ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume **26**, Number 1, 2023, 48–63

New Opportunities Model for Monitoring, Analyzing and Forecasting the Official Statistics on Coronavirus Disease Pandemic

Sergey M. ABRAMOV[1], Sergey TRAVIN[2], Gheorghe DUCA[3], and Radu-Emil PRECUP[4, 5, *]

[1]Russian Academy of Sciences, Program Systems Institute, 152140 Pereslavl-Zalessky, Russian Federation

[2]Russian Academy of Sciences, Semenov Federal Research Center for Chemical Physics, Kosygina Street 4, Building 1, 119991 Moscow, Russian Federation

[3]Institute of Chemistry, Research Center of Physical and Inorganic Chemistry, Str. Academiei 3, 2028 Chisinau, Republic of Moldova

[4]Politehnica University of Timisoara, Department of Automation and Applied Informatics, Bd. V. Parvan 2, 300223 Timisoara, Romania

[5]Romanian Academy – Timisoara Branch, Center for Fundamental and Advanced Technical Research, Bd. Mihai Viteazu 24, 300223 Timisoara, Romania

E-mails: abram@botik.ru, travinso@yandex.ru, ggduca@gmail.com, radu.precup@aut.upt.ro[*]

[*] Corresponding author

Abstract. At the beginning of 2020, it became obvious that the coronavirus disease 2019 (COVID-19) pandemic will have a fairly significant scale and duration. There was an unmet need for the analysis and forecast of the development of events. The forecast was needed to make the managerial decisions in terms of knowledge on the dynamics of the pandemic, considering and analyzing the incoming official statistics about the pandemic, modeling and predicting the behavior of this statistics. Due to the objective and subjective factors, the available statistics is far from the unknown true data regarding the pandemic. Therefore, strictly speaking, it was necessary to model and predict not the dynamics of the pandemic, but the dynamics of the official (*i.e.* government) statistics on the pandemic. This paper proposes a new model, referred to as the new opportunities model, to monitor, analyze and forecast the government statistics on COVID-19 pandemic. A modeling approach is offered in this regard. The modeling approach is important as it answers simple questions on what awaits us in the near future, which is the current phase of the pandemic and when all this will be over. The new opportunities model is applied to three different countries in terms of area, economy and population, namely Russia, Romania and Moldova, plus the Campania region in Italy, and proves to be efficient over other similar models including the classical *Susceptible-Infected* (SI) model.

Key-words: Coronavirus disease 2019 pandemic; new opportunities model; official statistics; optimization; parabolic regression; prediction.

1. Introduction

There are two general approaches to dynamic systems modeling and system identification. The first one is based on measuring the values affecting the system, next calculating the system trajectory, comparing it with experimental data, computing the system structure and models such that to fit the input-output data (or the experimental data). This approach is referred to as data-driven modeling, and it is currently popular in the context of machine learning as it usually employs nonlinear models as neural networks, fuzzy models and their combinations. Some recent models and results of this approach are briefly discussed as follows. Finite difference approaches are applied in [1] to financial pricing models. Bifurcation sequences are discussed in [2] in relation with piecewise linear maps. Data streams are subjected to clustering and investigated on the basis of a clustering approach in [3] and the scarcity of labeled samples in large-scale data streams is ensured in [4] using weakly supervised scalable teacher forcing networks. Decision-making models are deeply discussed in [5] in the context of human well-being and resilience. Fuzzy logic is widely used in this approach because of their transparency and ability to capture both the dynamics and the nonlinearities. Fuzzy rule interpolation-based models applied to student result prediction in [6]. Fuzzy cognitive maps are applied to traffic management in [7] in the framework of agent-based cloud computing systems. Evolving fuzzy models are involved in cloud-based identification [8], modeling the dynamics of shape memory alloy wire actuators [9] and adaptive cloud-based control [10]. The tensor product-based model transformation is generally presented in [11] focusing on model-based control, investigated in [12] in close relation to fuzzy modeling and applied in [13] to tower crane systems modeling. Signatures suggested in [14] are applied to the observation process modeling in [15] and the algebraic structure of fuzzy signatures is analyzed in [16]. Neural networks are proposed in [17] for modeling based on biomonitoring studies data, in [18] for relation patterns extraction from climate data and in [19] for drug development and biomedical applications.

