

## Estimates of the COVID-19 pandemic dynamics in Ukraine based on two data sets

Igor Nesteruk

Institute of Hydromechanics, National Academy of Sciences of Ukraine  
National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.  
[inesteruk@yahoo.com](mailto:inesteruk@yahoo.com)

### ABSTRACT

**Background.** To simulate how the number of COVID-19 cases increases versus time, various data sets for the number of new cases and different mathematical models can be used. Since there are some differences in statistical data, the results of simulations can be different. Complex mathematical models contain many unknown parameters, the values of which must be determined using a limited number of observations of the disease over time. Even long-term monitoring of the epidemic may not provide reliable estimates of its parameters due to the constant change of testing conditions, isolation of infected and quarantine. Therefore, simpler approaches are necessary. In particular, previous simulations of the COVID-19 epidemic dynamics in Ukraine were based on smoothing of the dependence of the number of cases on time and the generalized SIR (susceptible-infected-removed) model. These approaches allowed to detect the waves of pandemic and to make adequate predictions of their duration and final sizes. In particular, eight waves of the COVID-19 pandemic in Ukraine were investigated.

**Objective.** We will compare the results simulation of a new epidemic wave in Ukraine based on national statistics and data reported by Johns Hopkins University (JHU).

**Methods.** In this study we use the smoothing method for the dependences of the number of cases on time, the generalized SIR model for the dynamics of any epidemic wave, the exact solution of the linear differential equations, and statistical approach developed before.

**Results.** Ninth epidemic wave in Ukraine was simulated. The optimal values of the SIR model parameters were calculated and compared with the use of two data sets. Both predictions are not very optimistic: new cases will not stop appearing until June-July 2021.

**Conclusions.** New waves of COVID-19 pandemic can be detected, calculated and predicted with the use of rather simple mathematical models. The results of calculations depend on the data sets for the number of confirmed cases. The expected long duration of the pandemic forces us to be careful and in solidarity. The government and all Ukrainians must strictly adhere to quarantine measures in order to avoid fatal consequences. Probably the presented results could be useful in order to estimate the efficiency of future vaccinations.

**Keywords:** COVID-19 pandemic, epidemic dynamics in Ukraine, mathematical modeling of infection diseases, SIR model, parameter identification, statistical methods.

**NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

## Introduction

The studies of the COVID-19 pandemic dynamics in Ukraine are presented in [1-15] and summarized in the book [16]. Different simulation and comparison methods were based on official accumulated number of laboratory confirmed cases [17, 18] (national statistics). These figures coincides with the official WHO data sets, [19]. Unfortunately, WHO stopped to provide the daily information in August 2020.

The classical SIR model [20-22], connecting the number of susceptible  $S$ , infected and spreading the infection  $I$  and removed  $R$  persons, was applied in [2-5, 8-10] to simulate the first pandemic wave in Ukraine. The unknown parameters of this model were be estimated with the use of the cumulative number of cases  $V=I+R$  and the statistics-based method of parameter identification developed in [23, 24].

The weakening of quarantine restrictions, changes in the social behavior and probably in the coronavirus activity causes change in SIR characteristics and the epidemic dynamics. To detect these changes, a simple method of numerical differentiations of accumulated number of cases was proposed in [11, 15, 16]. To simulate these new pandemic waves, the SIR model was generalized in [12, 15, 16]. In [12, 14, 16] the results of simulation of the first six epidemic waves in Ukraine are presented with the use of a procedure for sequentially determining the parameters of the model for each epidemic wave, starting with the first one.

This method requires considerable effort and time. The book [16] introduced a new algorithm for determining the optimal parameter values for a particular epidemic wave without calculating the dynamics of previous waves and presented calculations for the seventh epidemic wave in Ukraine. The eighth epidemic wave in Ukraine was simulated in [15] with the use of this approach. In this paper, we will analyze the dynamics of the epidemic in Ukraine in the period from September 1 to December 20, 2020, calculate the parameters of the eighth wave and make some predictions.

Since the national statistics does not look complete (see, e.g., the results of total staff testing in two schools and two children gardens in the Ukrainian city of Chelnytskii, [25]), it is very interesting to compare different data sets and simulations results based on different statistical information. For this purpose we will use the data reported by Johns Hopkins University (JHU) [26]. The results of smoothing two data sets and SIR simulations will be presented.

## Materials and Methods

### *Data*

Day in December 2020, $t_j$	Number of cases, national statistics, [17,18] $V_{j1}$	Number of cases, JHU, [26] $V_{j2}$	Day in January 2021	Time in days $t_j$	Number of cases, national statistics, [17,18] $V_{j1}$	Number of cases, JHU, [26] $V_{j2}$
1	758264	765117	1	32	1069517	1096855
2	772760	778560	2	33	1074093	1102256
3	787891	793372	3	34	1078251	1107137
4	801716	808828	4	35	1083585	1111631
5	813306	822985	5	36	1090496	1117256
6	821947	834913	6	37	1099493	1124482
7	832758	843898	7	38	1105169	1133802
8	845343	855054	8	39	1110015	1139800
9	858714	867991	9	40	1115026	1144943
10	872228	881727	10	41	1119314	1150265
11	885039	895620	11	42	1124430	1154850
12	894215	908839	12	43	1130839	1160243
13	900666	918444	13	44	1138764	1166958
14	909082	925321	14	45	1146963	1175343
15	919704	934161	15	46	1154692	1183963
16	931751	945218	16	47	1160682	1192114
17	944381	957692	17	48	1163716	1198512
18	956123	970758	18	49	1167655	1201894
19	964448	982937	19	50	1172038	1206125
20	970993	991700	20	51	1177621	1210854
21	979506	998678	21	52	1182969	1216780
22	989642	1007627	22	53	1187897	1222459
23	1001132	1018199	23	54	1191812	1227723
24	1012167	1030125	24	55	1194328	1231965
25	1019876	1041583	25	56	1197107	1234772
26	1025989	1049717	26	57	1200883	1237810
27	1030374	1056265	27	58	1206412	1241863
28	1037362	1061074	28	59	1211593	1247674
29	1045348	1068476	29	60	1216278	1253127
30	1055047	1076880	30	61	1219455	1258093
31	1064479	1086997	31	62	1221485	1261546

