Predictive Modeling on the Number of Covid-19 Death Toll in the United States Considering the Effects of Coronavirus-Related Changes and Covid-19 Recovered Cases

Hoang Pham
Department of Industrial and Systems Engineering, Rutgers University,
Piscataway, New Jersey 08854, USA.
E-mail: hopham@soe.rutgers.edu

Abstract
COVID-19 is caused by a coronavirus called SARS-CoV-2. Many countries around the world implemented their own policies and restrictions designed to limit the spread of Covid-19 in recent months. Businesses and schools transitioned into working and learning remotely. In the United States, many states were under strict orders to stay home at least in the month of April. In recent weeks, there are some significant changes related restrictions include social-distancing, reopening states, and staying-at-home orders. The United States surpassed 2 million coronavirus cases on Monday, June 15, 2020 less than five months after the first case was confirmed in the country. The virus has killed at least 115,000 people in the United States as of Monday, June 15, 2020, according to data from Johns Hopkins University.

With the recent easing of coronavirus-related restrictions and changes on business and social activity such as stay-at-home, social distancing since late May 2020 hoping to restore economic and business activities, new Covid-19 outbreaks are on the rise in many states across the country. Some researchers expressed concern that the process of easing restrictions and relaxing stay-at-home orders too soon could quickly surge the number of infected Covid-19 cases as well as the death toll in the United States. Some of these increases, however, could be due to more testing sites in the communities while others may be are the results of easing restrictions due to recent reopening and changed policies, though the number of daily death toll does not appear to be going down in recent days due to Covid-19 in the U.S. This raises the challenging question:

- How can policy decision-makers and community leaders make the decision to implement public policies and restrictions and keep or lift staying-at-home orders of ongoing Covid-19 pandemic for their communities in a scientific way?

In this study, we aim to develop models addressing the effects of recent Covid-19 related changes in the communities such as reopening states, practicing social-distancing, and staying-at-home orders. Our models account for the fact that changes to these policies which can lead to a surge of coronavirus cases and deaths, especially in the United States. Specifically, in this paper we develop a novel generalized mathematical model and several explicit models considering the effects of recent reopening states, staying-at-home orders and social-distancing practice of different communities along with a set of selected indicators such as the total number of coronavirus recovered and new cases that can estimate the daily death toll and total number of deaths in the United States related to Covid-19 virus. We compare the modeling results among the developed models based on several existing criteria. The model also can be used to predict the number of death toll in Italy and the United Kingdom (UK). The results show very encouraging predictability for the proposed models in this study.

The model predicts that 128,500 to 140,100 people in the United States will have died of Covid-19 by July 4, 2020. The model also predicts that between 137,900 and 154,000 people will have died of Covid-19 by July 31, and 148,500 to 169,700 will have died by the end of August 2020, as a result of the SARS-CoV-2 coronavirus that causes COVID-19 based on the Covid-19 death data available on June 13, 2020.

The model also predicts that 34,900 to 37,200 people in Italy will have died of Covid-19 by July 4, and 36,900 to 40,400 people will have died by the end of August based on the data available on June 13, 2020. The model also predicts that between 43,500 and 46,700 people in the United Kingdom will have died of Covid-19 by July 4, and 48,700 to 51,900 people will have died by the end of August, as a result of the SARS-CoV-2 coronavirus that causes COVID-19 based on the data available on June 13, 2020.

The model can serve as a framework to help policy makers a scientific approach in quantifying decision-makings related to Covid-19 affairs.

1. Introduction

Coronaviruses are a large family of viruses that cause illness ranging from the common cold to more severe diseases. COVID-19, short for "coronavirus disease 2019", a virus strain, first identified in Wuhan City, Hubei Province, China, that has only spread in people since December 2019. COVID-19 is a new disease and many experts in the field are still learning how it spreads. It has rapidly spread to many countries around the world including the United States (Patch.com, 2020). Worldwide travel due to globalization has also results in an increased risk of disease transmission (Rothan and Byrareddy, 2020). The outbreak of Covid-19 has rapidly spread to more than 220 countries around the world including the United States.

There were more than 2 million confirmed coronavirus cases on Monday, June 15 and at least 115,000 people have died in the United States according to a tally by Johns Hopkins University (COVID-19, 2020). Already in June, about 300,000 new Covid-19 cases and at least 10,000 more have died in the United States as of Monday, June 15, 2020, according to data from Johns Hopkins University. Worldwide, there are about 8 million confirmed cases and more than 430,000 reported deaths as of Monday, June 15, 2020 (COVID-19, 2020).

Many countries around the world implemented their own policies and restrictions designed to limit the spread of Covid-19 in recent months. In the United States, many states were under strict orders to stay home at least in the month of April. The worst weeks for coronavirus deaths in the U.S. were in early April. In recent weeks, there are some significant developments related restrictions include social-distancing, reopening states, and staying-at-home orders. Some researchers expressed concern the process of easing restrictions and relaxing stay-at-home orders too soon can quickly surge the number of infected Covid-19 cases as well as the death toll in the United States.

COVID-19 is primarily spread through respiratory droplets which means to become infected people generally in within six feet of someone who is contagious and come into contact with these droplets. While the best way to prevent illness is to avoid virus exposure, the Centers for Disease Control and Prevention (CDC, 2020) recommends taking preventive actions to contain the spread of viruses. This includes: (i) avoid touching your eyes, nose, and mouth; (ii) stay home when you are sick: (iii) cover your cough or sneeze with a tissue, then throw the tissue in the trash; (iv) clean and disinfect frequently touched objects and surfaces using a regular household cleaning spray or wipes, and (v) wash your hands often with soap and water for at least 20 seconds, especially after going to the bathroom, before eating, after blowing your nose, coughing, or sneezing. Chin et al. (2020) recently investigated the stability and detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV 2) in different environmental conditions including temperatures, surfaces. Pham and Pham (2020) also investigated the presence of breast cancer based on biomarkers by developing a generalized logistic dependent model. The US Centers for Disease Control and Prevention projected more than 127,000 coronavirus deaths in the country by June 27 (CNN, 2020).