The second approach to dynamic systems modeling and system identification is based on knowing the analytical solution to the system of differential equations, which is actually the system model. The parameters of the model are next computed so that the analytical solution fits as closely as possible the input-output data.

Both approaches to dynamic systems modeling need the correct definition of an optimization problem, where the objective function is usually quadratic and depends on the modeling errors, *i.e.* the differences between the model output (or the solution) and the experimental data. The variables of the objective function in the framework of the optimization problem are the parameters of the models. Several classical optimization algorithms are applied in this regard as the ordinary least squares approach. However, metaheuristic algorithms are also popular recently as, for example, cellular genetic algorithms [20], NSGA-III [21], [22], hybrid quantum computing-tabu search algorithms [23], monarch butterfly optimization algorithms [24], slime mould algorithms [25], [26], particle swarm optimization algorithms [27], hybrid particle filter-particle swarm optimization algorithms [28], string theory algorithms [29], and harmony search optimization algorithms [30].

This paper uses the second approach described above and suggests a model, referred to as the new opportunities model, and a modeling approach to forecast the official (*i.e.* government) statistics on the coronavirus disease 2019 (COVID-19) pandemic. Different models of pandemics and epidemics are used in the literature, and they are referred to as SI, SIR and SEIR, where S indicates susceptible, E indicates exposed, I indicates infected, and R indicates recovered. If the primary data is not complete enough, and sometimes it could be contradictory and unreliable, there is no reason to use too complex models. That is the reason why this paper is focused on the simplest model, which is the SI (*i.e.* susceptible plus infected) model, because no reliable primary data is available.

The motivations for the new opportunities model proposed in this paper are:

- unsuccessful management decisions leading to the expansion of the field of spread of the pandemic,
- successful management decisions leading to a narrowing of the field of spread of the pandemic,
- change of the method of registration of cases,
- intensive mass vaccination,
- a new strain is also a change in the spread of the pandemic, a change in the number of those who fall ill.

The new opportunities model is important and advantageous with respect to the state-of-the-art briefly discussed here because of the following reasons:

- it is simple and also transparent, which makes it applicable to model other systems in different areas,
- it exhibits very good performance, *i.e.* high prediction accuracy,
- it outperforms other similar models.

This paper is structured as follows: the analysis of the SI model, the development of the new opportunities model and the modeling approach are presented in the next section. The validation of the model and the modeling approach is carried out in Section 3 in terms of the application to three different countries as far as their areas, economies and populations are concerned, namely Russia, Romania and Moldova, and the Campania region in Italy. The efficiency of the model is proved in comparison with other similar models. The conclusions are drawn in Section 4.

2. The New Opportunities Model and the Modeling Approach

The development of the model is discussed making use of an example, namely flying a ball over the surface of a planet. It is considered that a ball is thrown at an angle to the horizon as shown in Fig. 1 in order to illustrate the modeling approach by means of a relatively simple example.

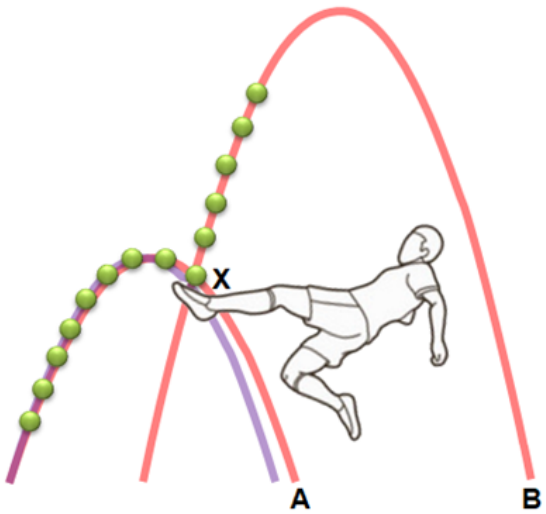

Fig. 1. Illustration of the trajectory of a ball thrown at an angle to the horizon.

