GENERAL PROBLEMS OF TRANSPORT THEORY

PHYSICAL KINETICS AND SIMULATION
OF THE SPREAD OF AN EPIDEMIC

P. S. Grinchuk and S. P. Fisenko

UDC 533.72+519.218.28

A self-consistent system of differential logistic equations has been constructed to describe the spread of an epidemic with account for the demography and geography of a country. An original method of numerical assessment of the number of unregistered cases in the country has been proposed.

Keywords: number of contacts, demography, geography, time of infection, similarity parameter.

Introduction. There are qualitative analogies between the problem of physical and chemical kinetics and the theory of propagation of epidemics in diverse populations of living beings [1, 2]. This makes it possible to transfer the experience accumulated in theoretical investigation of kinetic problems to epidemiology. Mathematical simulation of the spread of epidemics on the basis of a logistic differential equation or its discrete analogs has long been an important component of epidemiology and allows simulation of epidemic behavior [3, 4]. It should be pointed out that, due to the Covid-19 coronavirus epidemic, there is an accumulation of increasingly more publicly available statistical information on the spread of epidemics in different countries. It is already obvious that classical mathematical models of epidemics' propagation are rather rough and describe the observed phenomena only qualitatively.

The goal of this investigation is to develop a mathematical model of the spread of an epidemic with account for the geographic and demographic features of a country. Numerical results of simulation based on our model are given for the Republic of Belarus (RB) using the example of the Covid-19 coronavirus epidemic with processing of available statistical data [5].

"Zero" Approximation. Let us first consider the spread of the epidemic within the standard mathematical "zero-dimensional" approach when a country is characterized only by the size of the population N , and the spread of the epidemic is only determined by two coefficients k and b . In this case, the propagation of the epidemic is described by one ordinary logistic-type differential equation [3]

$$\frac{dy}{dt} = kyb \left(1 - \frac{y}{N} \right). \quad (1)$$

For the constant coefficients k and b , the exact solution of Eq. (1) with the initial value $y(0) = 1$ has the form

$$y(t) = \frac{N}{(N-1) \exp(-kbt) + 1}. \quad (2)$$

For short times t , expression (2) can be reduced to the form

$$y(t) \approx \frac{N}{N-1} \exp(kbt).$$

This expression makes obvious the exponential nature of the growth in the number of infected persons with time at the beginning of the epidemic. As follows from (2), the characteristic time of the spread of the epidemic is $\tau = 1/kb$. We designate the time during which, in practice, the entire population of the country becomes infected as t^* . From formula (2) we can obtain

A. V. Luikov Heat and Mass Transfer Institute, National Academy of Sciences of Belarus, 15 P. Brovka Str., Minsk, 220072, Belarus; email: fsp@hmti.ac.by. Translated from Inzhenerno-Fizicheskii Zhurnal, Vol. 94, No. 1, pp. 3–8, January–February, 2021. Original article submitted July 1, 2020.

Fig. 1. Variation in the number of cases in Belarus with time calculated in a "zero-dimensional" approximation at $kb = 3$ (1), 1 (2), and 0.3 (3).

Fig. 2. Infection rate in the population of Belarus calculated in a "zero-dimensional" approximation at $kb = 0.3$.

$$t^* = \ln N\tau. \quad (3)$$

For a country with a population of 10 million people, from expression (3) it follows that $t^* \cong 16.1/kb$.