Recently there are some researchers have been developed mathematical models to determine timelines for social distancing, healthcare policies and social distancing practice (De Brouwer et al., 2020; Dong et al., 2020; Prem et al., 2020; Sebastiani et al., 2020), case fatality rate (Onder et al., 2020; Rajgor et al., 2020; Battegay et al., 2020), disease transmission (Kucharski, 2020; Law et al., 2008) and disease detection in different environments (Chin et al., 2020). Pham (2020) recently develops a model that can estimate the cumulative number of deaths in the United States due to the ongoing Covid-19 virus based on a logistic function. This model, however, assumed that there is no significant change in the coming days due to various testing strategies, social-distancing policies, the reopening of community strategies, and staying-at-home orders policy. The model also did not take into account the daily new cases and the total number of new cases of Covid-19. Based on the Covid-19 global and United States death data, it appears that the cumulative number of deaths can follow an S-shaped curves. There are a number of existing S-shaped logistic models in the literature (Verhulst, 1845; Pham, 2007; Pham et al., 2014; Pham, 2018; Pham, 2019).

In this paper, we develop a novel generalized mathematical model and several explicit models considering the effects of recent coronavirus-related restrictions include reopening states, staying-at-home orders and social-distancing practice of different communities as well as coronavirus recovered and new cases. The models can estimate the number of daily deaths and total number of death toll in the United States related to Covid-19 virus.

We analyze the proposed models using several existing criteria for the Covid-19 data to predict the death toll in the United States, Italy and UK. The results show that the proposed models fit significantly well. The results show very encouraging predictability for the proposed models in this study. Section 2 discusses the model development and present several explicit models that can estimate the total number of deaths in the US, Italy and UK. Section 3 discusses the modeling results based on the Covid-19 deaths data. Section 4 briefly discuss the findings and conclusion remarks.

2. Model Development on Estimating the Number of Deaths
In this section, we discuss the models addressing the effects of recent Covid-19 related changes in the communities such as reopening states, practicing social-distancing, and staying-at-home orders. Our models account for the fact that changes to these policies which can lead to a surge of coronavirus cases and deaths, especially in the US. In other words, we discuss the models with considerations of the effects of recent reopening states, staying-at-home orders and practicing social-distancing of different communities along with a set of selected indicators such as the total number of coronavirus recovered and new cases that can estimate the daily death toll and total number of Covid-19 deaths in the US, Italy and U.K.

2.1 Model Considerations
We first present a generalized model addressing the Covid-19 related changes and restrictions such as reopening states, practicing social-distancing, staying-at-home orders along with a set of selected indicators include the total number of cases and the death toll. Specifically, the model takes into considerations that:

(1). There are a few people in the communities who have already infected Covid-19 virus from the beginning and are spreading the virus into the community but do not know if are infected

the virus. Consequently, the virus is spreading through the people who are in close contact with one another.

(2). The virus is spreading through the areas based on a time-dependent infected rate per person in which it will be spreading at a very slow rate from the beginning due to a small number of infected people and will be spreading at a growth rate much faster due to more number of people who have already infected the virus and are in close contact with non-infected individuals as the time progress, and hopefully with the help of testing strategies and those restrictions designed to limit the spread of Covid-19, then growing slowly until it reaches the maximum daily death toll related to Covid-19 virus.

(3). The model also considers the effects of recent Covid-19 related changes such as reopening states, practicing social-distancing, staying-at-home orders along with a set of selected indicators of the Covid-19 data that can lead to a surge of coronavirus cases and deaths.

(4). We assume deaths are proportional to infections, but with a lag. There can be a significant time lag between when someone is infected and when they die.

2.2 Model Development
We now present a novel generalized mathematical model as follows:

$$\frac{\partial p(t, \underset{\sim}{x})}{\partial t} = \alpha\left[I_{(t, \underset{\sim}{x})}\right]b(t, \underset{\sim}{x})p(t, \underset{\sim}{x})\left(a(t, \underset{\sim}{x}) - p(t, \underset{\sim}{x})\right) + \left(1 - \alpha\left[I_{(t, \underset{\sim}{x})}\right]\right)b(t, \underset{\sim}{x})\left(a(t, \underset{\sim}{x}) - p(t, \underset{\sim}{x})\right) \tag{1}$$

where

$a(t, \underset{\sim}{x}) =$ the total number of deaths plus the death toll to be introduced with respect to $\underset{\sim}{x}$ by time t

$b(t, \underset{\sim}{x}) =$ the time-dependent death rate per person per unit time (i.e., day) w.r.t. $\underset{\sim}{x}$

$p(t, \underset{\sim}{x}) =$ the cumulative number of deaths by time t w.r.t. $\underset{\sim}{x}$

$\underset{\sim}{x} =$ a set of indicators (i.e., recovered cases, daily new cases, total cases, et al)

$\alpha\left[I_{(t, \underset{\sim}{x})}\right] \in [0,1] =$ the overall Covid-19 recovery rate by time t.

The general solution $p(t, \underset{\sim}{x})$ for any given functions $a(t, \underset{\sim}{x})$ and $b(t, \underset{\sim}{x})$ can be found by solving the differential equation from eq. (1). It is expected that the solution for the function $p(t, \underset{\sim}{x})$ would be very complicated, if the solutions exist. In general, one can develop heuristic algorithms using machine learning approaches that can obtain a heuristic numerical solution for the function $p(t, \underset{\sim}{x})$ and the model parameter estimates.