**Table 1. Cumulative numbers of confirmed Covid-19 cases in Ukraine:  $V_{j1}$  -National statistics, [17, 18] and  $V_{j2}$  - according to COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU), [26].**

We will use two data sets regarding the accumulated numbers of confirmed COVID-19 cases in Ukraine: the official information from national sources [17,18] -  $V_{j1}$  and data set  $V_{j2}$  from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) [26]. These values and corresponding moments of time  $t_j$  (measured in days, zero point is November 30, 2020) are shown in Table 1.

It must be noted that the data presented in Table 1 does not show all the COVID-19 cases in Ukraine. Many infected persons are not identified, since they have no symptoms. For example, employees of two kindergartens and two schools in the Ukrainian city of Chmelnytskii were tested for antibodies to COVID-19, [25]. In total 292 people work in the surveyed institutions. Some of the staff had already fallen ill with COVID-19 or were hospitalized. Therefore, they were tested and registered accordingly. In the remaining tested 241 educators, antibodies were detected in 148, or 61.4%. Many people know that they are ill, since they have similar symptoms as other members of families, but avoid to make tests. Unfortunately, one laboratory confirmed case can correspond to several other cases which are not confirmed and displayed in the official statistics.

There are also significant differences in two data sets visible in Table 1. The numbers of cases reported by JHU in January 2021 are 30,000-40,000 higher than presented by the national statistics [17, 18]. However, both data sets are probably incomplete. Next, we will make a comparative analysis of these data and conduct appropriate SIR simulations.

### ***Detection of epidemic waves***

To control the changes of epidemic parameters, we can use daily numbers of new cases and their derivatives. Since these values are random, we need some smoothing. For example, we can use the smoothed daily number of accumulated cases proposed in [11, 12, 14-16]:

$$\bar{V}_i = \frac{1}{7} \sum_{j=i-3}^{j=i+3} V_j, \quad (1)$$

The first and second derivatives can be estimated with the use of following formulas:

$$\left. \frac{d\bar{V}}{dt} \right|_{t=t_i} \approx \frac{1}{2} (\bar{V}_{i+1} - \bar{V}_{i-1}) \quad (2)$$

$$\left. \frac{d^2\bar{V}}{dt^2} \right|_{t=t_i} \approx \bar{V}_{i+1} - 2\bar{V}_i + \bar{V}_{i-1} \quad (3)$$

### Generalized SIR model

The classical SIR model for an infectious disease [20-22] was generalized in [12, 14-16] in order to simulate different epidemic waves. We suppose that the SIR model parameters are constant for every epidemic wave, i.e. for the time periods:  $t_i^* \leq t \leq t_{i+1}^*$ ,  $i = 1, 2, 3, \dots$ . Then for every wave we can use the equations, similar to [20-22]:

$$\frac{dS}{dt} = -\alpha_i SI, \quad (4)$$

$$\frac{dI}{dt} = \alpha_i SI - \rho_i I, \quad (5)$$

$$\frac{dR}{dt} = \rho_i I. \quad (6)$$

Here  $S$  is the number of susceptible persons (who are sensitive to the pathogen and **not protected**);  $I$  is the number of infected persons (who are sick and **spread the infection**; please don't confuse with the number of still ill persons, so known active cases) and  $R$  is the number of removed persons (who **no longer spread the infection**; this number is the sum of isolated, recovered, dead, and infected people who left the region). Parameters  $\alpha_i$  and  $\rho_i$  are supposed to be constant for every epidemic wave.

Parameters  $\alpha_i$  show how quick the susceptible persons become infected (see (4)). Large values of this parameter correspond to severe epidemics with many victims. These parameters accumulate many characteristics. First they show how strong (virulent) is the pathogen and what is the way of its spreading. Parameters  $\alpha_i$  accumulate also the frequency of contacts and the way of contacting. In order to decrease the values of  $\alpha_i$ , we have to minimize the number of our contacts and change our contacting habits. For example, we have to avoid the public places and use masks there, minimize or cancel traveling. We have to change our contact habits: to avoid handshakes and kisses. First, all these simple things are very useful to protect yourself. In addition, if most people follow these recommendations, we have chance to diminish the values of parameters  $\alpha_i$  and reduce the negative effects of the pandemic.

The parameters  $\rho_i$  characterize the patient removal rates, since eq. (6) demonstrates the increase rate of  $R$ . The inverse values  $1/\rho_i$  are the estimations for time of spreading infection  $\tau_i$  during  $i$ -th epidemic wave. So, we are interested in increasing the values of parameters  $\rho_i$  and decreasing  $1/\rho_i$ . People and public authorities should work on this and organize immediate isolation of suspicious cases.

