions die between 1 to 4 weeks after contracting the virus [57–59].

Table 2
Granger causality H_0^d, H_0^e, H_0^g, tests of 36 countries' COVID-19 daily average cases, mortalities and ***C:M*** responses on demographic, economic and geographical covariates. P-values are provided.

	H_0^d **(Demographic)**		H_0^e **(Economic)**				H_0^g **(Geographical)**	
Covariates (8) Responses (3)	Population	P Density / Km^2	Median age	Total tests	Total tests / 1M	GDP / Capita	Latitude	Longitude
Cases	0.593	0.280	0.524	0.545	0.241	0.631	0.705	0.501
Mortalities	0.799	0.134	0.551	0.793	0.381	0.751	0.760	0.988
C:M ratios	**0.001**	0.106	0.937	0.373	**0.031**	**0.030**	0.335	0.109

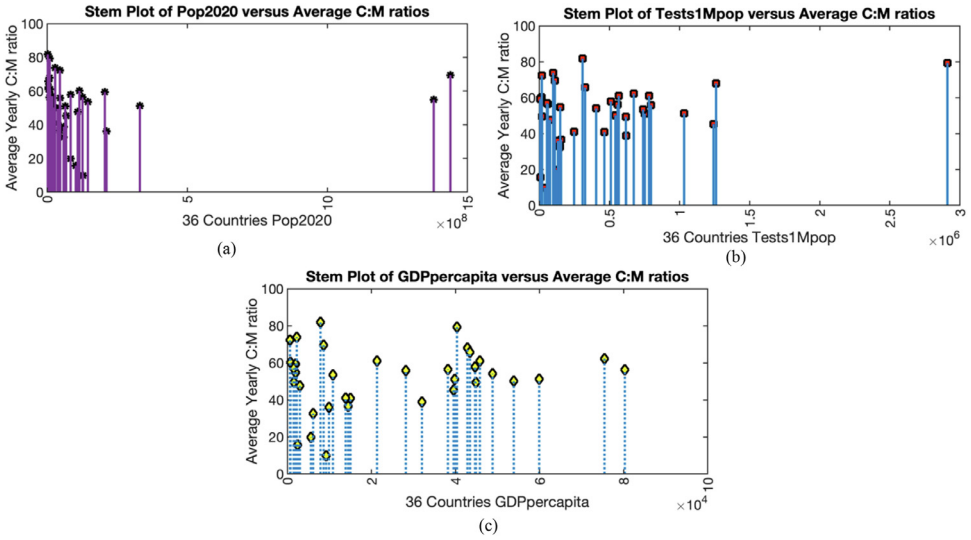

Fig. 4. Stem plots of 36 countries' (a) population in 2020 (b) Tests per 1M and (c) GDP per capita by 02/20/21 versus (***C:M***) ratios.

Therefore, forecasting was based on a cause and delayed-effect approach that took all factors above into consideration. More specifically, the weather parameters were offset by 4 weeks with respect to the response. Further, COVID-19 data between January 22 and March 31, 2020 were not used in this analysis.

7. Experiments

Three major experiments in this work constitute a series of Granger-causality tests on covariates; regression analysis for determining the impact of weather data on the COVID-19 responses for 36 countries; and training of models for forecasting the COVID-19 cases-to-mortality in clusters of countries with similar impact from weather conditions, determined through regression analysis.

This work was carried out in a MATLAB 2020a environment on a Macintosh Operating System with 16GB of RAM.

7.1. Experiments on covariate analysis

Granger-causality tests were carried out to test the demographic H_0^d, geographical H_0^g, economic H_0^e and weather H_0^w factors as described in Section 6.1. In the case of H_0^d, H_0^g, and H_0^e tests, the yearly average cases, mortalities and ***C:M*** responses were used for each country, since the covariate data of each country is singular. In the H_0^w tests, all daily cases, mortality and ***C:M*** time-series data from January 22, 2020 through June 15, 2021 were used since the weather data comprises daily average measurements. H_0s were rejected for all p-values < 0.05.

7.2. Experiments on regression analysis

C:M responses from 36 countries were regressed on the temperature, solar irradiation, rainfall and relative humidity data for each country, with emphasis on these parameters: regression equation, sum of squares due to regression, and error and the statistical p-value of each regression line's slope. The regression analysis was based on fortnightly grouping of the weather predictors and COVID-19 response data.

7.3. Experiments on COVID-19 forecasting

Training and testing data comprise COVID-19 and temperature data from February 1 to November 30, 2020, and December 1, 2020 to June 15, 2021, respectively. $r[n]$ responses were calculated from temperature data of the 36 countries. The COVID-19 data is a time-series data and, hence, the input and target data only differ in a time-step. The most impactful weather condition from the regression analysis was factored in; the 2D input features comprise the COVID-19 and temperature data. The output data comprises the COVID-19 data delayed by a time-step. All input data for each country were standardized to have a mean of zero and a standard deviation of 1.

Countries were grouped according to their (1) similar trend of ***C:M*** ratios and (2) similar weather condition, and following this order of priority. We determined the similarity in the trend of ***C:M*** ratios by normalizing all ***C:M*** ratio-series per country to values between 0 and 10 and computing an L2-norm similarity score from a difference matrix of all countries' ***C:M*** series. Given a minimum difference score, m_s, for each country's (say $\boldsymbol{c}$) comparison with other 35 countries, a search for countries with similarity scores below a threshold, $\{\tau \mid \tau = 1.2m_s\}$, selects countries similar to $\boldsymbol{c}$. It is remarkable that majority of the countries with similar weather conditions appear to have similar trends in ***C:M*** ratios. Hence, countries were grouped according to their similarity scores. Countries with noisy data were avoided in the forecasting experiments. The selected groups are as follows:

Table 3
Granger causality H_0^w tests and covariate analysis of cases, mortalities and ***C:M*** responses on measurements of rainfall, relative humidity, solar irradiation and temperature from 36 countries. P-values are provided.

