Dynamic modeling of the impact of socio-economic restrictions and behavior on COVID-19 outbreak

Mikhail Pomazanov¹ · Artem Arkhipov¹ · Alexander Karminsky¹

Received: 15 July 2020 / Revised: 22 May 2021 / Accepted: 22 May 2021 /

Published online: 29 July 2021

© Eurasia Business and Economics Society 2021

Abstract

The COVID-19 pandemic had an enormous social and economic impact on societies in 2020. The epidemiological situation is evolving on a daily basis, and the methodology of how to evaluate the impact of the pandemic and the severity of its consequences is lacking. The only available high-frequency data now are the number of people who have contracted the illness, and the restrictive measures that authorities have implemented to contain the outbreak. The most important question now is whether authorities can prevent subsequent waves. The contribution of the paper is a dynamic model of COVID-19 outbreaks, on the basis of which we investigated the possible impact of the socio-economic behavior and restrictions on its waves. To build the model, a large database for different countries with a wide range of economic and social institutions was collected. We give a detailed description of the model and a comparison of the results with trajectories of the outbreaks in the countries under consideration. The proposed model describes the empirical results and can be used for timely and contemporary predictions of the stages of pandemics. Despite this, the model needs future development and verification because the pandemic is not over, and the accumulation of empirical information continues. Yet the model might also be useful as a basis for researching the impacts of other socio-economic and medical actions for containing pandemics.

Keywords Contagion · Infection model · COVID-19 · Dynamic model · Empirical information

✉ Alexander Karminsky
akarminsky@hse.ru

Mikhail Pomazanov
mhubble@yandex.ru

Artem Arkhipov
a.v.arkhipov@mail.ru

¹ Faculty of Economic Science, National Research University Higher School of Economics, Moscow, Russia

1 Introduction

In 2020, the world faced an unprecedented challenge in the form of a new coronavirus outbreak. One of the characterizing features of the outbreak is the difficulty in forecasting its size, and how to link the pandemic to containment measures. The pandemic became a trigger which activated the need to predict the dynamics of infections and their impact on social and economic processes and conditions. However, the available approaches to model these aspects of epidemics do not appear to be plausible solutions to the current situation.

Among the classical epidemiological models, there are three main types of deterministic (parametric) models for infectious diseases that spread as a result of contact with a person in a population, which are well discussed in Hethcote (1989). The acronym SIR stands for a proportion of susceptible (or healthy) “S” people in the population, the proportion of infected “I”, and the percentage of recovered or dead “R” who are supposed to be immune, so that $S + I + R = 100\%$. SIR is used for diseases where the infection provides permanent immunity. The second approach, SIS (susceptible-infected-susceptible), does not take into account immunity, so that $S + I = 100\%$. The third type of model takes into account the natural birth and mortality rates in the population, so this type is focused on long-term modeling of epidemics. When a SIR outbreak happens over a relatively short time (less than 1 year), then this outbreak is called an epidemic. Since an epidemic is relatively fast, the model does not include birth and death dynamics. Epidemics are common in diseases such as flu, measles, rubella, and chickenpox. A good overview of non-parametric models is presented in Choisy et al. (2007).

Classical parametric models cannot be used to describe the COVID-19 pandemic, as they have serious drawbacks, e.g. they do not take into account active administrative restrictions to prevent the spread of the epidemic. On the contrary, the parameters of infection distribution are assumed constant, which implies that diseases are left to spread on their own. According to empirical observations, modeling attempts are based on a set of assumptions, e.g. defining the epidemic as a time period when the number of confirmed cases increases exponentially, and defining the starting date of the epidemic at a date when a certain artificially chosen threshold (“zero”) level has been overcome, which in many cases is set at the first 100 infected people. The number of infected people at the end of the outbreak can ensure the slow speed of the increase in the total number of infected. The total share of infected people is a small proportion of the population, so $S \gg I + R$, hence there is no sense to include immunity in the model (and thus SIR would not work properly). For the SIS approach there are solutions to differential equations in finite functions (the Bernoulli equation), yet it follows exponential growth at the beginning of outbreak only if the number of infected people does not tend to zero after the peak of the disease.