Since in Eq. (1), use is made of the coefficients k and b averaged over the entire population, it follows from this equation that fluctuation manifestations are inevitable in the spread of the epidemic. These manifestations are obvious if we are to study primary statistics [5, 6]. Additionally, we can point out that the coefficient b fluctuates in time with a period of about 7 days, which is due to the organization of life in modern human society. Fluctuation effects in epidemic propagation are evident and important when the right side of Eq. (1) is of the order of unity, i.e., at the beginning and end of the spread of an epidemic. This circumstance is well known [4]. It is important to note that, as follows from the derivation of Eq. (1), the parameter b is directly proportional to the product of the population density by its mobility. It is precisely for this reason that epidemics spread faster in large urban centers (megacities) where the high population density is due to a significant number of high-rise apartment buildings. In this case, the high mobility of the population is ensured by a developed public transport system. By analogy with problems of physical kinetics, the number of collisions of one molecule with others in the gas is directly proportional to the average velocity of this molecule's motion and the volumetric density of the molecules in the gas volume.

As a result of processing statistical data on the dynamics in the growth of the number of Covid-19 cases in Minsk at the beginning of the epidemic by the method of running medians smoothing, we obtained the value $kb \approx 0.3$ [5]. In the English-language scientific literature, the product kb is called reproduction number and is designated as R_0 [6]. Note that, based on statistical data obtained during the study of the spread of influenza epidemics in the USA, it was established that the product $kb = 2.7$. It seems that during the current Covid-19 pandemic, the widely publicized and advocated measures of personal hygiene and social distancing reduce the number of contacts and also the likelihood of infection on contact.

Figure 1 shows the results of calculation of the number of infected persons as a function of the value of the kb product for the Republic of Belarus. The kb value is, in fact, a similarity parameter for the problem of epidemic propagation. This figure shows that an epidemic can last for 7–70 days and, consistent with an analytical estimate, it is inversely proportional to the kb parameter. It is important to note that, according to the mathematical model described by Eq. (1), the peak in the epidemic growth rate occurs when half of the population is infected, irrespective of the value of the model parameters kb and N . It is useful to use this note for analysis of available statistical data.

Figure 2 shows the infection growth rate among the population of the Republic of Belarus, whose size is about 10 million people, calculated within the framework of zero-dimensional approximation for the case $kb = 0.3$. It is seen from this figure that the main number of cases occurs in the first 20 days. What stands out here is that the infection rate curve calculated by Eq. (1) is symmetrical (has the shape of a parabola) with regard to the position of the maximum u . It is important to stress that the available statistical data on all countries yield a nonsymmetrical infection rate curve.

Average Number of Contacts. Comparison of the results of calculation by Eq. (1) with the available statistical data on the dynamics in registration of cases is, in our view, a rather difficult task. It is obvious that the number of infected persons

is much higher than the number of registered cases. According to epidemiologists, these two numbers may in some cases differ by several orders of magnitude [7]. It is necessary to conduct special experiments (screening) to identify the number of infection cases in a rather large random representative sample from among the residents [6]. Such investigations with a sample of 70,000 people are known to have been started in Moscow in May 2020.

In our view, for the Republic of Belarus, it is only possible to obtain the lower estimate of the parameter kb from the growth dynamics in the number of registered infected individuals. We drew a conclusion on the value $kb \approx 0.3$, based on the solution of an inverse problem using statistical data on the number of registered cases. Therefore, we can argue that all our calculations of the duration of the epidemic spread in the Republic of Belarus are overestimated due to its exponential dependence on the kb product. Note that after a long lockdown period in many German states, the estimate $kb = 0.9$ was obtained [6]. Nevertheless, we can argue that, if the epidemic spread rate in a subpopulation of a compact community has reached its peak, a substantial share of the population (about 50%) will be infected. It is obvious that the "zero-dimensional" mathematical model of the epidemic spread is rather rough, since it does not reflect the geography and demography of a country (population size, residents, settlement structure, and the areas of the country's cities and towns). At the same time, this model captures quite accurately the kinetics of infection transmission.