Since the derivative  $d(S + I + R)/dt$  is equal to zero (it follows from summarizing Eqs. (4)-(6)), the sum

$$N_i = S + I + R \quad (7)$$

must be constant for every wave and is not the volume of population.

To determine the initial conditions for the set of equations (4)–(6), let us suppose that at the beginning of every epidemic wave  $t_i^*$ :

$$I(t_i^*) = I_i, \quad R(t_i^*) = R_i, \quad S(t_i^*) = N_i - I_i - R_i \quad (8)$$

It follows from (4) and (5) that

$$\frac{dI}{dS} = \frac{\nu_i}{S} - 1, \quad \nu_i = \frac{\rho_i}{\alpha_i} \quad (9)$$

Integration of (9) with the initial conditions (8) yields:

$$I = \nu_i \ln S - S + N_i - R_i - \nu_i \ln(N_i - I_i - R_i) \quad (10)$$

It follows from (9) that function  $I$  has a maximum at  $S = \nu_i$  and tends to zero at infinity. The corresponding number of susceptible persons at infinity  $S_{i\infty} > 0$  can be calculated from the non-linear equation:

$$S_{i\infty} = (N_i - I_i - R_i) e^{\frac{S_{i\infty} - N_i - R_i}{\nu_i}} \quad (11)$$

Formula (11) follows from (10) at  $I=0$ .

In [12, 15, 16] the set of differential equations (4)–(7) was solved by introducing the function

$$V(t) = I(t) + R(t), \quad (12)$$

corresponding to the number of victims or the cumulative confirmed number of cases. For many epidemics (including the COVID-9 pandemic) we cannot observe dependencies  $S(t)$ ,  $I(t)$  and  $R(t)$  but observations of the accumulated number of cases  $V_j$  corresponding to the moments of time  $t_j$  provide information for direct assessments of the dependence  $V(t)$ .

It follows from (5) and (6) that:

$$\frac{dV}{dt} = \alpha_i SI \quad (13)$$

Eqs. (7), (10) and (13) yield:

$$\frac{dV}{dt} = \alpha_i (N_i - V) [v_i \ln(N_i - V) + V - R_i - v_i \ln(N_i - R_i - I_i)] \quad (14)$$

Integration of (14) provides an analytical solution for the set of equations (4)–(6):

$$F_i^*(V, N_i, I_i, R_i, v_i) = \alpha_i (t - t_i^*), \quad (15)$$

$$F_i^* = \int_{R_i + I_i}^V \frac{dU}{(N_i - U) [v_i \ln(N_i - U) + U - R_i - v_i \ln(N_i - R_i - I_i)]}. \quad (16)$$

Thus, for every set of parameters  $N_i, I_i, R_i, v_i, \alpha_i$  and a fixed value of  $V$ , integral (16) can be calculated and the corresponding moment of time can be determined from (15). Then functions  $I(t)$  and  $R(t)$  can be easily calculated with the use of formulas (10) and:

$$S = N_i - V, \quad R = V - I \quad (17)$$

The final numbers of victims (final accumulated number of cases corresponding to the  $i$ -th epidemic wave) can be calculated from:

$$V_{i\infty} = N_i - S_{i\infty}. \quad (18)$$

To estimate the final day of the  $i$ -th epidemic wave, we can use the condition:

$$I(t_{if}) = 1. \quad (19)$$

which means that at  $t > t_{if}$  less than one person still spreads the infection.

### ***Parameter identification procedure***

In the case of a new epidemic, the values of its parameters are unknown and must be identified with the use of limited data sets. For the first wave of an epidemic starting with one infected person, the

number of unknown parameters is only four, since  $I_1 = 1$  and  $R_1 = 0$ . The corresponding statistical approach was proposed in [23, 24] and used in [2-5, 8-10] to simulate the first COVID-19 pandemic wave in Ukraine.

For the next epidemic waves ( $i > 1$ ), the moments of time  $t_i^*$  corresponding to their beginning are known. Therefore the exact solution (15)-(16) depend only on five parameters -  $N_i, I_i, R_i, \nu_i, \alpha_i$ . Then the registered number of victims  $V_j$  corresponding to the moments of time  $t_j$  can be used in eq. (16) in order to calculate  $F_{i,j} = F_i^*(V_j, N_i, \nu_i, I_i, R_i)$  for every fixed values of  $N_i, \nu_i, I_i, R_i$  and then to check how the registered points fit the straight line (15).

Eq. (15) can be rewritten as follows:

$$y \equiv F_i^*(V, N_i, I_i, R_i, \nu_i) = \alpha_i t - \alpha_i t_i^* \quad (20)$$

Assuming

$$\gamma = \alpha_i, \beta = -\alpha_i t_i^* \quad (21)$$

we can estimate the values of parameters  $\gamma$  and  $\beta$ , by treating the values  $y_j \equiv F_i^*(V_j, N_i, I_i, R_i, \nu_i)$  and corresponding time moments  $t_j$  as random variables. Then we can use the observations of the accumulated number of cases and the linear regression in order to calculate the coefficients  $\hat{\gamma}$  and  $\hat{\beta}$  of the regression line

$$\hat{y} = \hat{\gamma}t + \hat{\beta} \quad (22)$$

using the standard formulas from, e.g., [27]. Values  $\hat{\gamma}$  and  $\hat{\beta}$  can be treated as statistics-based estimations of parameters  $\gamma$  and  $\beta$  from relationships (21).