Without loss of generality, based on eq. (1) we specifically propose a new simplified model as follows:

$$\frac{\partial p(t)}{\partial t} = \alpha\left[I_{(t, \underset{\sim}{x})}\right]b(t)p(t)\left(a - p(t)\right) + \left(1 - \alpha\left[I_{(t, \underset{\sim}{x})}\right]\right)b(t)\left(a(t) - p(t)\right) \tag{2}$$

where

$a(t) =$ the total number of deaths (a) plus the death toll due to coronavirus-related restrictions by time t

$b(t) =$ the time-dependent death rate per person per unit time (i.e., day)

$p(t) =$ the cumulative number of deaths by time t

$\underset{\sim}{x} =$ a set of indicators (i.e., recovered cases, daily new cases, total cases, et al)

$\alpha\left[I_{(t,\underset{\sim}{x})} \right] \in [0,1] =$ the overall Covid-19 recovery rate by time t.

It is worth noting that the Covid-19 recovery rate $\alpha\left[I_{(t,\underset{\sim}{x})} \right]$ is the function of time t and $\underset{\sim}{x}$, a set of indicators such as recovered cases, daily new Covid-19 cases, total cases etc. We now present some explicit models based on $\alpha\left[I_{(t,\underset{\sim}{x})} \right]$ as follows:

Model 1: When $\alpha\left[I_{(t,\underset{\sim}{x})} \right] = 1$: We present the following model p(t) that can estimate the cumulative number of deaths at time t (Pham, 2020):

$$p_1(t) = \frac{a}{1 + \left(\dfrac{c}{\beta + e^{bt}} \right)} \tag{3}$$

where $a, b, c,$ and β are the unknown constant parameters. The daily death toll $r_1(t)$ is given by:

$$r_1(t) = \frac{abce^{bt}}{\left[\left(\beta + e^{bt} \right) + c \right]^2} \tag{4}$$

Model 2: When $\alpha\left[I_{(t,\underset{\sim}{x})} \right] = 0$: In this case, we present the following model:

$$p_2(t) = \frac{a(1 + dt)}{1 + \left(\dfrac{c}{\beta + e^{bt}} \right)} \tag{5}$$

where $a, b, c, d,$ and β are the unknown constant parameters The daily death toll $r_2(t)$ can be obtained as follows:

$$r_2(t) = \frac{ad\left(\beta + e^{bt} \right)}{\left(\beta + e^{bt} + c \right)} + \frac{(1 + dt)\left(abce^{bt} \right)}{\left(\beta + e^{bt} + c \right)^2} \tag{6}$$

2.2.1 Weighted Recovery Rate Model
When $0 \le \alpha \le 1$: We propose the following weighted recovery rate model:

$$p(t) = \alpha p_1(t) + (1 - \alpha) p_2(t) \tag{7}$$

where $p_1(t)$ and $p_2(t)$ are given as eq. (3) and (5), respectively. The daily death toll $r(t)$ can be obtained as follows:

$$r(t) = \alpha r_1(t) + (1-\alpha) r_2(t) \tag{8}$$

The value $\alpha\left[I_{(t,x)}\right]$ can be estimated based on the number of recovered cases and the total number of Covid-19 cases by the time t. For various values of $\alpha\left[I_{(t,x)}\right]$, and based on a preliminary study using machine learning approaches, we present the following models in this study:

Model 3:

$$p_3(t) = (2p_1(t) + p_2(t))/3 \tag{9}$$

Model 4:

$$p_4(t) = (3p_1(t) + 2p_2(t))/5 \tag{10}$$

Model 5:

$$p_5(t) = \frac{2p_1(t) + p_2(t)}{6} + \frac{3p_1(t) + 2p_2(t)}{10} \tag{11}$$

Model 6:

$$p_6(t) = \hat{\alpha} p_1(t) + (1-\hat{\alpha}) p_2(t) \tag{12}$$

where $\hat{\alpha}$ is the empirical estimated parameter based on the recovery rate and the total number of Covid-19 cases.

Model 7:

$$p_7(t) = (1-\hat{\alpha}) p_1(t) + \hat{\alpha} p_2(t) \tag{13}$$

A list of all seven model is shown in Table 1. Here we can estimate these unknown parameters *a, b, c, d* and β using a least squares method and compare their results based on various model criteria such as MSE, AIC, BIC, PC, PIC, PRR and PP (see Table A in the Appendix).

Table 1. Models

Model	$p(\mathbf{t})$	$r(\mathbf{t})$
Model 1	$p_1(t) = \dfrac{a}{1+\left(\dfrac{c}{\beta+e^{bt}}\right)}$	$r_1(t) = \dfrac{abce^{bt}}{\left[\left(\beta+e^{bt}\right)+c\right]^2}$
Model 2	$p_2(t) = \dfrac{a(1+dt)}{1+\left(\dfrac{c}{\beta+e^{bt}}\right)}$	$r_2(t) = \dfrac{ad\left(\beta+e^{bt}\right)}{\left(\beta+e^{bt}+c\right)} + \dfrac{(1+dt)\left(abce^{bt}\right)}{\left(\beta+e^{bt}+c\right)^2}$
Model 3	$p_3(t) = (2p_1(t) + p_2(t))/3$	$r_3(t) = (2r_1(t) + r_2(t))/3$
Model 4	$p_4(t) = (3p_1(t) + 2p_2(t))/5$	$r_4(t) = (3r_1(t) + 2r_2(t))/5$
Model 5	$p_5(t) = \dfrac{2p_1(t)+p_2(t)}{6} + \dfrac{3p_1(t)+2p_2(t)}{10}$	$r_5(t) = \dfrac{2r_1(t)+r_2(t)}{6} + \dfrac{3r_1(t)+2r_2(t)}{10}$
Model 6	$p_6(t) = \hat{\alpha}p_1(t)+(1-\hat{\alpha})p_2(t)$	$r_6(t) = \hat{\alpha}r_1(t)+(1-\hat{\alpha})r_2(t)$
Model 7	$p_7(t) = (1-\hat{\alpha})p_1(t)+\hat{\alpha}p_2(t)$	$r_7(t) = (1-\hat{\alpha})r_1(t)+\hat{\alpha}r_2(t)$

3. Modeling Analysis and Prediction Results

In this section, we calculate the total number of deaths and daily death toll for all seven models as shown in Table 1 based on the death data from Worldometer (2020) for the United States, Italy and United Kingdom. Worldometer manually analyzes, validates, and aggregates data from thousands of sources in real time and provides global COVID-19 live statistics for people around the world (Worldometer, 2020). Worldometer data is also used by many research organizations, government of many countries including the Johns Hopkins CSSE (Worldometer, 2020). This Worldometer database provides a comprehensive global data for COVID-19 consisting of many indicators for each countries such as the daily deaths, total deaths, total cases, new cases, active cases, total recovered, serious critical, total cases per one million populations, total tests, and total tests per one million populations.