The purpose of the paper is to build a model which would help in predicting the length and severity of COVID outbreaks in different countries on the basis of the relevant theoretical background, and to discuss reasons for deviations in the actual

dynamics from the predicted path. Such a model and a discussion would provide valuable support to instruments of containment of this and other outbreaks, and would help to minimize the negative impact such outlying events might have on economies and lives.

In Sect. 2, we propose a theoretical approach to modeling the number of those suffering from COVID-19 on the basis of the model developed in Volterra (1976) (which has never been applied to this type of problem) and substantially amended in Pomazanov (2020) to include the peculiarities of epidemic situations. In Sect. 3, we present a parametric contagion model that is being successfully tested for the Ebola virus epidemic. Data sources are described in Sect. 4. In Sect. 5, we consider the deviation from the static model. In Sect. 6, we compare and contrast the findings at the country level and try to identify the nexus between the easing of administrative restrictions and the chances of a second wave of the pandemic. In the last section, we develop an approach which could show how to deal with situations with high numbers of residual infections.

2 Model of contagion and peak (distribution) parameters

A good explanation for the dynamics of a short-term epidemic can be obtained with an amended version of the Lotka–Volterra model¹ (L–V). In its classical form, it describes the interaction of two species, one of which is a predator and the other a prey.

To apply the model to an epidemic, the prey could be defined as the share of the population available for infection $S(t)$, which at the time t equals to one, $S(t) = 1$. It is assumed that the population takes measures to dissociate the “prey” according to the simplest law

$$S'(t) = -\alpha \cdot S(t), \quad (1)$$

where α is a constant reflecting the efficiency of the containment measures.

The predator in this model is the share of infected $N(t)$. A predator, similar to the L–V model, reproduces according to the law,

$$N'(t) = \gamma \cdot \alpha \cdot S(t) \cdot N(t) + \tilde{\beta}, \quad (2)$$

where γ is a constant reflecting the scale of the outbreak; $\tilde{\beta}$ reflects the level of residual infection, which can be identified only at a late stage of the epidemic.²

The solution to Eqs. (1) and (2) is the function of the absolute number of infected people, which takes the form:

¹ The model historically arose in 1931 to explain fluctuations in fish catch in the Adriatic Sea (Volterra, 1976). The same system of differential equations was proposed by Lotka a few years earlier (1924), but Volterra conducted a much more comprehensive analysis of this system.

² For the initial stage and the stage of the peak of infection, this parameter is assumed to be zero. In one of the following paragraphs, there is a solution for $\neq 0$ and the results of recent calculations for some countries.

$$N(t) = \hat{N} \cdot e^{\gamma(1-e^{-\alpha(t-\hat{t})})}, \quad (3)$$

where $\hat{N}$ is the number of infected at time $\hat{t}$ when the epidemic starts.

Both variables ($\hat{N}$, $\hat{t}$) are dependent near the conditional start of the epidemic, but they can change without changing the type of function (3) by renormalizing, taking into account the parameters α, γ . For $\alpha(t - \hat{t}) \ll 1$, Eq. (3) would be distributed exponentially $N(t) \cong \hat{N} \cdot e^{-\gamma\alpha(t-\hat{t})}$, where $\hat{N} \cdot e^{\gamma\alpha\hat{t}}$ is invariant. To differentiate epidemics from one-offs it is worth considering $\hat{N} = 100$, while $\hat{t}$ define as t_{100} . Therefore, for a normalized function of the number of infected people (3) $N(t) = 100 \cdot \exp\left(\gamma\left(1 - e^{-\alpha(t-t_{100})}\right)\right)$ the three contagion parameters should be solved: α, γ, t_{100} . These constants, which describe the development of the epidemic, should be determined in line with the dynamics of infection by minimizing the distance function, which will be discussed below.

The next reasonable stage is to define the peak outbreak parameters. The first important parameter is the moment of the peak contagion rate for (3), i.e. time t^* , where $N'(t)$ is at maximum. This can be solved by $N'(t) = 0$. So,

$$t^* = \hat{t} + \frac{\ln(\gamma)}{\alpha}, \quad (4)$$

which at its peak results in a maximum increase in the number of new infections per day.

$$\dot{N}^* = \hat{N} \cdot \alpha \cdot e^{\gamma-1}, \quad (5)$$

where $\dot{N}^*$ allows us to evaluate the maximum load on the healthcare system.