	Response	**Cases**				**Mortalities**				**Case-to-Mortality ratios**			
	Country	Rain	Humid	Solar	Temp	Rain	Humid	Solar	Temp	Rain	Humid	Solar	Temp
1	Argentina	0.466	0.851	0.125	0.011	0.350	0.064	0.017	0.000	0.552	0.053	0.695	0.261
2	Australia	0.681	0.399	0.755	0.304	0.490	0.003	0.001	0.000	0.351	0.598	0.754	0.807
3	Botswana	0.483	0.330	0.723	0.101	0.448	0.054	0.703	0.014	0.873	0.185	0.201	0.021
4	Brazil	0.000	0.095	0.701	0.001	0.003	0.923	0.712	0.001	0.367	0.067	0.559	0.270
5	Canada	0.899	0.247	0.645	0.674	0.900	0.242	0.600	0.004	0.001	0.368	0.218	0.001
6	Chile	0.743	0.831	0.068	0.367	0.069	0.000	0.000	0.002	0.857	0.714	0.828	0.572
7	China	0.492	0.741	0.973	0.662	0.616	0.538	0.348	0.064	0.676	0.357	0.651	0.987
8	Egypt	0.771	0.565	0.815	0.205	0.389	0.900	0.696	0.579	0.970	0.122	0.296	0.001
9	Ethiopia	0.338	0.019	0.587	0.208	0.001	0.000	0.000	0.002	0.167	0.910	0.493	0.506
10	Finland	0.494	0.347	0.051	0.652	0.674	0.014	0.009	0.000	0.497	0.303	0.526	0.313
11	France	0.858	0.012	0.983	0.049	0.093	0.403	0.099	0.030	0.238	0.520	0.452	0.838
12	Germany	0.727	0.406	0.810	0.452	0.734	0.084	0.001	0.002	0.882	0.893	0.407	0.014
13	Iran	0.509	0.725	0.323	0.121	0.452	0.101	0.908	0.005	0.149	0.000	0.016	0.000
14	India	0.598	0.533	0.795	0.786	0.704	0.621	0.443	0.559	0.901	0.039	0.092	0.004
15	Israel	0.542	0.224	0.056	0.011	0.472	0.286	0.091	0.000	0.697	0.870	0.249	0.917
16	Italy	0.407	0.736	0.868	0.233	0.686	0.330	0.031	0.188	0.869	0.369	0.460	0.123
17	Japan	0.045	0.163	0.897	0.025	0.223	0.627	0.241	0.403	0.761	0.816	0.649	0.901
18	Mexico	0.132	0.216	0.064	0.000	0.427	0.866	0.007	0.000	0.011	0.009	0.015	0.031
19	Netherlands	0.430	0.403	0.545	0.087	0.085	0.102	0.005	0.079	0.939	0.160	0.446	0.181
20	New Zealand	0.902	0.133	0.244	0.057	0.091	0.076	0.065	0.001	0.790	0.457	0.625	0.377
21	Nigeria	0.498	0.068	0.785	0.311	0.170	0.040	0.267	0.001	0.117	0.621	0.098	0.261
22	Norway	0.966	0.146	0.013	0.651	0.713	0.002	0.002	0.000	0.650	0.114	0.993	0.089
23	Philippines	0.652	0.577	0.247	0.015	0.358	0.598	0.310	0.009	0.957	0.133	0.218	0.360
24	Poland	0.552	0.884	0.986	0.151	0.154	0.741	0.711	0.069	0.059	0.694	0.045	0.000
25	Portugal	0.322	0.032	0.585	0.088	0.594	0.688	0.113	0.638	0.046	0.000	0.000	0.000
26	Russia	0.190	0.194	0.118	0.487	0.993	0.743	0.373	0.013	0.080	0.000	0.000	0.000
27	South Africa	0.780	0.466	0.207	0.996	0.988	0.001	0.000	0.000	0.079	0.032	0.595	0.045
28	Spain	0.552	0.000	0.059	0.000	0.713	0.056	0.005	0.002	0.750	0.670	0.649	0.545
29	Switzerland	0.900	0.311	0.465	0.090	0.789	0.059	0.001	0.028	0.667	0.433	0.025	0.230
30	UAE	0.887	0.817	0.276	0.327	0.039	0.471	0.859	0.154	0.064	0.532	0.001	0.028
31	UK	0.067	0.257	0.138	0.150	0.303	0.801	0.399	0.001	0.870	0.309	0.854	0.306
32	USA	0.291	0.818	0.760	0.363	0.126	0.008	0.002	0.036	0.310	0.395	0.147	0.003
33	Uganda	0.584	0.289	0.176	0.160	0.097	0.203	0.448	0.181	0.023	0.824	0.270	0.000
34	Venezuela	0.041	0.002	0.683	0.008	0.338	0.000	0.003	0.000	0.743	0.139	0.546	0.026
35	Zambia	0.137	0.007	0.047	0.000	0.288	0.227	0.089	0.000	0.437	0.969	0.025	0.001
36	Zimbabwe	0.160	0.672	0.000	0.000	0.263	0.202	0.093	0.000	0.490	0.864	0.162	0.339
	H_0^w rejected	3	6	3	11	3	9	15	26	3	9	8	16

- Group 1 - USA
- Group 2 - UK
- Group 3 - Portugal, Canada, Ethiopia
- Group 4 - Italy, Poland
- Group 5 - Argentina, Brazil
- Group 6 - Russia, South Africa
- Group 7 - Iran, Nigeria
- Group 8 - Mexico
- Group 9 - Egypt
- Group 10 - Botswana, Uganda
- Group 11 - Switzerland
- Group 12 - Germany, Japan

In order to mark the improvements in trained forecast models due to incorporating a weather parameter, we trained models with and without temperature data features, using the same architecture, data and training parameters. There is only a difference in the input layer size, which is 2 in the case of COVID-19 and temperature data input, and 1 in the case of COVID-19 data input only. Each set of random forest and deep learning models had twelve (12) groups of models trained with (a) COVID-19 and temperature data and another twelve (12) groups of models trained (b) without temperature data, respectively.

In any group comprising data from multiple countries, data from the countries were not mixed together, in order to maintain the sequence or pattern of data for each country. Hence, a model is first trained with the set of data for a country, then retrained with the next set of data from the following country, and so on, within each group. All forecasting models were trained with the same parameters.

Bootstrap aggregation method was applied in the random forest training. The adaptive moment estimation (ADAM) [60] optimization method was used for the gradient descent in deep learning. Initial learning rate, learn rate drop factor, L2 regularization, minibatch size and maximum epochs were set to 0.01, 0.1, 0.0001, 15 and 250, respectively, in the deep learning approach. Data was never shuffled as that would disrupt the pattern of the training data.

8. Results and discussions

8.1. Results on covariate analysis

Results of Granger-causality tests on the eight (8) covariates on cases, mortalities and ***C:M*** responses across 36 countries for demographic H_0^d, economic H_0^e and geographical H_0^g factors, are shown in Table 2. Granger-causality tests of the null hypothesis for Population, Total tests / 1M and GDP per capita of 36 countries on ***C:M*** responses yielded p-values < 0.05, in the last row of Table 2, and hence H_0 is rejected for these cases for ***C:M***. It appears that population, total tests per 1M and GDP per capita affect ***C:M***.

An explanation for the rejection of H_0 in the three cases, is explored by plotting the distribution of covariates versus their respective yearly average ***C:M*** responses in Fig. 4. The plots in

Table 4
Regression analysis of ***C:M*** responses on measurements of temperature from 36 countries.

	Country	*C*	*S*	*Cf*	*Sc*	*Pv*	SSR	SSE	SSR > SSE	Temperature (Predictor)		***C:M*** ratios (response)	
										Min	Ave	Min	Max
1	Argentina	43.40	-0.50	100	-1.14	0.207	220	1759	No	4.97	14.19	14.83	55.10
2	Australia	20.76	1.39	100	6.67	0.186	1458	10572	No	11.98	19.85	8.79	84.36
3	Botswana	71.92	0.48	100	0.67	0.000	108	30	Yes	13.22	20.82	76.26	85.00
4	Brazil	141.90	-4.96	100	-3.49	0.044	873	2497	No	18.99	21.62	8.87	54.94
5	Canada	48.29	1.99	100	4.11	0.000	10978	914	Yes	-21.68	0.21	9.37	79.91
6	Chile	10.02	2.38	100	23.74	0.002	2052	2027	Yes	7.46	13.06	23.80	77.12
7	China	62.78	0.38	100	0.61	0.648	107	6891	No	3.19	16.02	11.91	85.00
8	Egypt	25.45	-0.44	100	-1.74	0.015	160	294	No	9.54	22.86	7.92	25.69
9	Ethiopia	148.03	-4.97	100	-3.36	0.116	677	3378	No	15.60	17.69	35.92	80.11
10	Finland	36.02	2.78	100	7.73	0.000	6093	2765	Yes	0.15	8.60	14.61	81.58
11	France	-4.36	4.50	-100	103.30	0.000	11418	3724	Yes	4.72	12.49	4.42	81.77
12	Germany	10.45	4.04	100	38.69	0.000	9525	3455	Yes	3.92	11.78	13.23	81.12
13	India	-19.62	2.63	-100	13.41	0.001	4822	3959	Yes	13.58	27.71	20.65	81.12
14	Iran	25.77	-0.34	100	-1.31	0.053	149	467	No	3.45	20.10	11.67	35.57
15	Israel	38.92	1.39	100	3.58	0.105	962	4483	No	11.78	20.25	18.32	81.50
16	Italy	-24.40	3.99	-100	16.36	0.000	12237	2390	Yes	5.76	15.80	6.15	80.17
17	Japan	20.09	2.73	100	13.57	0.000	8649	4423	Yes	1.09	12.86	7.44	80.35
18	Mexico	0.99	0.40	100	40.30	0.114	17	84	No	15.22	21.73	5.95	15.35
19	Netherlands	-7.75	4.84	-100	62.41	0.000	11489	2490	Yes	5.16	12.46	8.25	80.45
20	New Zealand	100.54	-4.69	100	-4.67	0.002	4930	5063	No	3.07	7.66	10.07	85.00
21	Nigeria	126.85	-2.69	100	-2.12	0.437	341	7453	No	22.86	25.57	13.45	80.70
22	Norway	54.21	2.35	100	4.34	0.000	5711	1897	Yes	-7.76	2.99	20.65	81.20
23	Philippines	-242.60	11.04	-100	4.55	0.000	4245	1581	Yes	23.42	26.28	13.29	79.52
24	Poland	21.44	1.63	100	7.61	0.000	2199	1426	Yes	1.89	11.94	16.46	62.48
25	Portugal	-11.97	4.33	-100	36.21	0.001	4729	3518	Yes	11.57	16.71	19.88	79.56
26	Russia	54.81	-0.50	100	-0.91	0.005	665	828	No	-22.27	2.19	43.20	77.22
27	South Africa	11.34	1.31	100	11.59	0.108	522	2477	No	10.51	16.33	16.44	67.12
28	Spain	10.02	3.36	100	33.57	0.001	7325	5418	Yes	5.29	13.77	5.38	81.63
29	Switzerland	24.22	4.34	100	17.91	0.000	12582	2713	Yes	-1.73	7.49	10.00	81.49
30	UAE	74.90	0.15	100	0.19	0.180	17	123	No	17.71	30.67	68.18	80.98
31	UK	9.28	3.85	100	41.53	0.000	4603	2731	Yes	4.37	10.45	13.95	80.42
32	USA	77.13	-0.38	100	-0.49	0.511	194	5967	No	0.45	13.56	18.42	85.00
33	Uganda	516.86	-21.1	100	-4.08	0.000	12160	3499	Yes	20.36	22.34	5.61	81.04
34	Venezuela	120.21	-1.78	100	-1.48	0.452	167	3910	No	24.18	26.32	21.03	83.14
35	Zambia	140.14	-4.13	100	-2.95	0.122	1529	7900	No	16.14	20.31	11.21	80.98
36	Zimbabwe	27.12	1.20	100	4.44	0.534	288	9947	No	13.49	18.75	14.32	85.00
						20			18				