Application 1: COVID-19 Death Data in the United States

We calculate the total number of Covid-19 deaths and daily death toll in the US of all seven models given in Table 1 based on the available data consisting of 106 days obtained from February 29, 2020 to June 13, 2020 as shown in Table 2. Figures 1, 2, and 3 show the real data on the total number of deaths, daily death toll, and daily number of infected cases of Covid-19 in the US, respectively. Table 3 shows the values of model 1 and 2 based on several existing criteria (see Table A in the appendix) such as MSE, AIC, BIC, PC, PIC, PRR and PP. Using Models 6 and 7, we calculate the upper and lower predictions on the number of deaths and the daily death toll in the US.

Table 2. US Deaths data (Worldometers, 2020) during 2/29/20 - 6/13/20

| Date	Cumulative number of deaths	Date	Cumulative number of deaths	Date	Cumulative number of deaths
2/29	1	4/5	11793	5/11	81847
3/1	1	4/6	13298	5/12	83718
3/2	6	4/7	15526	5/13	85540
3/3	9	4/8	17691	5/14	87293
3/4	11	4/9	19802	5/15	89104
3/5	12	4/10	22038	5/16	90324
3/6	15	4/11	24062	5/17	91189
3/7	19	4/12	25789	5/18	92193
3/8	22	4/13	27515	5/19	93750
3/9	26	4/14	30081	5/20	95155
3/10	30	4/15	32712	5/21	96569
3/11	38	4/16	34905	5/22	97868
3/12	41	4/17	37448	5/23	98904
3/13	48	4/18	39331	5/24	99519
3/14	58	4/19	40901	5/25	100025
3/15	73	4/20	42853	5/26	100800
3/16	95	4/21	45536	5/27	102338
3/17	121	4/22	47894	5/28	103563
3/18	171	4/23	50234	5/29	104776
3/19	239	4/24	52191	5/30	105793
3/20	309	4/25	54256	5/31	106432
3/21	374	4/26	55412	6/1	107153
3/22	509	4/27	56795	6/2	108299
3/23	689	4/28	59265	6/3	109383
3/24	957	4/29	61655	6/4	110415
3/25	1260	4/30	63856	6/5	111390
3/26	1614	5/1	65753	6/6	112096
3/27	2110	5/2	67444	6/7	112469
3/28	2754	5/3	68597	6/8	113055
3/29	3251	5/4	69921	6/9	114148
3/30	4066	5/5	72271	6/10	115130
3/31	5151	5/6	74799	6/11	116034
4/1	6394	5/7	76928	6/12	116825
4/2	7576	5/8	78615	6/13	117527
4/3	8839	5/9	80037		
4/4	10384	5/10	80787		
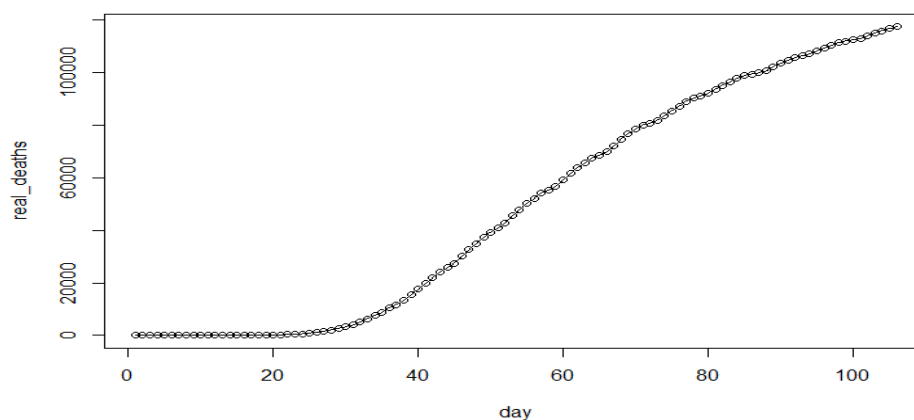
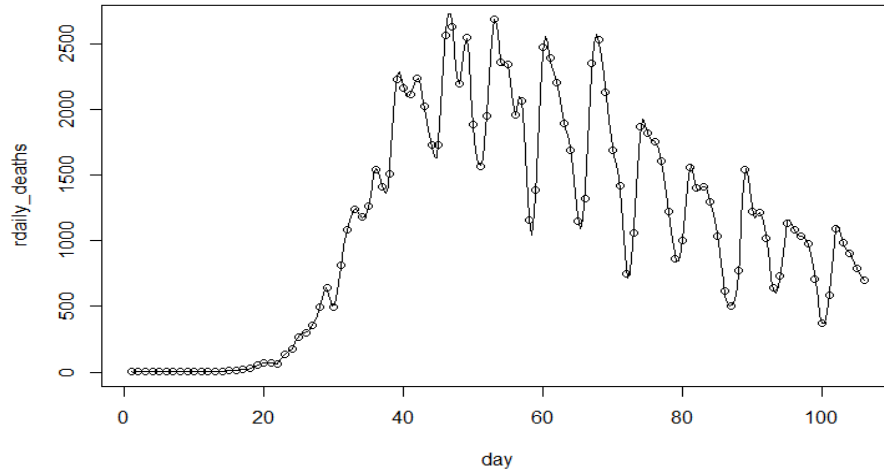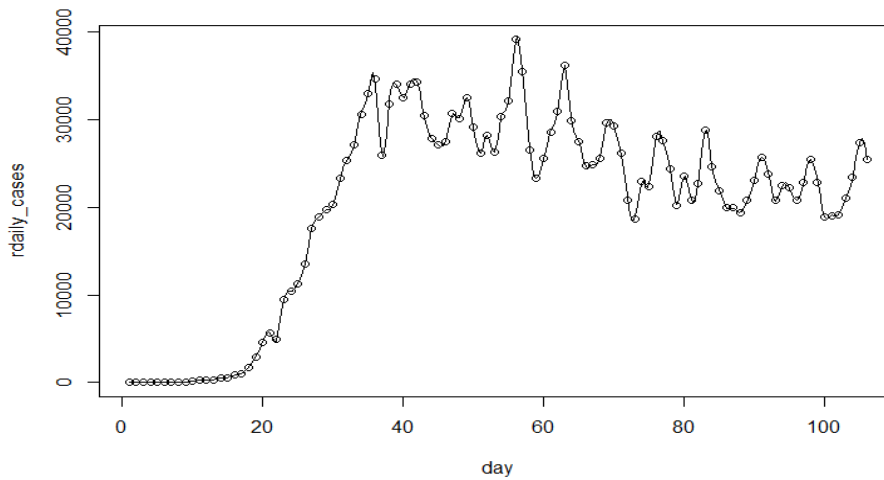Figure 1. Total death toll vs day in the United States