Assuming $\tilde{\beta} = 0$, from (2) one can forecast the maximum number of infected over the course of the epidemic:

$$N_{max} = \hat{N} \cdot e^{\gamma}. \quad (6)$$

The timing of the exit from containment measures (which should be highly correlated with the stages of epidemic ending) can be set up based on practical considerations, e.g., on the percentile p of the remaining infections until N_{max} is reached, so time T_p is calculated from the definition of the percentile,

$$p = \frac{N_{max} - \hat{N} \cdot e^{\gamma(1-e^{-\alpha(T_p-\hat{t})})}}{N_{max}}. \quad (7)$$

From (7) for small p one could easily infer a simple formula to define the date for easing restrictions:

$$T_p = t^* + \frac{1}{\alpha} \ln \left(\frac{1}{p} \right), \quad (8)$$

where t^* denotes the peak of the epidemic (4).

As a reasonable option, one could begin the relaxation of the regime of restrictions when $p=10\%$ (the end of the “high security regime”), and the second stage of easing when $p=1\%$ (the end of the self-isolation regime).

3 Validation of a parametric contagion model on Ebola outbreaks

Equation (3) was validated on data of the Ebola virus epidemic (WHO, 2016). Calculations are performed on a daily basis for each date T during the course of the epidemic, without relying on the information known after date T . The result is shown in Fig. 1. The graphs are aligned along the horizontal axis (identical for all graphs). In chart (A) the current forecast value of the peak infection date for the model is calculated on a daily basis, the dashed line is the actual peak, the straight dotted line is the current date of the forecast T . Chart (B) shows the forecast of the maximum number of infected (C_{max}) in all African countries, and the actual number of those who got the disease (dotted). Chart (C) shows the predicted and actual number of infections per day when all information is available.

The dynamics of the Ebola virus outbreak are well approximated by (3). Forecasts of the peaks of infections per day begin to approach their true values when the fluctuations subside. Approaching the values which explain the total number of infected (in the proximity of 10,000 people, or up to 30% of the true, ex-post measured number of infected) begins as the epidemic approaches the peak of infection. Before that, the forecasts are highly volatile and not significant.

One of the peculiarities of the approach is that it works well for outbreaks where the peak is over (see the Chart A). Earlier than that, the date of the epidemic peak might be less reliable.

Based on empirical observations of COVID-19 outbreaks, when the model starts to indicate that the peak is over, the true peak date should be already over for 7 days or more, and so a steady trend of the decreasing intensity of the infection rate begins.

4 Database description

For the purpose of this paper, we relied on data provided by the World Health Organization (2020), which monitors the outbreak on a daily basis. Country-wise information was obtained for the whole set of member countries available during the preparation of the paper. The data set consists of the total number of infected people in each country (for some countries the data was available for individual provinces or regions, e.g. China, the US) on a daily basis starting from January 22nd

Fig. 1 Modeling the Ebola virus epidemic

and ending June 29th, which are the first 160 days of observations of the COVID pandemic.

In addition, we used information on restrictions imposed by governments in different countries, which was collected by the Government Response tracker by Oxford Blavatnik School of Government (2020). This tracker is calculated by a team of 100 Oxford community members who have continuously updated a database of 17 indicators of government responses in 177 countries. These indicators examine containment policies such as school and workplace closings, restrictions on public events and public transport, and stay-at-home policies.

5 Indicators of deviation from the static model

Since the purpose of this study is to analyze the spread of the epidemic in different countries, it is necessary to propose indicators of the deviations of the actual dynamics from the static model. These indicators should show in which countries the observed dynamics differ significantly from the model, so that it is possible to give an interpretation of the systemic and non-systemic causes of the deviation from the “ideal” dynamics. Having identified these reasons, it is necessary to draw conclusions about the factors that lead to the unpredictability of the dynamics. In itself, such unpredictability is a significant cause of uncertainty both for the global economy and for individual countries.