The reliability of the method can be checked by calculating the correlation coefficients  $r_i$  (see e.g., [27]) for every epidemic wave checking how close its value is to unity. We can use also the F-test for the null hypothesis that says that the proposed linear relationship (20) fits the data set. The experimental values of the Fisher function can be calculated for every epidemic wave with the use of the formula:

$$F_i = \frac{r_i^2(n_i - m)}{(1 - r_i^2)(m - 1)}, \quad (23)$$

where  $n_i$  is the number of observations for the  $i$ -th epidemic wave,  $m = 2$  is the number of parameters in the regression equation. The corresponding experimental value  $F_i$  has to be compared with the critical value  $F_C(k_1, k_2)$  of the Fisher function at a desired significance or confidence level ( $k_1 = m - 1$ ,



$k_2 = n_i - m$ ). When the values  $n_i$  and  $m$  are fixed, the maximum of the Fisher function coincides with the maximum of the correlation coefficient. Therefore, to find the optimal values of parameters  $N_i, \nu_i, I_i, R_i$ , we have to find the maximum of the correlation coefficient for the linear dependence (20). To compare the reliability of different predictions (with different values of  $n_i$ ) it is useful to use the ratio  $F_i / F_C(1, n_i - 2)$  at fixed significance level. We will use the level 0.001; corresponding values of  $F_C(1, n_i - 2)$  can be taken from [28]. The most reliable prediction yields the highest  $F_i / F_C(1, n_i - 2)$  ratio.

The exact solution (15)-(16) allows avoiding numerical solutions of differential equations (4)-(6) and significantly reduce the time spent on calculations. In the case of sequential calculation of epidemic waves  $i = 1, 2, 3 \dots$ , it is possible to avoid determining the four optimal unknown parameters  $N_i, \nu_i, I_i, R_i$ , thereby reducing the amount of calculations and difficulties in isolation a maximum of the correlation coefficient. For parameters  $I_i, R_i$  it is possible to use the numbers of  $I$  and  $R$  calculated for the previous wave of epidemic at the moment of time when the following wave began. Then we need to calculate values  $F_i^*(V_i, N_i, \nu_i)$ , linear regression coefficients (22), correlation coefficient  $r_i$ ,  $F_i / F_C(1, n - 2)$  and to isolate the values of parameters  $N_i$  and  $\nu_i$  corresponding to the maximum of  $r_i$ . Knowing the optimal values of five parameters  $N_i, I_i, R_i, \nu_i, \alpha_i$ , the SIR curves and other characteristics of the corresponding epidemic wave can be calculated with the use of formulas (10)-(17). This approach has been successfully used in [12, 14-16]. In particular, six waves of the Covid-19 epidemic in Ukraine and four pandemic waves in the world were calculated.

Segmentation of epidemic waves and their sequential SIR simulations need a lot of efforts. To avoid this, a new method of obtaining the optimal values of SIR parameters was proposed in [15, 16]. First of all we can use the relationship

$$V_i = I_i + R_i \quad (24)$$

which follows from (12). To estimate the value  $V_i$ , we can use the smoothed accumulated number of cases (e.g., formula (1)). Then

$$V_i \approx \frac{1}{7} \sum_{j=i-3}^{j=i+3} V_j \quad (25)$$

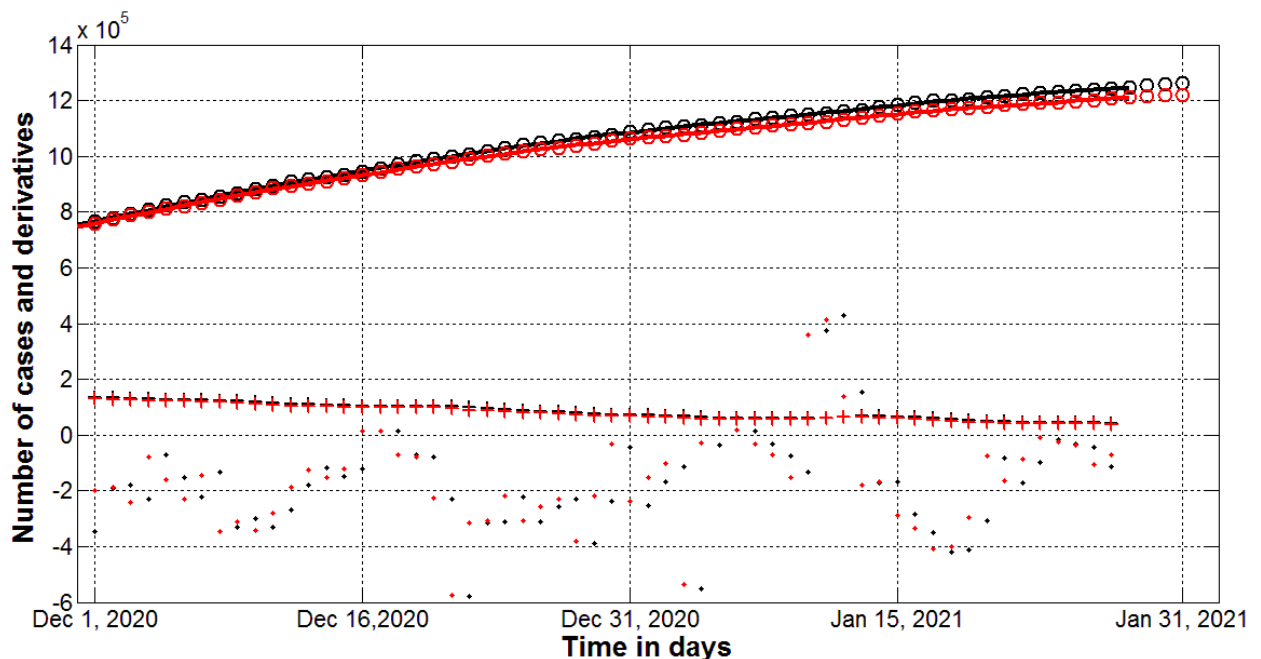
where  $i$  corresponds to the moment of time  $t_i^*$ . To obtain one more relationship, let us use (7) and (13)