Figure 2. Daily death toll vs day in the United States

Figures 3. Daily number of infected cases of Covid-19 vs. day in the US

Table 3. Model selection criteria based on available data in the United States

Selection Criteria	Model 1	Model 2
p(t)	$p_1(t) = \dfrac{a}{1 + \left(\dfrac{c}{\beta + e^{bt}}\right)}$	$p_2(t) = \dfrac{a(1+dt)}{1 + \left(\dfrac{c}{\beta + e^{bt}}\right)}$
MSE	119671330	2389277
AIC	1975.6	1561.65
BIC	1986.2	1574.97
PC	950.77	744.43
PIC	12206475707	241316954
PRR	36.2	24.78
PP	162855.7	188284.4

The model predicts that 128,500 to 140,100 people in the United States will have died of Covid-19 by July 4. The model also predicts that between 137,900 and 154,000 people will have died of Covid-19 by July 31, and 148,500 to 169,700 will have died by the end of August 2020, as a result of the SARS-CoV-2 coronavirus that causes COVID-19 based on the Covid-19 death data available on June 13, 2020. Figures 4 and 5 show the prediction on the total number of people and the number of people daily who have died of Covid-19 in the US, respectively. The results show very encouraging predictability for the proposed models. The model can serve as a framework to help policy makers a scientific approach in quantifying decision-makings related to Covid-19 affairs.
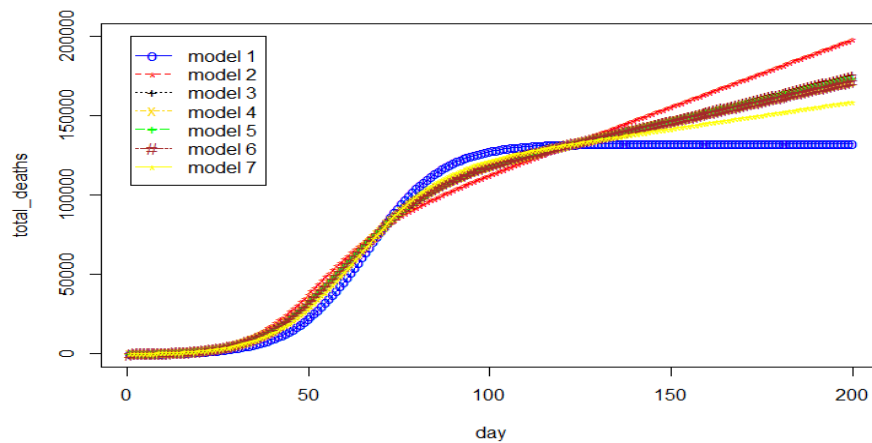
Figure 4. The estimated cumulative number of deaths vs. day in the United States
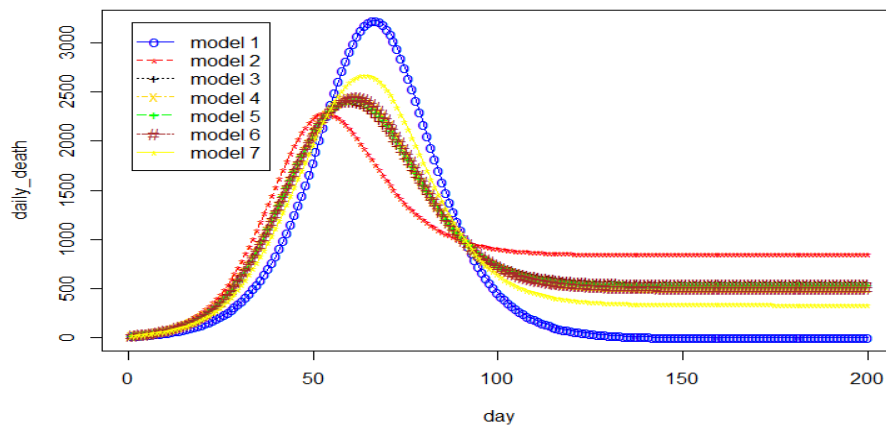
Figure 5. The estimated daily number of deaths vs. day in the United States

Application 2: COVID-19 Death Data in Italy

We calculate the total number of Covid-19 deaths and daily death toll in Italy of all seven models given in Table 1 based on the available data consisting of 114 days from February 21, 2020 to June 13, 2020 (Worldometer, 2020). Similarly, Figure 6 shows the real daily death toll data in