We consider a daily increase in the number of infected people $C'(t)$ and its static model $N'(t)$, where $t \in [\hat{t}, T]$ is a date in a specified region or country. $\hat{t}$ is defined as a model starting date when the first $\hat{N}$ cases are registered (it might be 10, 100 or 1,000 depending on definition in the calculations). In fact, the starting date would not have any effect on the calculation result, it should be chosen in accordance with the scale of the region, country, city, etc. Details of model calibration are in “Appendix”.

The Hodrick-Prescott deviation measures the relative difference of the model $N'(t)$ from its HP-filtered curve $C'_\lambda(t)$ based on actual data of daily cases.

To calculate this deviation, data for $C'(t)$ is filtered from WHO (2020) with the Hodrick-Prescott methodology with the parameter $\lambda = 1600 \cdot 90^4$ for the period $t \in [\hat{t}, T]$ (as discussed in Hodrick & Prescott, 1997), yielding a time series $C'_\lambda(t)$. The deviation metric from the level model λ is given by the function:

- over the entire observation interval:

$$HP2(\lambda) = \frac{1}{C_{max}} \sum_{t=\hat{t}}^T |C'_\lambda(t) - N'(t)|, \quad (9)$$

- over the observation period after the peak of contagion is over:

$$HP3(\lambda) = \frac{\sum_{t=Peak}^T |C'_\lambda(t) - N'(t)|}{C(T) - C(Peak)}, \quad (10)$$

where Peak is the date of the peak of infection (4), C_{max} is the forecast of the maximal number of people infected according to model (6).

In Fig. 2 there are two countries with very different dynamics for the infection outbreak. One of the examples, Iran (chart A), describes a significant deviation of the modeled curve from the HP-filtered average, while the other case, Italy (chart B), describes an “ideal” coincidence of the model with filtered and actual data: the number of people actually infected is aligned with the modeled number.

The statistical data quality Q metric is necessary to define those countries in which the statistical data of the number of COVID-19 infected is collected improperly, which leads to unexplained statistical fluctuations in the number of new

Fig. 2 An example of a country with a significant deviation from the model (Iran, **A**) and a country with an almost perfectly modeled dynamic fit (Italy, **B**)

infections per day. When analyzing the deviations between the actual and model dynamics, such countries should be ignored (or, at least, treated with caution when judgments and conclusions are made) since such deviations are probably associated with significant statistical errors or local random outbreaks (as in China).

The metric shows the difference in the deviation of the actual number of cases per day $C'(t)$ from the HP-filtered curve, adjusted for statistical error in relative terms (%):

$$Q(\lambda) = \sqrt{\max\left(0, \frac{\sum_{t=i}^T \left[(C'(t) - C'_\lambda(t))^2 - C'_\lambda(t) \right]}{\sum_{t=i}^T C'_\lambda(t)^2}\right)}. \quad (11)$$

The problem of statistical data quality is very important, as it looks very poor even in such European countries as Spain and France (Fig. 3, charts A and B, respectively).

The Stringency Index metric shows how intensively the restrictions (self-isolation, quarantine measures, etc.) were relaxed following the peak of the epidemic in countries (which had passed an infection peak at the time of the study):

$$SI = \frac{\sum_{t=Peak}^T IS(t)}{T - Peak + 1}, \quad (12)$$

$$SI\% = \frac{1}{\max_{t \in [i, T]} IS(t)} \cdot SI. \quad (13)$$

The lower the SI percentage, the faster the relaxation of the restrictions. For the sake of the analysis, we concentrated on countries with a large number of infected and where one peak of infection has already passed. The metric is based on daily data about the Stringency Index (from Oxford, 2020), which is among the metrics used by the Oxford COVID-19 Government Response Tracker.

The Stringency Index is defined in a range between 0 and 100 reflecting these indicators. The higher the index score, the higher the level of restrictions.

The Stringency Index provides a gauge that allows us to consider the timing of enforcing and relaxing epidemic containment measures.³

³ Oxford provides an overlay of countries' death curves and their stringency scores. Some countries saw their deaths begin to flatten as they reached the highest stringency, such as Italy, Spain, and France. As China pulled stronger measures, its death curve plateaued. In countries such as the UK, the US, and India, the Oxford graphs find that the death curve did not flatten even after the strictest measures were enforced.