First Approximation. To take a more accurate account of the geography and demography of a country in a classical model of the spread of an epidemic and to show possible variants of its development, we consider epidemic propagation using the example of the Republic of Belarus. Taking into consideration the geography of the country and its demography, we break down the population of the Republic of Belarus into three subpopulations. The first subpopulation includes Minsk (a population of 2 million people as of 2020) and Gomel (500,000 people). The second subpopulation includes the remaining provincial (oblast) capitals (Brest, Grodno, Mogilev, and Vitebsk with a total population of about 1,400,000 people), and the third subpopulation includes the residents of small towns and villages which currently constitute a substantial majority of the country's population. Such a description of the country's population with account for the geography and demography of the country is the next approximation in relation to the "zero-dimensional" one and, in our view, can be applied to an overwhelming majority of countries.

It is important to point out that the population density in Minsk is 2.2 times higher than on the average in the oblast (province) capitals [6]. Thus, whereas for Minsk $kb \approx 0.3$, for the provincial capitals, we will have $kb = 0.136$. For the third group, the product kb will be lower than for the provincial capitals. It is difficult to identify this parameter with good accuracy. It is assumed in our calculations that for small towns and villages, the quantity kb is two times lower than for the second subpopulation, i.e., for the third subpopulation $kb = 0.068$. Naturally, for each group, use is made of its own type-(1) equation with certain values of the parameters it includes. In this case, the total size of all the subpopulations is assumed to be equal to the population of the country. As shown by our numerical experiments, the initial conditions for the above-mentioned system of equations can be quite random. The numerical calculations were conducted for the number of infected individuals in the first subpopulation, viz. $4 \cdot 10^6$, in the second subpopulation, viz. $2 \cdot 10^6$, and in the third subpopulation, viz. $4 \cdot 10^6$. In the first approximation, we ignored the epidemiological "interaction" between the three subpopulations of the country's residents and their movement beyond the bounds of each subpopulation. In other words, it was assumed that the population was only mobile inside their own local community.

Figure 3 shows the variation in the number of infected individuals with time calculated within our first approximation for the Republic of Belarus. It is obvious that the epidemic spreads nonuniformly across the country. Due to the large share of the third subpopulation, the main mass of the population becomes infected rather late compared to the first subpopulation. At $kb \approx 0.3$ for the first subpopulation (country's capital), the total duration of the epidemic in the Republic of Belarus is expected to last about 500 days. Figure 4 shows the epidemic growth rate for the first subpopulation of the country. It is seen that the maximum epidemic growth rate is significantly lower compared to the case presented in Fig. 2. The peak width at half the maximum amplitude is 13 days. Figure 5 shows a trimodal curve of infection growth per day. What calls for attention is that the second peak is nonsymmetrical, as indeed are most of the epidemic growth rate peaks in all countries [7]. In our view, the empirical data on the number of registered infection cases are, in fact, envelopes for a great number of local peaks. In our case, the curve is a sum of three symmetrical curves of different height with different positions of the maxima. For other countries, the number of subpopulations into which a country's population is broken down can be significantly different. In particular, in countries with high density and mobility of the population, like Italy or Germany, the number of subpopulations is lower. The first highest peak in the epidemic growth rate is largely due to its spread in the largest cities of the Republic of Belarus, i.e., in the first subpopulation. The halfwidth of the peak for this subpopulation is about 13 days.

Fig. 3. Variation in the number of cases in Belarus with time since the epidemic onset.

Fig. 4. Epidemic spread rate in the first subpopulation.

Fig. 5. Epidemic spread rate in Belarus with account for subpopulations.

Fig. 6. Numerical estimate of the total number of COVID-19 cases in Belarus.

The duration of the epidemic in Belarus as a function of the number of initially infected people was investigated numerically at $kb = 1$. The calculations have shown that the epidemic duration is basically determined by the time of its propagation in the third subpopulation. In this case, the duration of the epidemic, when the number of the initially infected people (initial conditions) varies within the range of 10 to 100 persons, is 70 ± 7 days for all the three populations.