$$I_i = \frac{1}{\alpha_i(N_i - V_i)} \frac{dV}{dt} \Big|_{t=t_i^*} \quad (26)$$

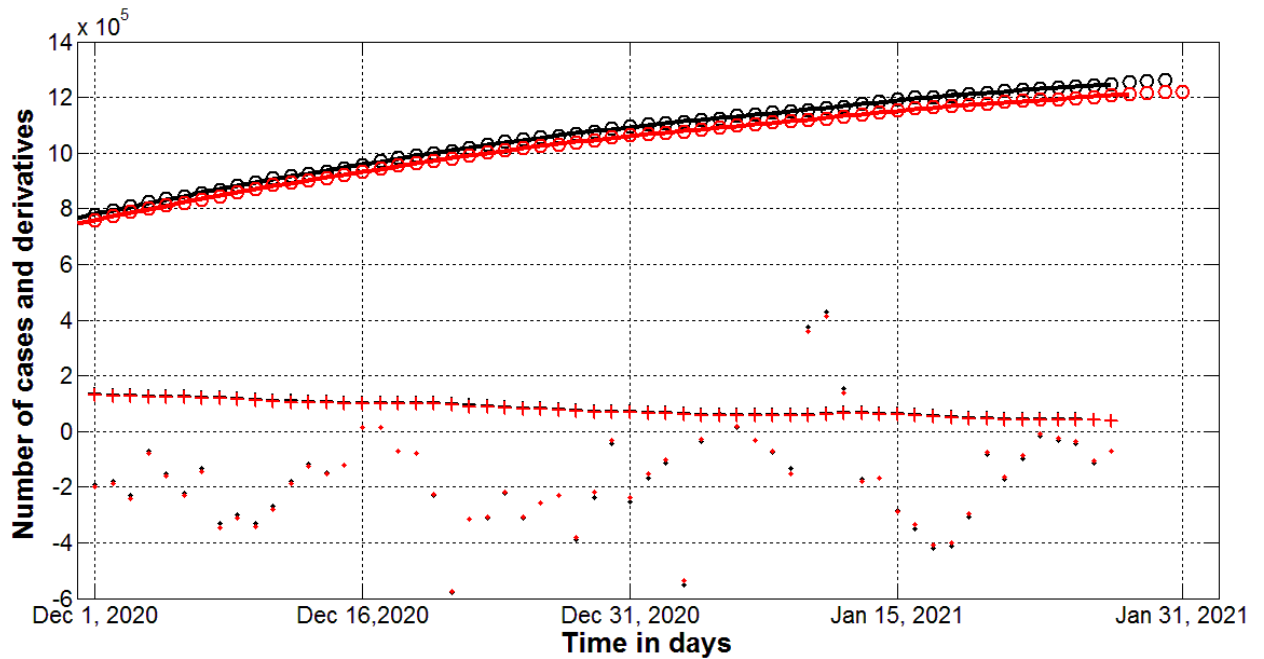
To estimate the average number of new cases  $dV/dt$  at the moment of time  $t_i^*$ , we can use (2). Thus we have only two independent parameters  $N_i$  and  $v_i$ . To calculate the value of parameter  $\alpha_i$ , some iterations can be used (see details in [15, 16]).

## Results and discussion

The COVID-19 pandemic characteristics for Ukraine in December 2020 and January 2021 are shown in Figs.1 and 2. Red color corresponds to the national statistics [17,18]; black - to JHU data [26]. “Circles” show the corresponding accumulated numbers of cases; lines represent the smoothed number of accumulated cases (eq. (1)); “crosses” – the first derivative (eq. (2)) multiplied by 10; “dots” – second derivative (eq. (3)) multiplied by 1000. Fig. 1 illustrate the data sets presented in Table 1. It can be seen one day shift for the values of the second derivatives calculated with the use of different data sets. If the numbers of cases reported by JHU are attributed to the previous day, the differences in the values of the second derivatives become almost imperceptible (see Fig. 2), but the numbers of accumulated cases according to JHU are still much higher than for the national statistics (see Fig. 2).



**Fig. 1. Pandemic dynamics in Ukraine in December 2020 and January 2021 calculated with the use of data sets from Table 1: national statistics [17, 18] (red) and data set reported by JHU [26] (black). Accumulated number of cases - “circles”; smoothed values - lines, eq. (1). “Crosses” show the first derivatives (eq. (2)) multiplied by 10, “dots” - the second derivative (eq. (3)) multiplied by 1000.**



**Fig. 2. Pandemic dynamics in Ukraine in December 2020 and January 2021 with the numbers of cases reported by JHU attributed to the previous day: national statistics [17, 18] (red) and data set reported by JHU [26] (black).** Accumulated number of cases - “circles”; smoothed values - lines, eq. (1)). “Crosses” show the first derivatives (eq. (2)) multiplied by 10, “dots” - the second derivative (eq. (3)) multiplied by 1000.

The jump of the second derivatives occurred on January 10-12, 2021 (see “dots” in Fig. 2) can be explained the New Year and Christmas celebrations. The increase in the number of contacts caused the increase in the number of new cases after some incubation period (see “crosses” in Figs. 1 and 2). The national lockdown in the period January 8-24 allowed stopping this tendency.