Fig. 4 focus on the distribution across the 36 countries, rather than on each country. The stem plots show weak but significant relationships between the three covariates: population, Tests / 1Mpop, GDP per capita and ***C:M*** response. Other than a few countries with a low population, it appears that the yearly average ***C:M*** response increases as population increases in Fig. 4 (a). Yearly average ***C:M*** responses also increase as Tests / 1 M population increase (Fig. 4 (b)). Other than some countries with low GDPs / Capita, probably in Africa where solar irradiation and high temperatures give some added advantage, yearly average ***C:M*** responses increase as GDP per Capita increase (Fig. 4(c)). These might contribute to some further explanations in this paper.

Results of Granger-causality tests on the four (4) weather data on cases, mortalities and ***C:M*** responses within each country are shown for 36 countries in Table 3. H_0^w is rejected for all cases where the p-value < 0.05. The last row of Table 3 shows the total number of H_0^w rejected for each weather parameter for the three responses: cases, mortalities and ***C:M***. The parameters, solar and temperature, have a very strong impact on mortalities, from the high number of H_0^w rejected. The temperature parameter appears to show the strongest relationship for a number of countries for the three responses. H_0^w of 26 countries and 16 countries were rejected for temperature effects on mortality and ***C:M*** responses, respectively. 26 and 16 are the two highest total numbers in the last row of Table 3. Other than Portugal and Netherlands, temperature impacts COVID-19 mortalities in all temperate countries, as well as many tropical countries. ***C:M*** responses are affected by temperature in ten (10) temperate countries. This shows that temperature is indeed a significant parameter in forecasting the COVID-19 pandemic. It also shows that the response, ***C:M***, is reliable.

8.2. Results on regression analysis

The results for the regression analysis of ***C:M*** responses on the four predictors from 36 countries are shown in Table 4 through Table 7. These tables provide the linear regression line equation intercept ***C*** and slope for each country, ***S***. The intercept fixed at a height of ± 100, ***Cf,*** is provided with the corresponding slope ***Sc***. This helps compare the magnitude of the slopes across all 36 countries and across all four (4) predictors: temperature, solar irradiation, rainfall and relative humidity. The rest of the results portray the p-value of the slope, ***Pv***; the regression sum of squares, **SSR**; total sum of squares, **SST**; the error sum of squares, **SSE**; SSR > SSE – which is indicated by a "yes" or "no"; the average and minimum or maximum temperatures (predictors); and the minimum and maximum ***C:M*** ratios, for each country. The major results of the regression analysis are the SSR and SSE. The relationship is assumed to be due to regression when SSR > SSE. When SSR < SSE, the regression line is mostly due to random noise and could affect the slope. In the results, any country with an extraordinary large slope, where SSR > SSE, normally has a steep regression line due to a large range in ***C:M*** responses and / or small range in minimum and maximum predictor values.

The annual minimum and average temperatures are presented in Table 4. In Table 4, the p-values of the gradients are less than 0.05 in 20 countries indicating that there is a significant relationship between temperature and the spread of COVID-19 in those

Table 5
Regression analysis of ***C:M*** responses on measurements of solar irradiation from 36 countries.

	Country	*C*	*S*	*Cf*	*Sc*	*Pv*	SSR	SSE	SSR > SSE	Solar irradiation		
										Min	Average	Max
1	Argentina	48.30	-2.87	100	-5.94	0.046	504	1475	No	1.72	4.16	8.17
2	Australia	26.47	4.19	100	15.81	0.413	581	11449	No	3.70	5.21	8.27
3	Botswana	70.91	2.04	100	2.88	0.020	45	93	No	4.50	5.43	7.48
4	Brazil	21.94	2.45	100	11.18	0.810	14	3355	No	4.43	5.19	6.08
5	Canada	-10.69	15.31	-100	143.15	0.000	8095	3796	Yes	0.66	3.88	6.11
6	Chile	18.22	5.55	100	30.47	0.000	2614	1465	Yes	1.35	4.13	8.53
7	China	58.79	2.93	100	4.99	0.660	100	6898	No	1.62	3.45	4.68
8	Egypt	35.18	-2.79	100	-7.93	0.014	162	292	No	4.88	7.11	8.35
9	Ethiopia	105.77	-8.33	100	-7.88	0.021	1325	2731	No	3.00	5.47	6.93
10	Finland	16.66	11.37	100	68.23	0.000	7135	1723	Yes	0.31	3.81	6.80
11	France	-24.15	16.17	-100	66.95	0.000	10981	4161	Yes	1.46	4.70	6.91
12	Germany	0.84	15.14	100	1810	0.000	9911	3069	Yes	0.53	3.78	5.89
13	India	62.89	-1.82	100	-2.90	0.816	35	8746	No	4.11	5.26	6.64
14	Iran	35.44	-2.63	100	-7.41	0.042	162	453	No	4.06	6.27	7.94
15	Israel	19.59	7.44	100	37.96	0.005	2438	3007	No	3.33	6.39	8.15
16	Italy	-38.41	15.51	-100	40.39	0.000	10295	4331	Yes	1.82	4.97	7.32
17	Japan	-5.40	13.87	-100	257.08	0.055	3124	9948	No	2.65	4.36	5.94
18	Mexico	8.25	0.21	100	2.58	0.757	1	101	No	4.00	6.47	7.99
19	Netherlands	-4.81	14.20	-100	295.36	0.000	9660	4319	Yes	0.64	4.04	5.73
20	New Zealand	98.84	-10.81	100	-10.94	0.003	4776	5217	No	1.48	3.17	6.47
21	Nigeria	137.09	-15.52	100	-11.32	0.013	2869	4925	No	3.62	5.09	6.12
22	Norway	28.89	9.68	100	33.53	0.008	3086	4521	No	0.39	3.34	5.22
23	Philippines	-34.99	15.48	-100	44.23	0.005	2559	3266	No	3.89	5.34	6.62
24	Poland	10.83	7.99	100	73.73	0.001	1990	1635	Yes	0.78	3.77	5.32
25	Portugal	6.30	9.84	100	156.28	0.000	5370	2877	Yes	1.88	5.50	8.02
26	Russia	68.78	-4.16	100	-6.06	0.001	830	663	Yes	0.52	3.62	6.38
27	South Africa	25.67	1.54	100	6.00	0.442	128	2871	No	2.57	4.63	8.14
28	Spain	-15.54	13.94	-100	89.69	0.000	8575	4168	Yes	1.82	5.16	7.41
29	Switzerland	-28.59	20.03	-100	70.07	0.000	11074	4221	Yes	1.64	4.26	6.19
30	UAE	69.28	1.55	100	2.24	0.135	21	119	No	4.83	6.50	7.63
31	UK	10.42	10.93	100	104.90	0.001	4125	3208	Yes	0.68	3.58	5.76
32	USA	104.39	-6.94	100	-6.65	0.055	1467	4694	No	2.08	4.67	6.58
33	Uganda	227.82	-33.59	100	-14.75	0.009	6269	9391	No	4.13	5.43	6.65
34	Venezuela	129.44	-10.56	100	-8.16	0.059	942	3134	No	4.18	5.30	6.34
35	Zambia	175.74	-21.99	100	-12.51	0.099	1720	7709	No	4.48	5.43	5.93
36	Zimbabwe	101.36	-9.93	100	-9.80	0.335	680	9556	No	4.18	5.20	7.10
						18			13			