Italy. Table 4 shows the values of model 1 and 2 based on several existing criteria (see Table A in the appendix). From Model 6, we can calculate the upper and lower predictions on the number of deaths and the daily death toll. The model predicts that 34,900 to 37,200 people in Italy will have died of Covid-19 by July 4, and 36,900 to 40,400 people will have died by the end of August 2020 based on the data available on June 13, 2020, as a result of the SARS-CoV-2 coronavirus that causes COVID-19 based on the data available on June 13. Figures 7 and 8 show the predictions on the number of people and the number of people daily who have died of Covid-19 in Italy, respectively. The results show very encouraging predictability for the proposed models.
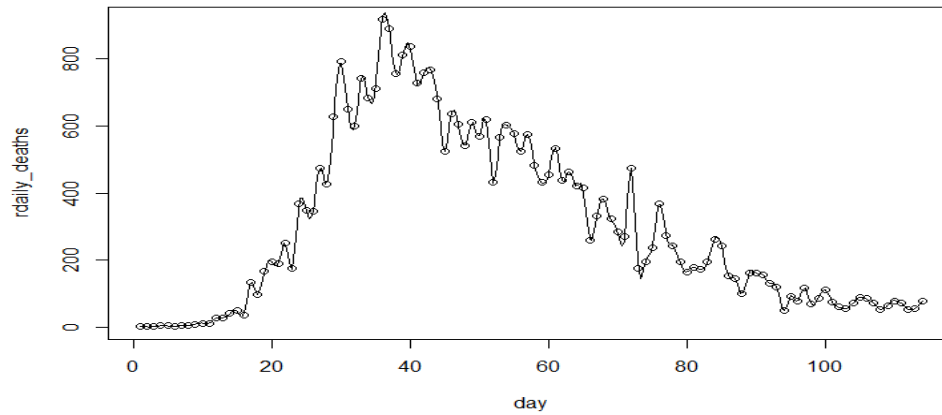
Figure 6: Daily death toll vs. day in Italy

Table 4. Model selection criteria based on available data in Italy

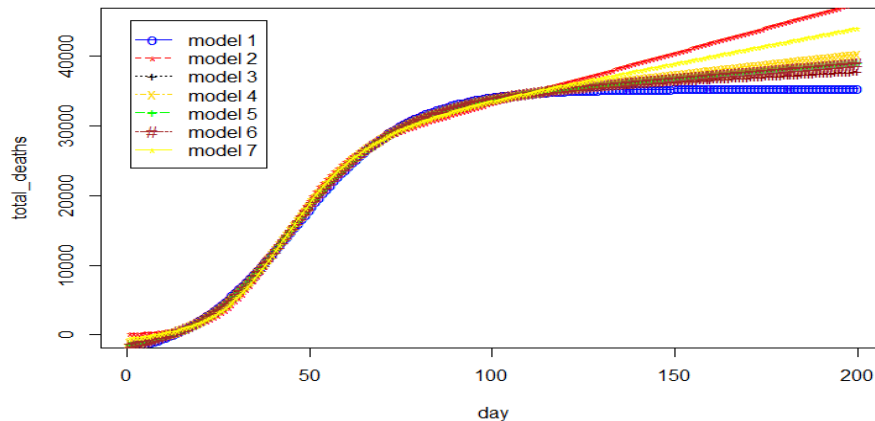| Criteria	Model 1	Model 2
MSE	848915.2	330886.4
AIC	1560.22	1453.77
BIC	1571.17	1467.46
PC	753.0	695.41
PIC	93380681	36066627
PRR	19.77	15.67
PP	5093919	52224.5

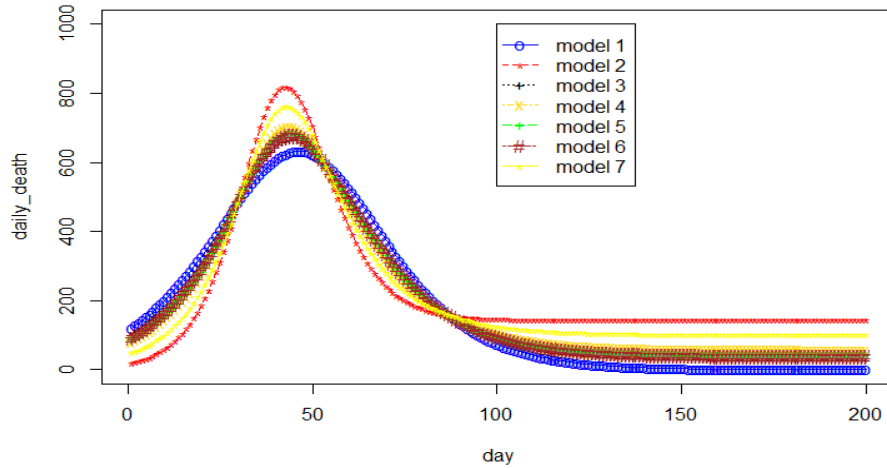Figure 7. The estimated cumulative number of deaths vs. day in Italy