As pointed out above, statistical data are only available on the number of registered cases. A question arises: how can we estimate the total number of infected people, including symptom-free cases? It appears reasonable to use the following technique. For any initial number of infected people but fewer or equal to 100, a calculation is conducted within the framework of our model at $kb = 1$ for the first subpopulation (the largest cities). As shown by the European experience in 2020, the epidemic spread coefficient drops to a value quite close to unity after a long and strict limitation of social contacts. Then the number of infected people calculated using the model's equations will yield a reasonable estimate of the total number of cases in the country at the instant of time t . At the beginning of the epidemic, this estimate will naturally be understated.

Figure 6 shows the total number of infected people in the Republic of Belarus calculated for the first subpopulation based on our model at $kb = 1$. Note that, according to [5], on the 30th day of the epidemic spread, the number of registered cases was about 14,000 people, and the total number of infected individuals calculated by our model is 2.01 million people. The ratio of the total number of infected persons to the number of registered cases can be estimated as 142. This ratio is not constant and varies slowly with time. This is a distinctive feature of the Covid-19 spread in the Republic of Belarus.

Conclusions. A method has been proposed to account for geographic and demographic features of a country in simulating the spread of an epidemic in it. In particular, a generalization of a classical mathematical model to the Republic of Belarus represents a system of three self-consistent ordinary differential logistic-type equations with their coefficients.

The coefficients in the equations take into account the dependence of the similarity parameter of the problem of epidemic spread on the population density in each of its group. The curve of the population infection rate in Belarus obtained using our approach is widened, and its peak has a lower value compared to a similar curve constructed under a classical model. The curve of the population infection rate in Belarus constructed within the framework of our model is trimodal.

A numerical investigation has been made into the sensitivity of the obtained results to the initial problem conditions for each of the three subpopulations of the country. It has been shown that as the initial number of cases varies in each population group, the total duration of the epidemic may only differ by a few days. On the whole, the duration of the epidemic is determined by its spread in a subpopulation for which the product kb has the lowest value. For Belarus this subpopulation is also the most numerous.

Based on the developed mathematical model, a computational method has been proposed for estimating the number of infected people in Belarus at random instants of time since the start of the epidemic. It turned out that the calculated number of cases exceeds the number of infectees (registered cases) by approximately two orders of magnitude and varies rather insignificantly with time.

Acknowledgment. The investigation was carried out with support from the Belarusian Republican Foundation for Fundamental Research.

NOTATION

b , average number of contacts per day for an average individual; k , total-population-averaged probability of infection transmission from a sick individual to a healthy person; N , population size of a country; t , time, day; u , infection rate, number of cases per day; y , number of cases; τ , characteristic time of epidemic spread, day.

REFERENCES

1. Ya. B. Zel'dovich, *Selected Works*, Vol. 1 [in Russian], Nauka, Moscow (1984).
2. A. N. Kolmogorov, I. G. Petrovskii, and N. S. Piskunov, Investigation into the equation of diffusion combined with the substance growth and its application to one biological problem, *Byull. MGU, Ser. A, Mat. Mekh.*, **1**, No. 6, 1–26 (1937).
3. A. M. Nakhushiev, *Equations of Mathematical Biology* [in Russian], Vysshaya Shkola, Moscow (1995).
4. R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, *Rev. Mod. Phys.*, **87**, 925 (2015).
5. P. S. Grinchuk and S. P. Fisenko, *Two Waves in the Dynamics of Coronavirus Rate in the Republic of Belarus*, Preprint; DOI: 10.24108/preprints-3112048.
6. M. Meyer-Hermann et al., Estimate of the development of the epidemic reproduction number R_t from Coronavirus SARS-CoV-2 case data and implications for political measures based on prognostics; doi.org/10.1101/2020.04.04.20053637 medRxiv preprint.
7. M. Lipsitch, Who is immune to the Coronavirus? *The New York Times*, 14.04.2020.
8. *Population of the Republic of Belarus*; ru.wikipedia.org.
9. *Database of The New York Times*, March–June 2020; free access.