countries. The regression analysis in 24 out of 36 countries in Table 4 have a positive gradient showing that temperature is directly proportional with ***C:M*** ratios in these countries. Other than New Zealand, the countries with an impactful negative slope (< -1.2) at a fixed intercept of ± 100, are countries with average temperatures $> 17^{\circ}C$ (tropical weather countries). In most of the temperate countries, temperature is positively correlated with ***C:M*** responses. This shows that the ***C:M*** ratio increases as temperature increases, with the exception of Russia and USA, where Tests per 1M, GDP per capita or other covariates that are not tested in this paper, might be influencing the results. The Tests per 1M from source [40] are high for both USA and Russia. Additionally, USA has one of the highest GDP per Capita [40]. It is not surprising that New Zealand has a negative slope because New Zealand has the most minimal daily cases and mortalities amongst all other 35 countries, as seen in the box plots in Fig. 1(c) and (d). It is, therefore, much less affected by COVID-19. In tropical countries, there is a short variation in temperature across all seasons and, hence, temperature does not impact COVID-19 in tropical countries as much as in temperate countries, resulting in mostly negative slopes in the tropical countries. For the set of intercepts fixed at a height of ± 100, the corresponding slopes are significant as well. SSR > SSE in 18 out of 36 countries in Table 4, where 15 of them are temperate countries. This is remarkable for the temperature predictor.

The results for the regression analysis of ***C:M*** responses on measurements of solar irradiation from 36 countries are shown in Table 5. ***C:M*** responses were not repeated in Tables 5–7, since the values are the same as those in Table 4. The p-values for 18 of the slopes in Table 5 are less than 0.01. Solar irradiation is positively correlated with ***C:M*** in 23 countries: this includes all temperate countries other than Russia, USA and New Zealand, for the same reasons explained above. SSR > SSE in 13 out of 36 countries in Table 5 and the magnitude of the slopes when the intercepts fixed at a height of ± 100 are high, signifying a high impact of solar irradiation on COVID-19 in the 13 countries. The p-values for the 13 countries in Table 5 are ≤ 0.001. Solar irradiation impacts COVID-19 in the 13 countries where the minimum solar irradiation is less than 1.90 $kw\text{-}hr/m^2/day$ and the difference between the average and minimum solar irradiation is greater than 2.6 $kw\text{-}hr/m^2/day$ in the country. Other than Norway, all the 13 countries where solar irradiation have a higher SSR are temperate countries, clearly showing that a high differential in solar intensity impacts ***C:M*** responses. Norway has few cases and mortalities compared to other temperate countries and has the second highest GDP per Capita amongst the 36 countries [40]. Therefore, ***C:M*** ratios increase with significant increase in solar irradiation.

The results for the regression analysis of ***C:M*** responses on measurements of rainfall from 36 countries are shown in Table 6. The minimum rainfall is below 0.05 *mm* for majority of the 36 countries. SSR > SSE in only six (6) out of 36 countries with p-values ≤ 0.002 in Table 6. Other than Brazil, rainfall positively correlates with ***C:M*** responses in the six countries with SSR > SSE. The frequency and high differential in rainfall intensities could have contributed to a higher SSR in the six countries. Therefore, rainfall reduces the impact of COVID-19 in these countries.

Table 6
Regression analysis of ***C:M*** responses on measurements of rainfall from 36 countries.

	Country	*C*	*S*	*Cf*	*Sc*	*Pv*	SSR	SSE	SSR > SSE	Rainfall		
										Min	Ave	Max
1	Argentina	36.46	-0.08	100	-0.23	0.975	0	1979	No	0.00	1.15	3.27
2	Australia	43.73	12.70	100	29.04	0.252	1112	10918	No	0.00	0.36	2.01
3	Botswana	80.51	2.08	100	2.58	0.002	70	68	Yes	0.00	0.71	3.12
4	Brazil	45.85	-3.17	100	-6.92	0.000	2771	599	Yes	0.00	3.52	11.61
5	Canada	15.97	12.36	100	77.37	0.000	7847	4045	Yes	0.43	2.65	5.73
6	Chile	45.45	-1.38	100	-3.04	0.173	523	3557	No	0.00	3.13	14.20
7	China	64.35	0.80	100	1.24	0.403	354	6645	No	0.59	5.69	22.95
8	Egypt	14.67	13.18	100	89.85	0.070	98	356	No	0.00	0.05	0.78
9	Ethiopia	55.62	0.81	100	1.46	0.387	219	3837	No	0.20	5.60	13.96
10	Finland	65.27	-2.13	100	-3.26	0.632	149	8710	No	0.45	2.49	5.94
11	France	51.59	0.14	100	0.26	0.983	1	15142	No	0.18	2.14	4.82
12	Germany	74.09	-8.16	100	-11.02	0.178	1628	11352	No	0.01	1.96	4.86
13	India	38.87	5.15	100	13.25	0.001	4827	3954	Yes	0.00	2.80	10.55
14	Iran	18.76	0.90	100	4.77	0.834	2	614	No	0.00	0.25	1.19
15	Israel	71.05	-3.31	100	-4.66	0.292	429	5016	No	0.00	1.18	5.01
16	Italy	22.05	7.42	100	33.68	0.112	2492	12134	No	0.04	2.24	6.06
17	Japan	45.79	1.71	100	3.74	0.313	947	12125	No	0.59	5.47	15.93
18	Mexico	7.92	0.83	100	10.44	0.001	60	42	Yes	0.00	2.07	7.04
19	Netherlands	61.70	-4.33	100	-7.02	0.395	730	13250	No	0.11	2.12	6.08
20	New Zealand	49.14	4.14	100	8.43	0.382	549	9444	No	1.43	3.73	6.88
21	Nigeria	38.10	3.55	100	9.33	0.001	4409	3385	Yes	0.00	5.63	15.04
22	Norway	38.68	7.83	100	20.25	0.017	2597	5011	No	0.18	2.88	5.81
23	Philippines	51.82	-0.30	100	-0.59	0.511	183	5643	No	2.67	13.76	40.38
24	Poland	35.84	2.39	100	6.66	0.371	209	3416	No	0.06	2.13	6.01
25	Portugal	61.64	-0.64	100	-1.04	0.841	25	8222	No	0.10	1.89	7.15
26	Russia	58.77	-2.17	100	-3.69	0.187	180	1313	No	0.31	2.33	5.11
27	South Africa	38.20	-1.66	100	-4.34	0.064	674	2325	No	0.13	3.26	13.52
28	Spain	52.59	2.14	100	4.07	0.671	169	12574	No	0.10	1.74	5.30
29	Switzerland	35.74	4.97	100	13.89	0.148	2193	13102	No	0.34	4.22	9.77
30	UAE	79.39	-0.14	100	-0.18	0.954	0	140	No	0.00	0.15	1.32
31	UK	57.89	-3.11	100	-5.38	0.315	528	6806	No	0.28	2.68	6.01
32	USA	61.45	3.25	100	5.28	0.375	349	5812	No	1.28	3.24	6.23
33	Uganda	66.79	-5.70	100	-8.53	0.087	3053	12606	No	1.03	3.74	9.84
34	Venezuela	67.86	0.97	100	1.43	0.170	531	3546	No	0.00	5.78	22.70
35	Zambia	55.67	0.33	100	0.60	0.875	17	9411	No	0.00	1.71	10.52
36	Zimbabwe	46.30	3.49	100	7.54	0.227 6	1048	9188	No 6	0.01	0.98	9.76

The results for the regression analysis of ***C:M*** responses on measurements of relative humidity from 36 countries are shown in Table 7. The slopes in only 13 out of 36 countries are < 0.05. SSR > SSE in nine (9) countries with p-values ≤ 0.002. Tropical countries are usually humid compared to temperate countries. In the nine countries where SSR > SSE, only Nigeria has a significant corresponding slope of 6.93 when the intercept is fixed at -100 and is positively correlated with ***C:M*** responses. The corresponding slopes for other eight countries are $< |2.00|$, showing that relative humidity has no significant impact in those countries.