Figure 8. The estimated daily number of deaths vs. day in Italy

Application 3: COVID-19 Death Data in the United Kingdom

We now calculate the total number of Covid-19 deaths and daily death toll in the United Kingdom of all seven models given in Table 1 based on the available data consisting of 101 days from March 5, 2020 to June 13, 2020 (Worldometer, 2020). Figure 9 shows the daily death toll data in the U.K. Table 5 shows the values of model 1 and 2 based on several existing criteria. From Model 6, we can obtain the upper and lower predictions on the number of deaths and the daily death toll in the U.K. The model also predicts that between 43,500 and 46,700 people in the United Kingdom will have died of Covid-19 by July 4, and 48,700 to 51,900 people will have died by the end of August 2020, as a result of the SARS-CoV-2 coronavirus that causes COVID-19 based on the data available on June 13, 2020. Figures 10 and 11 show the predictions on the number of people and the number of people daily who have died of Covid-19 in the U.K., respectively. The results show very encouraging predictability for the proposed models.
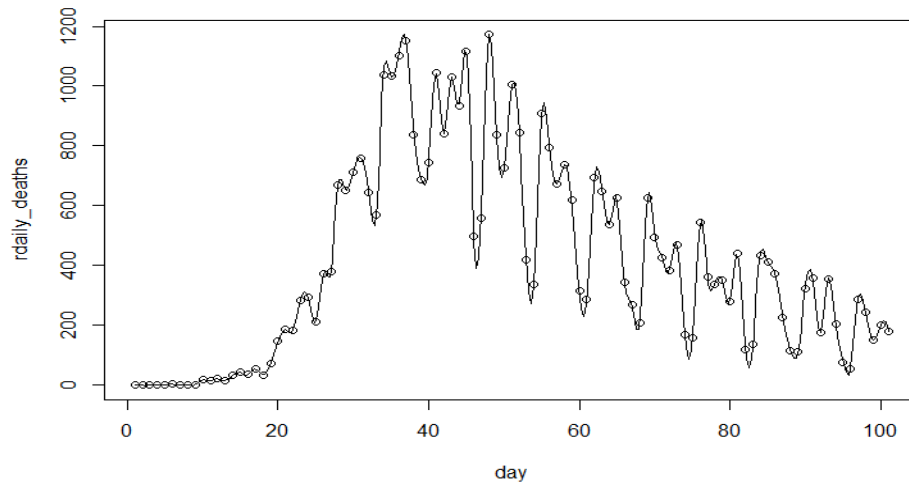

Figure 9. Daily death toll vs day in United Kingdom

Table 5. Model selection criteria based on available data in the United Kingdom

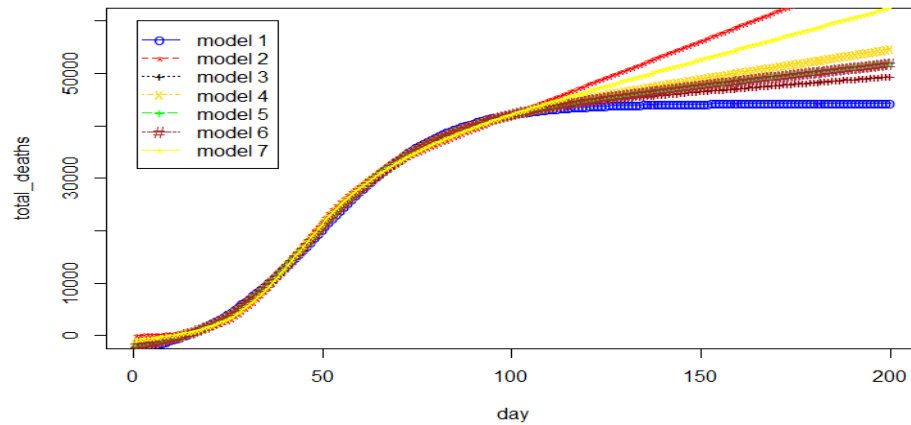| Criteria	Model 1	Model 2
MSE	1826546	275169
AIC	1460.1	1267
BIC	1470.6	1283
PC	701.4	603.9
PIC	177174958	26416219
PRR	21.6	17.87
PP	18579538	97525.6

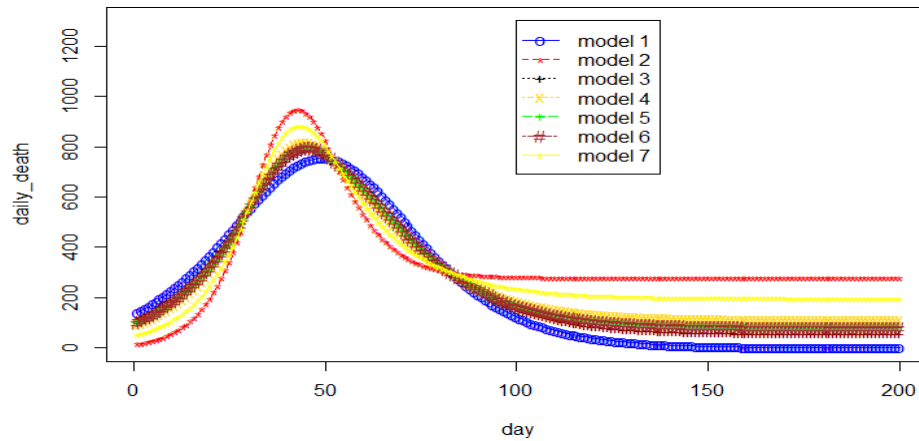Figure 10. The estimated cumulative number of deaths vs day in the United Kingdom (UK)

Figure 11: The estimated daily number of deaths vs day in the United Kingdom (UK)

4. Conclusion

This paper presents a generalized mathematical model and several explicit models considering the effects of coronavirus-related restrictions and changes such as reopening states, staying-at-home orders and social-distancing practice of different communities along with a set of selected indicators such as the total number of coronavirus recovered cases and new cases in the US and

other countries such as Italy and the United Kingdom. The models can predict the daily death toll and total number of deaths in the US related to Covid-19 virus. The models also predict the daily death toll and total number of deaths in Italy and the United Kingdom. The results show very encouraging predictability for the proposed models. The models can serve as a framework to help policy makers a scientific approach in quantifying decision-makings related to Covid-19 affairs.