Fig. 3 Unexplainable data errors in measuring the number of people with COVID-19 in Spain and France

6 The static infection model of COVID-19 and its results

In a setup of the static model with zero residual infection rate ($\tilde{\beta} = 0$) we compared the predicted results and the actual dynamics for 90 countries in the WHO (out of 177). However, in the paper we present results for a lower number of countries, as we have applied a set of filters and criteria which excluded many of them from the applicability of the analysis. The first criterion was to exclude countries which might have not passed the first peak, this excluded 29 countries.

The average deviation of the forecasted peak on the basis of the static infection model from the HP-filtered model (which is generally very similar to the actual data) is 5 days (with a standard deviation of 18 days), and when countries with a strong second wave are excluded, the deviation between the HP-filtered peak and the actual peak falls to zero (with a standard deviation of 5 days). If the distances of a predicted peak from the HP-filtered data are taken at absolute values, the averages increase, while the accuracy improves. This highlights that the model reflects reality well, and therefore it is worth considering this model as a basis for future investigations into the topic.

For the analysis of events and trends that impact the dynamics of the COVID-19 spread over countries we applied a set of additional filters. For instance, we decided to exclude the countries with a low number of infected (less than 5000 people), which cut 11 other countries from the analysis, some of which are important for the future research (e.g. Hungary, which shows clear signals of a second wave—in line with the many other CEE countries). The idea behind that was to avoid situations in which the outbreaks are too small to make reasonable conclusions about.

At the next stage, we excluded countries with a low quality of statistics (high Q metric)—another nine locations. Some of them are worth monitoring in the future, such as France, China, Sweden, South Korea. Finally, we choose 20 countries with relatively a high HP metric (see Table 1).

The countries from Table 1 were analyzed to identify whether there is any relationship between the strength of the 2nd wave of the outbreak and the restrictions imposed by governments. The basic hypothesis was that easing restrictions would be directly related to the strength of the second wave of the coronavirus. Summarizing the findings of, for countries where the second wave is observed, the idea of a nexus between the speed of lifting restrictions and the speed of epidemic spread is supported in most cases (see Table 2).

For Afghanistan and Qatar, the evidence on the absence of a second wave might be not convincing, as the peaks of infections in those countries are too close to date when we stopped collecting the data (June 29th, 2020). Hopefully, those countries avoid the second wave, as the speed of lifting restrictions is low, but the situation should be monitored.

For countries with a second wave, the analysis suggests that the faster the speed of lifting restrictions, the higher the strength of the second wave. Moreover, the countries in the table can be additionally subdivided according to the following approach. Some countries with a weak or very weak second wave (these

Table 1 Countries analyzed and their key metrics

Country/location	Actual number of already infected as of 29 June 2020	% of infected in total population (%)	Forecasted number of infected	Peak date forecast	Date of first 100 infected (t_100 date)	Forecasted % of infected in total population (%)	HP metric (%)	IS metric (%)	Q metric (%)
Qatar	95,106	3.93	142,082	28.05.2020	26.03.2020	5.87	13.1	80.3	12.8
Norway	8862	0.17	8563	24.03.2020	03.03.2020	0.16	14.7	72.1	35.6
Spain	248,970	0.54	241,168	30.03.2020	03.03.2020	0.52	15.4	85.4	49.6
Netherlands	50,433	0.30	49,112	05.04.2020	05.03.2020	0.29	16.9	88.7	19.7
Iran	225,205	0.29	378,403	22.05.2020	06.01.2020	0.48	17.2	61.4	14.8
Afghanistan	31,238	0.12	50,233	04.06.2020	13.04.2020	0.19	17.6	72.0	19.3
Germany	195,042	0.24	186,489	30.03.2020	02.03.2020	0.23	19.4	81.0	28.9
Austria	17,723	0.20	16,627	25.03.2020	08.03.2020	0.19	23.0	74.8	40.9
Algeria	13,571	0.03	17,296	13.05.2020	11.03.2020	0.04	23.9	80.8	14.3
Ukraine	44,538	0.10	87,369	12.06.2020	01.03.2020	0.20	28.0	38.9	17.0
Turkey	198,613	0.26	186,870	15.04.2020	08.03.2020	0.24	28.8	89.5	17.3
US	2,590,552	0.80	2,884,897	01.05.2020	08.02.2020	0.89	29.4	90.7	11.6
Malaysia	8637	0.03	8644	03.04.2020	28.02.2020	0.03	29.7	79.4	42.4
Romania	26,582	0.13	27,153	24.04.2020	03.03.2020	0.14	32.1	77.5	24.4
Australia	7834	0.03	7161	25.03.2020	13.03.2020	0.03	37.2	80.2	49.4
Serbia	14,288	0.20	12,723	14.04.2020	19.03.2020	0.18	37.3	60.6	31.3
Portugal	41,912	0.40	39,933	12.04.2020	26.02.2020	0.38	42.2	80.2	34.9
Czechia	11,805	0.11	10,244	03.04.2020	02.03.2020	0.10	46.2	66.4	37.5
Morocco	12,290	0.04	12,460	02.05.2020	11.03.2020	0.04	49.7	92.5	47.7
Israel	24,441	0.29	19,222	03.04.2020	08.03.2020	0.23	59.0	75.1	40.9