In summary, the regression analysis studies in this Section show that temperature, solar irradiation, and rainfall are positively correlated with ***C:M*** responses. Relative humidity plays a role in humid countries. The higher SSR obtained for the temperature predictor for 18 countries, including most of the temperate countries, shows that temperature has the greatest impact on ***C:M*** responses. In majority of the temperate countries, ***C:M*** ratios increase with rise in temperature, and vice versa, as experienced from April 2020 till February 2021.

8.3. Results on COVID-19 forecasting

The most impactful weather condition is, hence, temperature going by the conclusions of Section 8.2 on regression analysis. Two sets of twenty-four (24) COVID-19 forecasting models were trained using (1) Deep learning and (2) Random forest approaches. A set of 24 trained models comprises (a) 12 models with temperature data input and (b) 12 models without temperature data input. All models were evaluated in this Section using COVID-19 cases and deaths test data [40] between December 1, 2020 and June 15, 2021 pertaining to twenty (20) countries. Temperature test data comprises average daily readings from November 1, 2020 to May 15, 2021 for the countries within groups 1 to 12 listed in Section 7.3.

COVID-19 data comprises a 197-day test set which eventually gets reduced to a time-series of 182 daily cases and daily mortalities due to the averaging based on Eq. (1). The series, $\{r[n] \mid n = 0, \ldots (197 - 15)\}$, were each obtained for the daily cases, daily mortalities, and temperature data per country. Subsequently, a 175-sample point ***C:M*** response (12/23/2020 - 06/15/2021) was obtained from the cases and mortalities in this manner, $\{c_1/m_8, \ldots c_k/m_{k+7}, \ \ldots c_{182-7}/m_{182}\}$.

Test sets of twenty (20) countries grouped into twelve (12) categories based on Section 7.3 were evaluated on (a) the 12 models trained with temperature input and (b) the other 12 models trained without temperature input for both deep learning and random forest models. In this evaluation, we pay particular attention to the trend of the predicted outcome, ***F***, versus the observed outcome, ***O***, in a plot for each country. The root means square error (RMSE) for a set of ***n+1*** samples is calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=0}^{n} (F_i - O_i)}{n+1}} \tag{3}$$

The results for the evaluation on Random Forest (RF) and Deep learning (DL) models, with and without temperature (temp), are shown as RMSE in Table 8. The kinds of models are RF no-temp, DL no-temp, RF temp and DL temp. The lowest RMSE i.e., the best results per country, are emphasized in bold letters. It can

Table 7
Regression analysis of ***C:M*** responses on measurements of relative humidity from 36 countries.

	Country	*C*	*S*	*Cf*	*Sc*	*Pv*	SSR	SSE	SSR > SSE	R. Humidity		
										Min	Ave	Max
1	Argentina	19.44	0.31	100	1.58	2E-01	216	1764	No	30.18	55.11	75.86
2	Australia	90.54	-1.09	100	-1.20	2E-01	1485	10545	No	23.55	38.85	52.24
3	Botswana	72.50	0.25	100	0.35	2E-02	43	95	No	25.02	37.29	51.88
4	Brazil	81.65	-0.67	100	-0.82	1E-03	1852	1518	Yes	38.72	70.16	86.90
5	Canada	375.87	-3.84	100	-1.02	2E-04	7500	4392	Yes	75.88	85.12	95.24
6	Chile	93.19	-0.72	100	-0.77	1E-03	2165	1915	Yes	46.03	72.23	91.93
7	China	49.58	0.25	100	0.50	8E-01	42	6956	No	65.43	77.64	88.25
8	Egypt	7.23	0.25	100	3.50	3E-02	128	326	No	20.60	32.10	55.93
9	Ethiopia	37.69	0.32	100	0.86	3E-01	266	3790	No	46.21	69.12	86.46
10	Finland	359.61	-3.66	100	-1.02	3E-07	7558	1301	Yes	76.18	81.86	93.65
11	France	171.17	-1.56	100	-0.91	4E-03	6847	8295	No	50.41	76.47	94.06
12	Germany	246.68	-2.51	100	-1.02	2E-03	6762	6218	Yes	63.74	75.00	92.86
13	India	21.66	0.62	100	2.86	4E-02	2316	6466	No	20.50	51.06	83.83
14	Iran	15.12	0.13	100	0.89	3E-01	45	571	No	11.25	28.79	47.60
15	Israel	182.09	-1.66	100	-0.91	2E-02	1689	3755	No	55.74	69.27	77.44
16	Italy	200.00	-2.31	100	-1.16	1E-01	2650	11976	No	58.42	69.82	81.39
17	Japan	-80.64	1.70	-100	2.11	3E-01	892	12180	No	74.20	79.81	87.97
18	Mexico	6.36	0.07	100	1.06	7E-02	22	79	No	24.65	48.59	75.92
19	Netherlands	273.18	-2.87	100	-1.05	4E-03	6452	7527	No	67.44	76.98	92.89
20	New Zealand	-244.46	3.48	-100	1.42	1E-03	5342	4651	Yes	79.79	88.79	94.54
21	Nigeria	-14.86	1.03	-100	6.93	8E-08	6852	942	Yes	34.50	70.89	89.46
22	Norway	264.75	-2.38	100	-0.90	6E-06	5914	1693	Yes	74.39	85.53	95.97
23	Philippines	270.90	-2.67	100	-0.99	3E-01	534	5292	No	77.68	83.64	86.32
24	Poland	104.01	-0.90	100	-0.86	2E-02	1258	2367	No	56.62	70.24	90.86
25	Portugal	215.32	-2.05	100	-0.95	3E-03	3926	4321	No	59.07	75.53	89.17
26	Russia	36.24	0.20	100	0.56	6E-01	30	1463	No	73.68	86.08	98.07
27	South Africa	79.45	-0.68	100	-0.85	4E-02	796	2203	No	51.55	69.03	84.50
28	Spain	124.90	-1.03	100	-0.83	4E-02	3549	9194	No	43.28	66.53	90.30
29	Switzerland	426.56	-4.71	100	-1.10	2E-03	7517	7777	No	71.74	78.51	86.58
30	UAE	86.28	-0.17	100	-0.19	7E-02	30	110	No	32.03	41.04	58.61
31	Uganda	319.30	-3.18	100	-1.00	5E-04	4373	2961	Yes	77.68	84.71	95.51
32	UK	-1.90	0.96	-100	50.39	4E-01	319	5842	No	69.13	77.28	86.17
33	USA	-175.76	2.85	-100	1.62	5E-02	3793	11866	No	67.26	77.59	86.55
34	Venezuela	44.62	0.40	100	0.90	1E-01	666	3411	No	49.71	71.53	90.98
35	Zambia	85.44	-0.47	100	-0.55	3E-01	809	8620	No	36.52	61.74	85.44
36	Zimbabwe	-38.45	1.69	-100	4.38	2E-02	3580	6655	No	35.22	52.29	74.54
						13			**9**			

Table 8
Results of RMSE for the evaluation of test sets on trained models: RF no-temp, DL no-temp, RF temp and DL temp.