The results of SIR simulations with the use of two data sets presented in Table 1 are shown in Table 2 and Fig. 3. Since eight epidemic wave was already calculated in [12, 14-16], we took the period December 11-24, 2020 to calculate the ninth wave with the use of two datasets (presented in Table 1). The optimal parameters of SIR model and other epidemic characteristics can be compared with the results obtained in [15] for the eighth wave (see Table 2). The number of observations taken for calculations  $n_i$  was 14 in all cases.

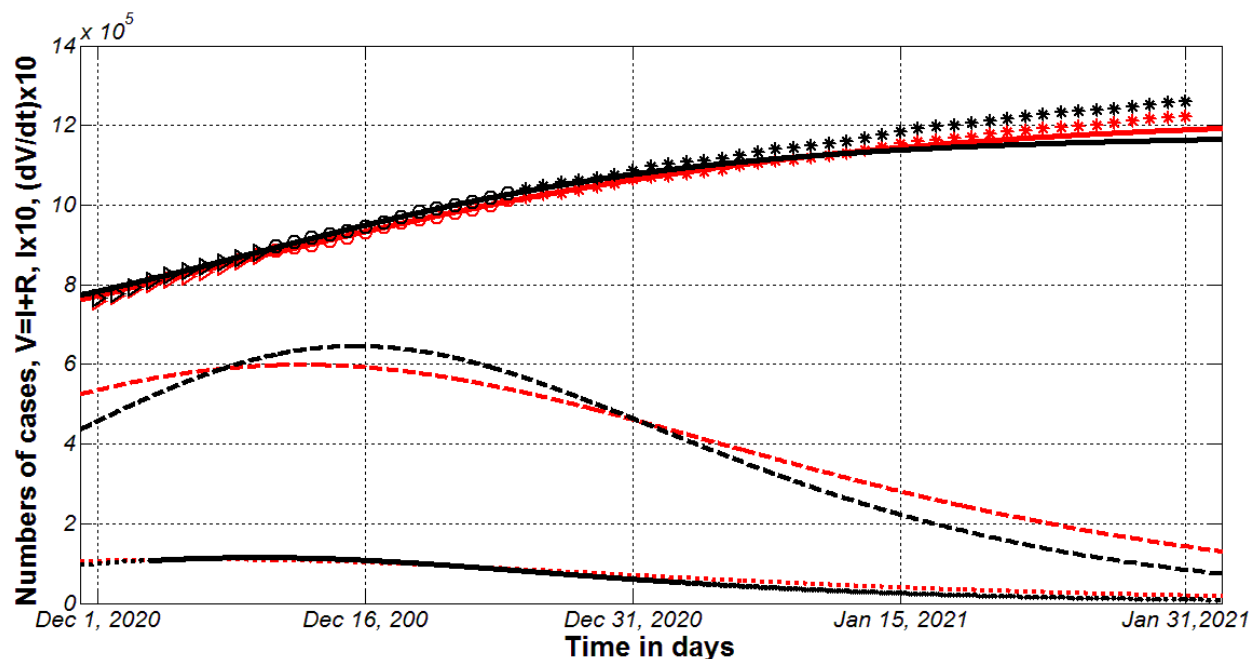
It can be seen two data sets yield rather different values of the optimal parameters for the ninth wave (especially for  $N_9$ ,  $S_{9\infty}$ , and  $\nu_9$ ), nevertheless the final sizes of this wave  $V_{9\infty}$  and  $\rho_9$  are rather close; the duration based on the national statistics is one month longer in comparison with the calculations based on JHU data set. Both simulations for the ninth epidemic wave in Ukraine yield slightly higher final sizes in comparison with the eighth wave (see Table 2).

Characteristics	Ukraine ninth wave, $i=9, n=14,$ JHU, $V_{j2}$	Ukraine ninth wave, $i=9, n=14,$ national statistics, $V_{j1}$	Eighth wave, $i=8, n=14,$ national statistics, [15]
Period taken for calculations $T_c$	December 11-24, 2020	December 11-24, 2020	Nov 21-Dec 4, 2020
$I_i$	62,384.6901000672	59,686.0031221196	74,07.6732607497
$R_i$	830,900.452757076	821,069.282592166	548,996.183882108
$N_i$	1,524,200.0384	2,037,235.2	1,454,508.4416
$V_i$	580,121.592967068	1,141,207.02807484	667,978.033362204
$\alpha_i$	2.91326081249e-07	1.55930616187218e-07	2.13865037324172e-07
$\rho_i$	0.169004550327192	0.177949115084894	0.142857147036735
$1/\rho_i$	5.91700044799982	5.6195840003078	6.9999997952
$r_i$	0.998082602967150	0.998062592012312	0.999333327224955
$F_i$ , eq. (23)	3120.24523044173	3087.92417596763	8990.91853799112
$F_i / F_C(1, n_i - 2)$	167.755119916222	166.017428815464	483.382717096297
$S_{i\infty}$ , eq. (11)	347,782	810,539	376,90
$V_{i\infty}$ , eq. (18)	1,176,418	1,226,696	1,078,218
Final day of the epidemic wave, eq. (19)	June 15, 2021	July 14, 2021	June 23, 2021

**Table 2. Calculated optimal values of SIR parameters for the eighth and ninth waves of the COVID-19 epidemic in Ukraine**