APPENDIX

Table A. Some criteria model selection

No.	Criteria	Formula	
1	SSE	$$SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$	Measures the total deviations between the predicted values with the actual data observation.
2	MSE (Pham, 2007)	$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-k}$$	Measures the difference between the estimated values and the actual observation.
3	AIC (Akaike, 1973)	$$AIC = -2\log(L) + 2k$$	The model improves the goodness of the fit but also increases the penalty by adding more parameters.
4	BIC (Schwarz, 1978)	$$BIC = -2\log(L) + k\log(n)$$	Depend on the sample size n that shows how strongly BIC impacts the penalty of the number of parameters in the model.
5	PIC (Pham, 2019)	$$PIC = SSE + k\left(\frac{n-1}{n-k}\right)$$ where $SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	PIC takes into account a larger the penalty when adding too many coefficients in the model when there is too small a sample.
6	PRR (Pham, 2007)	$$PRR = \sum_{i=1}^{n}\left(\frac{\hat{m}(t_i) - y_i}{\hat{m}(t_i)}\right)^2$$	Measures the distance of model estimates from the actual data against the model estimate.
7	PP (Pham, 2007)	$$PP = \sum_{i=1}^{n}\left(\frac{\hat{m}(t_i) - y_i}{y_i}\right)^2$$	Measures the distance of model estimates from the actual data against the actual data.
8	PC (Pham, 2020)	$$PC = \left(\frac{n-k}{2}\right)\log\left(\frac{SSE}{n}\right) + k\left(\frac{n-1}{n-k}\right)$$ where $SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	This new criterion measures the tradeoff between the uncertainty in the model and the number of parameters in the model by slightly increasing the penalty when each time adding parameters in the model when there is too small a sample.

Abbreviations
SSE   sum of squared error
MSE   mean squared error
LSE   least squared estimate
AIC   Akaike's information criterion
BIC   Bayesian information criterion
PC   Pham's criterion
PP   the predictive power
PIC   Pham's information criterion
PRR   the predictive ratio-risk

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B.N. Petrov & F. Caski (eds.), Proceedings of the Second International Symposium on Information Theory (pp. 267-281). Budapest: Akademiai Kiado.

Battegay, M., Kuehl, R., Tschudin-Sutter, S., Hirsch, H.H., Widmer, A.F., & Neher, R.A. (2020). 2019-novel Coronavirus (2019-nCoV): estimating the case fatality rate–a word of caution. *Swiss Medical Weekly*, 150(0506). https://doi.org/10.4414/smw.2020.20203.

CDC (2020). Centers for disease control and prevention; https://www.cdc.gov.

Chin, A.W.H., Chu, J.T. S., Perara, M.R.A., Hui, K.P.Y., Yen, H.-L, Chan, M.C.W., Paris, M., & Poon, L.L.M. (2020). Stability of SARS-CoV-2 in different environmental conditions. *The Lancet Mircobe*, vol. 20, https://doi.org/10.1016/S2666-5247(20)30003-3.

CNN (2020). https://edition.cnn.com/world/live-news/coronavirus-pandemic-06-24-20-intl/index.html

COVID-19 (2020). Dashboard by the center for systems science and engineering. Johns Hopkins University, https://coronavirus.jhu.edu/map.html.

De Brouwer, E., Raimondi, D., & Moreau, Y. (2020). Modeling the COVID-19 outbreaks and the effectiveness of the containment measures adopted across countries. medRxiv. https://doi.org/10.1101/2020.04.02.20046375.

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, *20*(5), 533-534.

Kucharski, A.J., Russell, T.W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R.M., Sun, F., Jit, M., Munday, J.D., Davies, N., Gimma, A., van Zandvoort, K., Gibbs, H., Hellewell, J., Jarvis, C.I., Clifford, S., Quilty, B.J., Bosse, N.I., Abbott, S., Klepac, P., &, Flasche, S. (2020). Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases*. *20*(5), 553-558.

Law, G.R., Feltbower, R.G., Taylor, J.C., Parslow, R.C., Gilthorpe, M.S., Boyle, P., & McKinney, P.A. (2008). What do epidemiologists mean by 'population mixing'? *Pediatric Blood & Cancer*, *51*(2), 155-160.

Onder, G., Rezza, G., & Brusaferro, S. (2020). Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *Jama*, *323*(18), 1775-1776.

Patch.com (2020). https://patch.com/new-jersey/oceancity/nj-coronavirus-update-gov-murphy-considers-curfew-31-new-cases.

Pham, H. (2007). *System software reliability*. Springer Science & Business Media.

Pham, H. (2018). A logistic fault-dependent detection software reliability model. *Journal of Universal Computer Science, 24*(12), 1717-1730.

Pham, H. (2019). A New Criterion for Model Selection. *Mathematics*, *7*(12), 1215.

Pham, H. (2020). On estimating the number of deaths related to Covid-19. *Mathematics*, *8*(5), 655.

Pham, H., & Pham, D.H. (2020). A novel generalized logistic dependent model to predict the presence of breast cancer based on biomarkers. *Concurrency and Computation: Practice and Experience*, *32*(1), e5467.

Pham, H., Pham, D.H., & Pham Jr, H. (2014). A new mathematical logistic model and its applications. *International Journal of Information and Management Sciences*, *25*(2), 79-99.

Prem, K., Liu, Y., Russell, T.W., Kucharski, A.J., Eggo, R.M., Davies, N., ... & Abbott, S. (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet Public Health*, *5*(5), e261-e270. https://doi.org/10.1016/S2468-2667(20)30073-6.

Rajgor, D.D., Lee, M.H., Archuleta, S., Bagdasarian, N., & Quek, S.C. (2020). The many estimates of the COVID-19 case fatality rate. *The Lancet Infectious Diseases*, *20*(7), 776-777. https://doi.org/10.1016/S1473-3099(20)30244-9.

Rothan, H.A., & Byrareddy, S.N. (2020). The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *Journal of Autoimmunity*, *109*, 102433.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461-464.

Sebastiani, G., Massa, M., & Riboli, E. (2020). Covid-19 epidemic in Italy: evolution, projections and impact of government measures. *European Journal of Epidemiology*, *35*(4), 341-345.

Verhulst, P.F. (1845). Resherches mathematiques sur la loi d'accroissement de la population. *Nouveaux Memoires De L'academie Royale Des Sciences*, *18*, 1-41.

Worldometers (2020). https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1?#countries.