RMSE on Models	USA	UK	Italy	Poland	Argentina	Brazil	Russia	South Africa	Iran	Nigeria	Total RMSE
RF no-temp	8.74	**10.11**	17.83	17.76	18.61	**3.21**	13.44	17.37	34.57	19.33	160.96
DL no-temp	12.02	12.31	6.61	**10.52**	17.50	5.36	6.10	10.11	**8.97**	**12.31**	101.80
RF temp	8.85	11.49	12.62	17.72	17.96	4.57	15.10	21.43	31.57	18.64	159.94
DL temp	**5.46**	12.23	**5.25**	11.55	**15.88**	6.46	**3.69**	**6.70**	9.45	15.51	**92.20**
	Mexico	Egypt	Botswana	Uganda	Switzerland	Germany	Japan	Ethiopia	Canada	Portugal	Total RMSE
RF no-temp	**1.08**	3.55	23.84	11.76	19.65	7.19	7.35	9.52	13.50	**5.09**	102.51
DL no-temp	8.53	2.73	18.51	10.74	18.82	6.30	**5.61**	5.99	**7.33**	7.72	92.27
RF temp	1.77	**2.06**	25.58	**10.10**	11.90	11.16	8.07	**4.15**	10.92	6.53	92.23
DL temp	2.43	2.07	**15.52**	12.84	**11.82**	**6.13**	6.55	5.74	7.83	6.95	**77.89**

be observed that the deep learning model with temperature input presents the lowest RMSE for many of the temperate countries in the evaluation in Table 8. The rightmost column shows the total RMSE for each kind of model. The total RMSE is lowest for DL temp, followed by DL no-temp, RF temp and RF no-temp.

We further examine the results of the best two approaches, DL temp and DL no-temp, by plotting the forecast from 12/23/2020 through 07/15/2021. Results for the evaluation on models trained with both COVID-19 and temperature input (DL temp) and those with COVID-19 data input only (DL no-temp) are presented in plots in Fig. 5 and Fig. 6, respectively. Each sub plot in Figs. 5 and 6 portrays the observed and forecasted outcome for each of the 20 countries. The legend of the plots for the observed and forecasted outcomes are symbolized as ***O*** and ***F***, respectively, in Figs. 5 and 6. The root means square error (RMSE) between the predicted and observed responses are also provided per country in the plots in Figs. 5 and 6. The RMSE values for most of the temperate countries in Fig. 5, USA, UK, Italy, Argentina, Russia, Switzerland, Germany, and Portugal are remarkably less than those for the same countries in Fig. 6. RMSE for DL temp is also lower for other countries, namely, South Africa, Mexico, Egypt, Botswana and Ethiopia. RMSE in thirteen (13) countries out of 20 countries are lower for DL temp in Fig. 5. In the remaining seven (7) countries, the RMSE of DL no-temp in Fig. 6 is slightly, but not significantly, lower than those of DL temp in Fig. 5. Further, lower RMSE of Brazil, Iran, Nigeria and Uganda, shows that temperature input is not essential in forecasting models for tropical countries.

Johannesburg has cold seasons with average temperatures as low as 5 degrees Celcius so it is not surprising that there is a strong impact of temperature on COVID-19 in South Africa. The

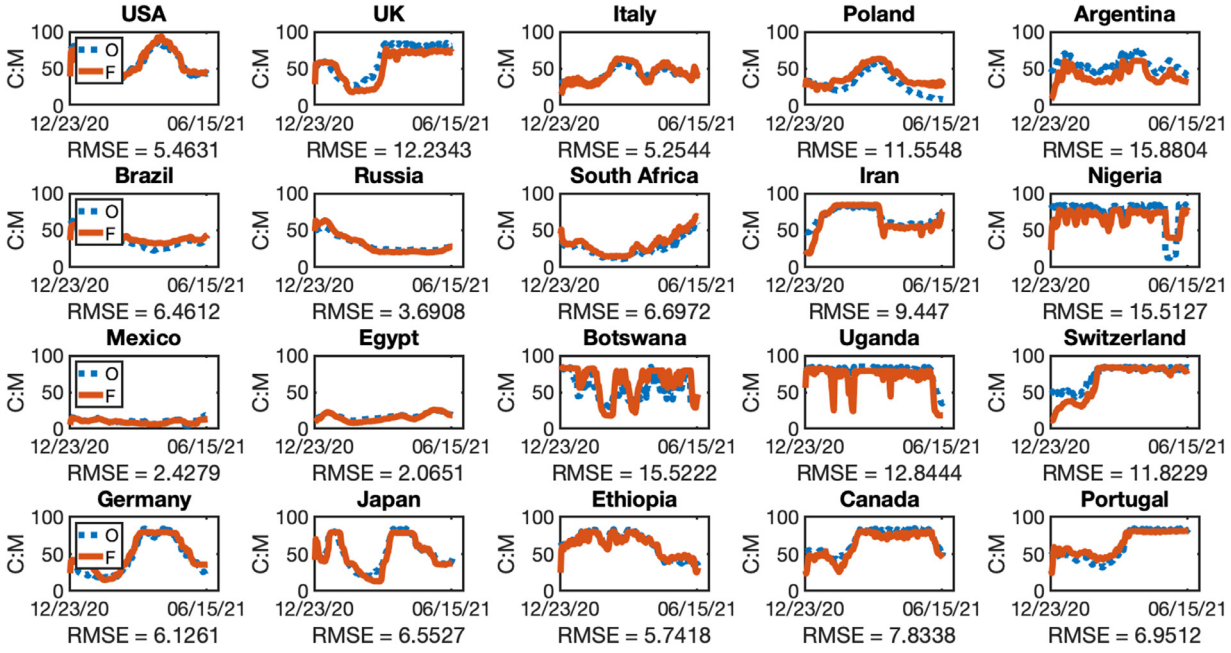

Fig. 5. Results of predicted (F) versus observed (O) COVID-19 cases-to-mortality responses for twenty (20) countries on twelve models trained with COVID-19 and temperature data input (DL temp).

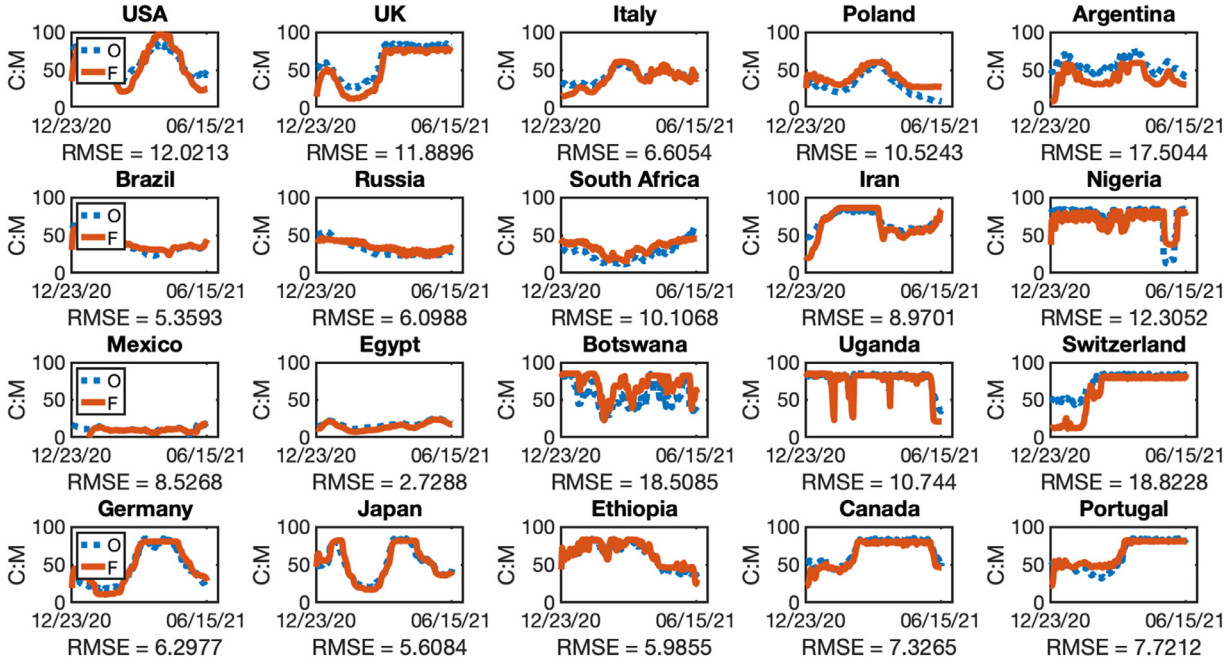

Fig. 6. Results of predicted (F) versus observed (O) COVID-19 cases-to-mortality responses for twenty (20) countries on twelve models trained with COVID-19 data only, without temperature data input (DL no-temp).

trend of the forecasted outcomes in most of the countries in Fig. 5 are also seen to follow the trend of the observed outcomes, showing that incorporating temperature features does indeed improve prediction accuracy for the temperate countries.