Fig. 3 illustrates the results of SIR simulations of the ninth epidemic wave in Ukraine obtained with the use of two data sets: national statistics [17, 18] (red) and data set reported by JHU [26] (black). Numbers of victims  $V(t)=I(t)+R(t)$  (eq. (12)) are represented by solid lines. Numbers of infected and spreading the infection persons  $I(t)$  are shown by dashed lines. Dotted lines represent the derivatives  $dV/dt$  calculated with the use of eq. (13). This derivative yields the estimation of the daily number of new cases. Markers show accumulated numbers of cases  $V_{j1}$  and  $V_{j2}$  from Table 1. “Circles” correspond to the accumulated numbers of cases taken for calculations (during period of time  $T_c$ ); “triangles” – numbers of cases before  $T_c$ ; “stars” – number of cases after  $T_c$ . It can be seen that the accuracy of simulations based on the national statistics is rather good (the deviations between red “stars” and red solid line are small). The use of JHU data sets yields worth accuracy. Nevertheless, the real numbers of cases already exceed the predicted saturations levels  $V_{9\infty}$  for both data sets and corresponding simulations. As of February 14, 2021 the national statistics yield the figure 1,271,143 of

accumulated cases in Ukraine [17, 18]. Thus the epidemic observations during 52 days (after  $T_c$ ) demonstrated only 3.5% exceeding of the saturation level in the case of national statistics.



**Fig. 3. The ninth COVID-19 epidemic wave in Ukraine. SIR simulations (lines) with the use with the use of data sets from Table 1: national statistics [17, 18] (red) and data set reported by JHU [26] (black). Numbers of victims  $V(t)=I(t)+R(t)$  – solid lines; numbers of infected and spreading  $I(t)$  multiplied by 10 – dashed; derivatives  $dV/dt$  (eq. (13)) multiplied by 10 – dotted. Markers show accumulated numbers of cases  $V_{j1}$  and  $V_{j2}$  from Table 1. “Circles” correspond to the accumulated numbers of cases taken for calculations (during period of time  $T_c$ ); “triangles” – numbers of cases before  $T_c$ ; “stars” – number of cases after  $T_c$ .**

## Conclusions

Constant changes in the Covid-19 pandemic conditions (i.e., in the peculiarities of quarantine and its violation, in situations with testing and isolation of patients, in coronavirus activity due to its mutations etc.) lead to new pandemic waves and cause the changes in the values of parameters of the mathematical models. Incomplete data and different methods of cases recording can give quite different values of model parameters and forecasts, which was demonstrated by the example of two data sets for Ukraine. The forecast of the duration of the epidemic in Ukraine is not optimistic. New cases will appear at least until July 2021. Let us hope that the number of new cases will grow slowly according to the results of both SIR simulations. To maintain the existing positive dynamics, it is necessary to adhere to quarantine measures and social distancing. Pandemic surveillance experience shows that neglecting simple precautions immediately leads to new waves.

The obtained results can be also useful for assessing the effectiveness of mass vaccination in Ukraine and other countries.

## Acknowledgements

I am very grateful to Professor Noureddine Benlgha and Oleksii Rodionov for their support and help in collecting and processing data.

## References

1. Nesteruk, I. "Comparison of the coronavirus pandemic dynamics in Ukraine and neighboring countries" [Preprint.] *ResearchGate*. 2020 Mar. doi: [10.13140/RG.2.2.31170.53448](https://doi.org/10.13140/RG.2.2.31170.53448)
2. Nesteruk I. Long-term predictions for COVID-19 pandemic dynamics in Ukraine, Austria and Italy [Preprint.] MEDRXIV, 2020 Apr. DOI: [10.13140/RG.2.2.31170.53448](https://doi.org/10.13140/RG.2.2.31170.53448)
3. Nesteruk I. Як довго українці сидітимуть на карантині? How long will the Ukrainians stay in quarantine? (in Ukrainian) [Preprint.] *ResearchGate*. 2020 Apr. doi: [10.13140/RG.2.2.15732.71046](https://doi.org/10.13140/RG.2.2.15732.71046)
4. Nesteruk I. SIR-simulation of Corona pandemic dynamics in Europe. [Preprint.] MEDRXIV, 2020 Apr. DOI: <https://medrxiv.org/cgi/content/short/2020.04.22.20075135v1>
5. Nesteruk I. Динаміка COVID-19 епідемії в Україні та Києві після покращання тестування. COVID-19 epidemic dynamics in Ukraine and Kyiv after testing has improved (In Ukrainian). [Preprint.] *ResearchGate*. 2020 Apr. doi: [10.13140/RG.2.2.36705.86885](https://doi.org/10.13140/RG.2.2.36705.86885)
6. Nesteruk, I., Kydybyn, I., Demelmair, G. Corona pandemic global stabilization? [Preprint.] *ResearchGate*. 2020 April. doi: [10.13140/RG.2.2.21570.32961](https://doi.org/10.13140/RG.2.2.21570.32961)  
[https://www.researchgate.net/publication/340917471\\_Corona\\_pandemic\\_global\\_stabilization](https://www.researchgate.net/publication/340917471_Corona_pandemic_global_stabilization)
7. Nesteruk, I., Kydybyn, I., Demelmair, G. GLOBAL STABILIZATION TRENDS OF COVID-19 PANDEMIC. KPI Science News, 2020, No. 2, pp. 55-62. DOI: [10.20535/kpi-sn.2020.2.205124](https://doi.org/10.20535/kpi-sn.2020.2.205124)
8. Nesteruk I. Simulations and predictions of COVID-19 pandemic with the use of SIR model. *Innov Biosyst Bioeng*, 2020, vol. 4, no. 2, 110–121, doi: [10.20535/ibb.2020.4.2.204274](https://doi.org/10.20535/ibb.2020.4.2.204274).  
<http://ibb.kpi.ua/article/view/204274>
9. Nesteruk I. Hidden periods, duration and final size of COVID-19 pandemic. Preprint. *MedRxiv*. May 10, 2020. <https://medrxiv.org/cgi/content/short/2020.05.10.20097147v1>
10. Nesteruk I. Статистика пандемії COVID-19 в Україні та світі COVID-19 pandemic statistics in Ukraine and world. (In Ukrainian). *ResearchGate*. 2020 May doi:[10.13140/RG.2.2.30999.39842](https://doi.org/10.13140/RG.2.2.30999.39842)  
[https://www.researchgate.net/publication/341542317\\_Statistika\\_pandemii\\_COVID-19\\_v\\_Ukraini\\_ta\\_sviti\\_COVID-19\\_pandemic\\_statistics\\_in\\_Ukraine\\_and\\_world](https://www.researchgate.net/publication/341542317_Statistika_pandemii_COVID-19_v_Ukraini_ta_sviti_COVID-19_pandemic_statistics_in_Ukraine_and_world)
11. Nesteruk I. Coronasummer in Ukraine and Austria. [Preprint.] *ResearchGate*. 2020 June. DOI: [10.13140/RG.2.2.32738.56002](https://doi.org/10.13140/RG.2.2.32738.56002)
12. Nesteruk I. Waves of COVID-19 pandemic. Detection and SIR simulations. *MedRxiv*. 2020 Aug. doi: <https://doi.org/10.1101/2020.08.03.20167098>  
<https://www.medrxiv.org/content/10.1101/2020.08.03.20167098v1>
13. Nesteruk I, Nikitin A, Shepetyuk B. Study of the COVID-19 epidemic dynamics in Ukraine and its regions by methods of probabilistic and statistical analysis of evolutionary models. [Preprint.] *ResearchGate*. 2020 Aug. DOI: [10.13140/RG.2.2.30998.98884](https://doi.org/10.13140/RG.2.2.30998.98884)