Source: Authors' calculations

Table 2 Comparison of the outbreak dynamics in selected countries with the Index of Stringency (summary)

Country	Second wave (Y/N)	Strength of the 2nd wave	Policies regarding restrictions	Evidence
Afghanistan	N		Lifting restrictions	
Norway	N		Lifting restrictions	
Qatar	N		Almost no lifting	
Malaysia	Y	Very weak	Almost no lifting	Supportive
Netherlands	Y	Very weak	Somewhat lifted	Supportive
Spain	Y	Very weak	Lifted	Non-convincing
Germany	Y	Weak	Somewhat lifted	Supportive
Austria	Y	Weak	Lifted	Non-convincing
Australia	Y	Weak	Lifted	Non-convincing
Turkey	Y	Strong	Almost no lifting	Non-convincing
Iran	Y	Strong	Almost no lifting	Non-convincing
US	Y	Strong	Almost no lifting	Non-convincing
Israel	Y	Strong	Somewhat lifted	Non-convincing
Portugal	Y	Strong	Somewhat lifted	Supportive
Ukraine	Y	Strong	Somewhat lifted	Supportive
Algeria	Y	Strong	Somewhat lifted	Supportive
Morocco	Y	Strong	Somewhat lifted	Supportive
Serbia	Y	Strong	Lifted	Supportive
Czechia	Y	Strong	Lifted	Supportive
Romania	Y	Strong	Lifted	Supportive

Source: Authors' calculations

are the countries where the number of infected in the second wave is substantially less than the number of infected at the first peak)—Spain, Australia and Austria—do not fit into the overall picture well. However, for the last two countries the speed of the second wave compared to the minimal number of infections after the peak is substantially growing, so it might be the case that the second wave would be “strong”, and so they will fit the hypothesis.

In countries with a strong second wave (i.e. the number of maximal infected in the last two weeks of June is more than 40% higher than the previous peak), e.g. Iran, Israel, Turkey and the US, the link between lifting restrictions and the probability of the second wave might be weak. However, these cases seem to be very special for the following reasons: Middle-Eastern countries might have a somewhat different pattern of behavior in personal communication between people, the density of the population in cities, and the large proportion of small street retail businesses. Those two aspects might have impacted the link under investigation, although future research into formal relations between social and economic factors which might undermine the relation is needed. In the US, the situation is very different from county to county, and deeper research is required,

Fig. 4 Outbreak waves and Index of Stringency dynamics in CEE countries

yet, between the first and the second wave the country faced a series of mass public events when thousands of people met in the streets while the epidemic outbreak might not have been over.

The link between substantially eased restrictions and the increase in the number of infected during the second wave is specifically strong in some CEE economies (Fig. 4): when the restrictions are eased rapidly, the outbreak increases within a few weeks, and the strength of the second wave seems to be very powerful. This is even more important as the modeled dynamic of the outbreak implies that the number of infected should approach zero instead of a second wave occurring.