In summary, accuracy of predictions on COVID-19 forecasting models for temperate countries can be improved by incorporating temperature data as input in the forecasting (Figs. 5 and 6).

9. Conclusion

Granger-causality tests were used to determine the influences of certain geographical, economic, demographic, and weather covariates on COVID-19 response. Other covariates not considered in this paper constitute areas for future studies. Regression analysis was used in analyzing the impact of weather parameters, viz., temperature, solar irradiation, rainfall, and relative humidity, on filtered responses of COVID-19 cases-to-mortality ratios. Results show that few covariates influence COVID-19 spread, and weather parameters have varied effects on the COVID-19 cases to mortality ratios (***C:M***). Out of the four weather parameters, temperature has a significant effect on COVID-19 ***C:M*** ratios. Results show a great impact in temperate countries with widely varying temperature profiles, while tropical countries appear to be unaffected by the temperature parameter. COVID-19 ***C:M*** ratios decrease as it gets

colder in the temperate countries and this calls for urgent attention: the need for more vaccines for COVID-19 virus, for instance. Prediction results from COVID-19 forecasting models trained with temperature features are better than those of forecasting models trained without temperature features, showing that the temperature parameter should be factored-in for accurate forecasting of COVID-19 responses.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

All owners of the free public data used in this paper are duly acknowledged.

References

[1] John Hopkins University, John Hopkins University's Center for Systems Science and Engineering (JHU CSSE), (2020). https://data.humdata.org/.
[2] Ioannidis JPA, Cripps S, Tanner MA. Forecasting for COVID-19 has failed. Int J Forecast 2020. doi:10.1016/j.ijforecast.2020.08.004.
[3] Shefrin H. The psychology underlying biased forecasts of COVID-19 cases and deaths in the United States. Front Psychol 2020;11:1–11. doi:10.3389/fpsyg.2020.590594.
[4] Chin V, Samia NI, Marchant R, Rosen O, Ioannidis JPA, Tanner MA, Cripps S. A case study in model failure? COVID-19 daily deaths and ICU bed utilisation predictions in New York state. Eur J Epidemiol 2020;35:733–42. doi:10.1007/s10654-020-00669-6.
[5] Nadella P, Swaminathan A, Subramanian SV. Forecasting efforts from prior epidemics and COVID-19 predictions. Eur J Epidemiol 2020;35:727–9. doi:10.1007/s10654-020-00661-0.
[6] Han E, Tan MMJ, Turk E, Sridhar D, Leung GM, Shibuya K, Asgari N, Oh J, García-Basteiro AL, Hanefeld J, Cook AR, Hsu LY, Teo YY, Heymann D, Clark H, McKee M, Legido-Quigley H. Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe. Lancet 2020;396:1525–34. doi:10.1016/S0140-6736(20)32007-9.
[7] Peto J, Alwan NA, Godfrey KM, Burgess RA, Hunter DJ, Riboli E, Romer P, Buchan I, Colbourn T, Costelloe C, Smith GDavey, Elliott P, Ezzati M, Gilbert R, Gilthorpe MS, Foy R, Houlston R, Inskip H, Lawlor DA, Martineau AR, McGrath N, McCoy D, Mckee M, McPherson K, Orcutt M, Pankhania B, Pearce N, Peto R, Phillips A, Rahi J, Roderick P, Saxena S, Wilson A, Yao GL. Universal weekly testing as the UK COVID-19 lockdown exit strategy. Lancet 2020;395:1420–1. doi:10.1016/S0140-6736(20)30936-3.
[8] Hamidi S, Sabouri S, Ewing R. Does density aggravate the COVID-19 pandemic?: Early findings and lessons for planners. J Am Plan Assoc 2020;86:495–509. doi:10.1080/01944363.2020.1777891.
[9] Yang Y, Shang W, Rao X. Facing the COVID-19 outbreak: what should we know and what could we do? J Med Virol 2020;92:536–7. doi:10.1002/jmv.25720.
[10] Mcclymont H, Hu W. Weather variability and COVID-19 transmission: a review of recent research. Int J Environ Res Public Health 2021;18:96. doi:10.3390/ijerph18020396.
[11] Naudé W. Artificial intelligence vs COVID-19: limitations, constraints and pitfalls. AI Soc. 2020;35:761–5. doi:10.1007/s00146-020-00978-0.
[12] Zeng X, Ghanem R. Dynamics identification and forecasting of COVID-19 by switching Kalman filters. Comput Mech 2020:1–15. doi:10.1007/s00466-020-01911-4.
[13] Ketu S, Mishra PK. Enhanced Gaussian process regression-based forecasting model for COVID-19 outbreak and significance of IoT for its detection. Appl Intell 2020:1–21. doi:10.1007/s10489-020-01889-9.
[14] Zhong L, Mu L, Li J, Wang J, Yin Z, Liu D. Early prediction of the 2019 novel coronavirus outbreak in the mainland China based on simple mathematical model. IEEE Access 2020;8:51761–9. doi:10.1109/ACCESS.2020.2979599.
[15] Sun D, Duan L, Xiong J, Wang D. Modeling and forecasting the spread tendency of the COVID-19 in China. Adv Differ Equations 2020;2020:1–16. doi:10.1186/s13662-020-02940-2.
[16] Nandi AK. Data modeling with polynomial representations and autoregressive time-series representations, and their connections. IEEE Access 2020;8:110412–24. doi:10.1109/ACCESS.2020.3000860.
[17] Alzahrani SI, Aljamaan IA, Al-Fakih EA. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. J Infect Public Health 2020;13:914–19. doi:10.1016/j.jiph.2020.06.001.
[18] Castillo O, Melin P. Forecasting of COVID-19 time series for countries in the world based on a hybrid approach combining the fractal dimension and fuzzy logic. Chaos Solitons Fractals 2020;140:110242. doi:10.1016/j.chaos.2020.110242.
[19] Cooper I, Mondal A, Antonopoulos CG. Dynamic tracking with model-based forecasting for the spread of the COVID-19 pandemic. Chaos Solitons Fractals 2020;139:110298. doi:10.1016/j.chaos.2020.110298.
[20] Zhao Z, Li X, Liu F, Zhu G, Ma C, Wang L. Prediction of the COVID-19 spread in African countries and implications for prevention and control: a case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. Sci Total Environ 2020;729:138959. doi:10.1016/j.scitotenv.2020.138959.
[21] Chen D-G, Chen X, Chen JK. Reconstructing and forecasting the COVID-19 epidemic in the United States using a 5-parameter logistic growth model. Glob Health Res Policy 2020;5:1–7. doi:10.1186/s41256-020-00152-5.
[22] Ribeiro MHDM, da Silva RG, Mariani VC, dos S Coelho L. Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. Chaos Solitons Fractals 2020;135:109853. doi:10.1016/j.chaos.2020.109853.
[23] Zheng N, Du S, Wang J, Zhang H, Cui W, Kang Z, Yang T, Lou B, Chi Y, Long H, Ma M, Yuan Q, Zhang S, Zhang D, Ye F, Xin J. Predicting COVID-19 in China using hybrid AI model. IEEE Trans Cybern 2020;50:2891–904. doi:10.1109/TCYB.2020.2990162.
[24] Verhagen MD, Brazel DM, Dowd JB, Kashnitsky I, Kashnitsky I, Mills MC. Forecasting spatial, socioeconomic and demographic variation in COVID-19 health care demand in England and Wales. BMC Med. 2020;18:1–11. doi:10.1186/s12916-020-01646-2.
[25] Zeroual A, Harrou F, Dairi A, Sun Y. Deep learning methods for forecasting COVID-19 time-series data: a comparative study. Chaos Solitons Fractals 2020;140:110121. doi:10.1016/j.chaos.2020.110121.
[26] Wieczorek M, Siłka J, Woźniak M. Neural network powered COVID-19 spread forecasting model. Chaos Solitons Fractals 2020;140:110203. doi:10.1016/j.chaos.2020.110203.
[27] Ren H, Zhao L, Zhang A, Song L, Liao Y, Lu W, Cui C. Early forecasting of the potential risk zones of COVID-19 in China's megacities. Sci Total Environ 2020;729:138995. doi:10.1016/j.scitotenv.2020.138995.
[28] Ramchandani A, Fan C, Mostafavi A. DeepCOVIDNet: an interpretable deep learning model for predictive surveillance of COVID-19 using heterogeneous features and their interactions. IEEE Access 2020;8:159915–30. doi:10.1109/access.2020.3019989.
[29] Shastri S, Singh K, Kumar S, Kour P, Mansotra V. Time series forecasting of COVID-19 using deep learning models: India-USA comparative case study. Chaos Solitons Fractals 2020;140:110227. doi:10.1016/j.chaos.2020.110227.
[30] Rustam F, Reshi AA, Mehmood A, Ullah S, On BW, Aslam W, Choi GS. COVID-19 future forecasting using supervised machine learning models. IEEE Access 2020;8:101489–99. doi:10.1109/ACCESS.2020.2997311.
[31] Decock K, Debackere K, Vandamme AM, Van Looy B. Scenario-driven forecasting: modeling peaks and paths. Insights from the COVID-19 pandemic in Belgium. Scientometrics 2020;124:2703–15. doi:10.1007/s11192-020-03591-6.
[32] Méndez-Arriaga F. The temperature and regional climate effects on communitarian COVID-19 contagion in Mexico throughout phase 1. Sci Total Environ 2020;735:139560. doi:10.1016/j.scitotenv.2020.139560.
[33] Gupta S, Raghuwanshi GS, Chanda A. Effect of weather on COVID-19 spread in the US: a prediction model for India in 2020. Sci Total Environ 2020;728:138860. doi:10.1016/j.scitotenv.2020.138860.
[34] Rosario DKA, Mutz YS, Bernardes PC, Conte-Junior CA. Relationship between COVID-19 and weather: case study in a tropical country. Int J Hyg Environ Health 2020;229:113587. doi:10.1016/j.ijheh.2020.113587.
[35] Demongeot J, Flet-Berliac Y, Seligmann H. Temperature decreases spread parameters of the new COVID-19 case dynamics. Biology 2020;9:94. doi:10.3390/biology9050094.
[36] Byass P. Eco-epidemiological assessment of the COVID-19 epidemic in China, January–February 2020. Glob Health Action 2020;13:1760490. doi:10.1080/16549716.2020.1760490.
[37] Menebo MM. Temperature and precipitation associate with COVID-19 new daily cases: a correlation study between weather and Covid-19 pandemic in Oslo, Norway. Sci Total Environ 2020;737:139659. doi:10.1016/j.scitotenv.2020.139659.
[38] Sharma P, Singh AK, Agrawal B, Sharma A. Correlation between weather and <scp>COVID</scp>-19 pandemic in India: an empirical investigation. J Public Aff 2020. doi:10.1002/pa.2222.
[39] Runkle JD, Sugg MM, Leeper RD, Rao Y, Matthews JL, Rennie JJ. Short-term effects of specific humidity and temperature on COVID-19 morbidity in select US cities. Sci Total Environ 2020;740:140093. doi:10.1016/j.scitotenv.2020.140093.
[40] Worldometer, Coronavirus update (live): cases and deaths from COVID-19 virus pandemic - Worldometer, (2021). https://www.worldometers.info/coronavirus/ (accessed February 20, 2021).
[41] NASA, National Aeronautics and Space Administration's (NASA) solar and meteorological data sets, (2020). https://power.larc.nasa.gov/.
[42] Yang W, Kandula S, Huynh M, Greene SK, Van Wye G, Li W, Chan HT, McGibbon E, Yeung A, Olson D, Fine A, Shaman J. Estimating the infection-fatality risk of SARS-CoV-2 in New York City during the spring 2020 pandemic wave: a model-based analysis. Lancet Infect Dis 2021;21:203–12. doi:10.1016/S1473-3099(20)30769-6.
[43] Perone G. The determinants of COVID-19 case fatality rate (CFR) in the Italian regions and provinces: an analysis of environmental, demographic, and healthcare factors. Sci Total Environ 2021;755:142523. doi:10.1016/j.scitotenv.2020.142523.
[44] Schuchat A. Public health response to the initiation and spread of pandemic COVID-19 in the United States, February 24–April 21, 2020. MMWR. Morb Mortal Wkly Rep 2020;69:551–6. doi:10.15585/mmwr.mm6918e2.