[https://www.researchgate.net/publication/343651320\\_Study\\_of\\_the\\_COVID-19\\_epidemic\\_dynamics\\_in\\_Ukraine\\_and\\_its\\_regions\\_by\\_methods\\_of\\_probabilistic\\_and\\_statistical\\_analysis\\_of\\_evolutionary\\_models](https://www.researchgate.net/publication/343651320_Study_of_the_COVID-19_epidemic_dynamics_in_Ukraine_and_its_regions_by_methods_of_probabilistic_and_statistical_analysis_of_evolutionary_models)

14. Nesteruk I. Нові хвилі пандемії COVID-19 в Україні. New waves of COVID-19 pandemic in Ukraine. (In Ukrainian). [Preprint.] ResearchGate. 2020 Aug. DOI:[10.13140/RG.2.2.26172.26244](https://doi.org/10.13140/RG.2.2.26172.26244)  
[https://www.researchgate.net/publication/343979510\\_Novi\\_hvili\\_pandemii\\_COVID-19\\_v\\_Ukraini\\_New\\_waves\\_of\\_COVID-19\\_pandemic\\_in\\_Ukraine](https://www.researchgate.net/publication/343979510_Novi_hvili_pandemii_COVID-19_v_Ukraini_New_waves_of_COVID-19_pandemic_in_Ukraine)
15. Nesteruk I. COVID-19 pandemic dynamics in Ukraine after September 1, 2020. MedRxiv. Posted December 22, 2020. Doi: <https://doi.org/10.1101/2020.12.21.20248627>
16. I. Nesteruk. *COVID19 pandemic dynamics*. Springer Nature, DOI: 10.1007/978-981-33-6416-5
17. <https://index.minfin.com.ua/ua/reference/coronavirus/ukraine/>
18. <https://www.kmu.gov.ua/>
19. World Health Organization. “Coronavirus disease (COVID-2019) situation reports”. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>.
20. Kermack WO, McKendrick AG. A Contribution to the mathematical theory of epidemics. J Royal Stat Soc Ser A. 1927;115:700-21.
21. Murray JD. *Mathematical Biology I/II*. New York: Springer; 2002.
22. Langemann D, Nesteruk I, Prestin J. Comparison of mathematical models for the dynamics of the Chernivtsi children disease. *Mathematics in Computers and Simulation*. 2016;123:68-79. DOI: 10.1016/j.matcom.2016.01.003
23. Nesteruk I. Statistics based models for the dynamics of Chernivtsi children disease. *Naukovi Visti NTUU KPI*. 2017;5:26-34. DOI: 10.20535/1810-0546.2017.5.108577
24. Nesteruk I. Statistics-based predictions of coronavirus epidemic spreading in mainland China. *Innov Biosyst Bioeng*. 2020; 4(1):13–18. DOI: 10.20535/ibb.2020.4.1.195074.
25. <https://podillyanews.com/2020/12/17/u-shkolah-hmelnytskogo-provely-eksperyment-z-testuvannyam-na-covid-19/>
26. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). <https://github.com/owid/covid-19-data/tree/master/public/data>
27. Draper NR, Smith H. *Applied regression analysis*. 3rd ed. John Wiley; 1998.
28. <https://onlinepubs.trb.org/onlinepubs/nchrp/cd-22/manual/v2appendixc.pdf>