7 The residual infection model of COVID-19 and its results

We solve differential Eqs. (1) and (2) for the normalized $\tilde{\beta} = \beta \cdot \hat{N} \cdot \alpha \cdot e^{\gamma-1} > 0$. To obtain an absolute number of infected $N(t)$ and the speed of the daily contagion $N'(t)$ one uses the following equations:

$$\begin{aligned} N(t) &= \hat{N} \cdot e^{\gamma(1-e^{-\alpha(t-\hat{t})})} \cdot \left[1 + \frac{\beta}{e} \left(Ei(\gamma) - Ei(\gamma) \cdot e^{-\alpha(t-\hat{t})} \right) \right] N'(t) \\ &= \gamma \cdot \alpha \cdot e^{-\alpha(t-\hat{t})} \cdot N(t) + \beta \cdot \hat{N} \cdot \alpha \cdot e^{\gamma-1} \\ Ei(x) &= \ln \ln(x) + \sum_{k=1}^{\infty} \frac{x^k}{k \cdot k!} \end{aligned} \quad (14)$$

Equation (14) contains a vector of four unknown parameters $\vec{p} = (\alpha, \beta, \gamma, \hat{t})$, which are defined by the optimization (minimization) of a metric function $F(\vec{p})$ (see “Appendix”) under the initial approximation of $\beta = 0$.

Our calculations during the preparation of this paper suggest that β depends significantly on the current phase of the pandemic. Therefore, for some countries β may still not be zero, which would result in a predicted increase in the number of cases (which, in theory, is unlimited if previously infected people do not have immunity) after the successful completion of the main phase. From a practical point of view, this should indicate caution in relaxing restrictions in this country. For example, at the time of writing, the international tourist season was open in Turkey, yet our model implies one of the world’s highest residual infection levels, which might result in new outbreaks in other countries (Fig. 5).

As of July 19th, the formulae (14) suggest that high residual infection rates will be observed in countries like Turkey, Ecuador, Romania, Portugal, Bosnia and Herzegovina, Belarus (ranked descending vs. the peak).

The level of residual infection is needed to serve as a starting value to implement this model for calculating the second and subsequent waves of the outbreak. This can be done by simply inputting $\hat{N}$, $\hat{t}$ into (2) and (3) as the initial conditions of the new wave.

Fig. 5 The case of Turkey: a country with constant residual infection rate, calculated and forecasted as of 19th July 2020

8 Conclusions and directions for future research

The paper proposed a COVID-19 impact dynamic model, whose result is a high quality prediction of the dates and scales of peaks as compared to actual data about infections all over the world for countries with reliable data.

It is the first wave of a pandemic which should attract a special fundamental scientific interest, because it is usually free from impacts of previous socio-economic measures to contain outbreaks, and from re-infection from other regions and countries, vaccinations, and other effects. The importance of such analysis is usually underestimated by the scientific community despite it connecting the efficiency of containment measures and the scale of outbreaks. Our approach to modeling the spread of a pandemic proposed and substantiated in this article has the value of providing a rapid and continual measure of whether restrictions should be lifted, as opposed to classical models (SIR and SIS) which require a much larger data set with revealed peaks.

The research provides convincing evidence that the proposed model predicts the dates of peaks and the scale of outbreaks well, and so might help avoiding the earlier-than-needed lifting of restrictions. Most important is that the model has been developed before the end of the pandemic, so although future data might impact both the results and conclusions, this continual modeling is a solid basis for future research in epidemiological studies.

The model has revealed a nexus between the appearance of the second wave of the pandemic and the easing of restrictions. The paper can also serve as a basis for hypothesizing about the quantitative analysis of this nexus. Such an analysis might help to design policy regarding the timing and speed of the easing of restrictions against such pandemics. We leave the testing of the quantitative hypotheses for future investigations.

As this modeling is done while the statistics have not been collected for all countries, it is difficult to judge whether any nation has already passed the peak of infection, the conclusions might change with the new information.

The results for specific countries obtained in this work at the end of June 2020 may not be relevant at the time of its publication, but forecasting and warning methods based on the presented model will be relevant when choosing the intensity and socio-economic methods of suppressing possible future pandemics.

The generalization of the model on the basis of residual infection would help in identifying effective criteria to recognize the appearance of the second and subsequent waves of COVID-19 infections. It seems that the waves would follow the same law of contagion as the first wave, but with different parameters.