[45] Daoust J-F. Elderly people and responses to COVID-19 in 27 Countries. PLoS ONE 2020;15:e0235590. doi:10.1371/journal.pone.0235590.
[46] Milani F. COVID-19 outbreak, social response, and early economic effects: a global VAR analysis of cross-country interdependencies. J Popul Econ 2021;34:223–52. doi:10.1007/s00148-020-00792-4.
[47] Haug N, Geyrhofer L, Londei A, Dervic E, Desvars-Larrive A, Loreto V, Pinior B, Thurner S, Klimek P. Ranking the effectiveness of worldwide COVID-19 government interventions. Nat Hum Behav 2020;4:1303–12. doi:10.1038/s41562-020-01009-0.
[48] Alfano V, Ercolano S. The efficacy of lockdown against COVID-19: a cross-country panel analysis. Appl Health Econ Health Policy 2020;18:509–17. doi:10.1007/s40258-020-00596-3.
[49] Friston KJ, Bastos AM, Oswal A, van Wijk B, Richter C, Litvak V. Granger causality revisited. Neuroimage 2014;101:796–808. doi:10.1016/j.neuroimage.2014.06.062.
[50] Shahid F, Zameer A, Muneeb M. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. Chaos Solitons Fractals 2020;140:110212. doi:10.1016/j.chaos.2020.110212.
[51] Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos Solitons Fractals 2020;135:109864. doi:10.1016/j.chaos.2020.109864.
[52] Bonacini L, Gallo G, Patriarca F. Identifying policy challenges of COVID-19 in hardly reliable data and judging the success of lockdown measures. J Popul Econ 2021;34:275–301. doi:10.1007/s00148-020-00799-x.
[53] Roda WC, Varughese MB, Han D, Li MY. Why is it difficult to accurately predict the COVID-19 epidemic? Infect Dis Model 2020;5:271–81. doi:10.1016/j.idm.2020.03.001.
[54] Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. Int J Infect Dis 2020;93:284–6. doi:10.1016/j.ijid.2020.02.060.
[55] Huang L, Zhang X, Zhang X, Wei Z, Zhang L, Xu J, Liang P, Xu Y, Zhang C, Xu A. Rapid asymptomatic transmission of COVID-19 during the incubation period demonstrating strong infectivity in a cluster of youngsters aged 16-23 years outside Wuhan and characteristics of young patients with COVID-19: a prospective contact-tracing study. J Infect 2020;80:e1–e13. doi:10.1016/j.jinf.2020.03.006.
[56] Yang L, Dai J, Zhao J, Wang Y, Deng P, Wang J. Estimation of incubation period and serial interval of COVID-19: analysis of 178 cases and 131 transmission chains in Hubei province, China. Epidemiol Infect 2020;148:1–6. doi:10.1017/S0950268820001338.
[57] Verma V, Vishwakarma RK, Verma A, Nath DC, Khan HTA. Time-to-death approach in revealing chronicity and severity of COVID-19 across the world. PLoS ONE 2020;15:e0233074. doi:10.1371/journal.pone.0233074.
[58] Wilson N, Kvalsvig A, Barnard LT, Baker MG. Case-fatality risk estimates for COVID-19 calculated by using a lag time for fatality. Emerg Infect Dis 2020;26:1339–41. doi:10.3201/EID2606.200320.
[59] Yang W, Kandula S, Huynh M, Greene SK, Van Wye G, Li W, Chan HT, McGibbon E, Yeung A, Olson D, Fine A, Shaman J. Estimating the infection-fatality risk of SARS-CoV-2 in New York City during the spring 2020 pandemic wave: a model-based analysis. Lancet Infect Dis 2020;0. doi:10.1016/S1473-3099(20)30769-6.
[60] Kingma DP, Ba JL. Adam: a method for stochastic optimization. In: Presented in the 3rd International Conference for Learning Representations (ICLR), San Diego, CA, USA, May 7-9, 2015. Arxiv; 2017. p. 1–15. 1412.6980. https://arxiv.org/abs/1412.6980v9.