Despite this, the model requires future development in several ways, e.g., the implementation of updated empirical information as well as new socio-economic and medical instruments for pandemic containment. It also needs more careful verification for a wider exploration for policy advice. One possible direction could be research into the links between the proposed model and risk-management models including systemic risk models (Springer, 2021). In addition to methods referred to or presented in this paper, it is worth considering a comparison with

machine learning and artificial intelligence-based approaches in a manner similar to that conducted in Karminsky and Burekhin (2019).

Appendix 1 Model calibration during the contagion period with actual data

Using the approach presented in Sect. 3, it is possible to monitor and predict the most important characteristics of the contagion during its outbreak, updating the parameters on a daily basis.

The criteria to determine the parameters of the epidemic is the distance to the empirical infection curve ($C(t_i)$ $i = 1 \dots T$, where T is the date of the current observation) in the RMS form, also taking into account infection rates. Hence the dynamic model should be defined as

$$F(\vec{p}) = \sum_{i=1}^T \frac{(C(t_i) - N(t_i, \vec{p}))^2}{C(T)^2} + \lambda \cdot \sum_{i=2}^T \left(C(t_i) - C(t_{i-1}) - \dot{N}\left(t_i - \frac{1}{2}, \vec{p}\right) \right)^2 \cdot \frac{(i-1)^2}{C(t_i)^2},$$

where $\vec{p}$ is the vector of parameters ($\alpha, \gamma, \hat{t}$), t_i are daily data publication dates, t_1 is the first date when $\hat{N}$ of infected ($C(t_i) \geq \hat{N}$, e.g., $\hat{N} = 100$), $C(t_i)$ is the number of infections, T is the current date of infection outbreak.

The parameter λ is chosen empirically and characterizes the weight smoothing rate of infections by function (3) in Sect. 2. The authors' experience, on the basis of empirical calculations, suggests $\lambda \in (1; 4)$. We suggest using the following initial values to find the minimum distance:

$$\alpha = 0, \gamma = \ln \left(\frac{C(T)}{\hat{N}} \right), \hat{t} = t_1.$$

This initial approximation means no further contagion and is suitable for starting the convergence to the correct solution.

References

- Choisy, M., Guégan, J.-F., & Rohani, P. (2007). *Mathematical Modeling of Infectious Diseases Dynamics*. Chapter 22, *Encyclopedia of Infectious Diseases: Modern Methodologies*, by M. Tibayrenc. Wiley. <https://doi.org/10.1002/9780470114209.ch22>.
- Hethcote, H.W. (1989). Three basic epidemiological models. In S.A. Levin, T.G. Hallam, L.J. Gross (Eds.), *Applied Mathematical Ecology* (pp. 119–144). Springer. <https://doi.org/10.1007/978-3-642-61317-3>.
- Hodrick, R., Prescott, E. C. (1997). Postwar U.S. business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, 29(1), 1–16. <http://www.jstor.org/stable/2953682>.
- Karminsky, A. M., & Burekhin, R. N. (2019). Comparative analysis of methods for forecasting bankruptcies of Russian construction companies. *Business Informatics*, 13(3), 52–66. <https://doi.org/10.17323/1998-0663.2019.3.52.66>.
- Oxford. (2020). *Our World in Data*. <https://ourworldindata.org/grapher/covid-stringency-index?tab=table>.
- Pomazanov, M.V. (2020). Modeling of COVID-19 pandemic indices and their relationships with socio-economic indicators. In *6th International Scientific Conference titled: Knowledge Based Sustainable Development – ERAZ, May 21, 2020*. <https://eraz-conference.com/#moreinfo> (in printing).

- Springer. (2021). Risk assessment and financial regulation in emerging markets' banking, trends and prospects. In A. Karminsky, et al. (Eds.). Springer Publishing House (**in printing**).
- Volterra, V. (1976). *Mathematical Theory of Struggle for Existence*. Nauka. [https://www.scirp.org/\(S\(i43dyn45teexjx455qlt3d2q\)\)/reference/ReferencesPapers.aspx?ReferenceID=1762958](https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=1762958).
- WHO. (2020). *Coronavirus disease (COVID-2019) situation reports*. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.
- WHO Situation Reports. (2016). *Number of Cases and Deaths in Guinea, Liberia, and Sierra Leone during the 2014–2016 West Africa Ebola Outbreak*. <https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/case-counts.html>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.