The gravitational acceleration on the planet and the parameters of the throw might be unknown. However, that does not matter because it is known that neglecting the air resistance and more complex aerodynamic effects (only inertia and gravity), the ball will move along a polynomial including a parabola.

It is supposed that some experimental data is available, namely some timestamps and coordinates of the ball. Then the parabola is drawn as close as possible to these coordinates, and the well-known polynomial regression algorithm can be used in this regard. This will next allow the analysis of the movement process, *i.e.* angles, speeds and gravitational acceleration can be calculated. Finally, predictions can be made, that means the start, top (vertex) and end points and the ball coordinates at any time can be computed.

The system (or the process) subjected to modeling can also be monitored. When new data is applied, it is checked that the new points do not deviate much from the parabola. The parabola is also recalculated in order to describe more precisely the true trajectory of the ball.

If the new points deviate greatly from the parabola, this indicates that obviously something happened. Therefore, it is needed to accumulate some new data and calculate a new parabola. In other words, the ball lost the opportunity to fly along the parabola A, and the ball obtained the opportunity in the point X to complete its flight along the parabola B.

The illustration given in Fig. 1 and the above physical interpretation introduce the following elements used in the sequel:

- modeling using experimental data and parabolic regression,
- analysis, *i.e.* calculation of model parameters,
- forecast by calculating the start and end points, vertex, coordinates at any time,
- monitoring that new points do not deviate from the parabola, and the refinement of the parabola,
- transition recognition from the previous opportunity to the current one.

The subject of the approach proposed in this paper is to model, monitor, forecast using official statistics and not the true pandemic process. The authors understand that this data is far from true data about the pandemic, but no other data is available. Only official daily statistics are available, namely the numbers of COVID-19 cases, recoveries and deaths. The available statistics are far from unknown true data about the pandemic due to objective and subjective factors as, for example, in Russia the difference is six to ten times.

The official sites are the sources of primary data for the model development. For example, these sites are [31] for Russia, [32] for Romania, [33] for Moldova and [34] for the Campania region in Italy. Every day, the number of cases, the number of recovered and the number of deaths are extracted from the site. These values and the timestamp are noted by the letters V, H, D and t, respectively. In this regard, the dataset is S

$$S = \{(t_i, V_i, H_i, D_i)|i = 0...m\}, \tag{1}$$

where i is the index of the current data sample, t_i (days) is the timestamp, m is the number of data samples or quadruples (t_i, V_i, H_i, D_i) available at the time timestamp t_m, V_i is the number of COVID-19 cases, H_i is the number of recovered cases (or recoveries), and D_i is the number of deaths. The measuring unit for V_i, H_i and D_i is number of people per day.

Secondary data is calculated on the basis of primary data. Three categories of secondary data are computed. First, the number of people who have finished getting sick, with the notation E_i, is the sum of recovered cases and deaths, *i.e.*

$$E_i = H_i + D_i,\ i = 1 \ldots m. \tag{2}$$

Second, the number of people who are still sick, with the notation S_i, is calculated in terms of

$$S_i = V_i - E_i,\ i = 1 \ldots m, \tag{3}$$

and it represents a burden on the healthcare system.

Third, seven days average velocities are calculated for all V_i, E_i, and S_i. The velocity of a certain variable is denoted by the same letter as the variable, prefixed with a small letter v, leading to the notations vV_i, for the velocity of number of COVID-19 cases vE_i, for the number of people who have finished getting sick and S_i for the number of people who are still sick. These three average velocities are obtained as follows as finite difference approximations:

$$\begin{aligned} &v\Gamma = (\Gamma - \Gamma_{i-7})/(t_i - t_{i-7}), i = 7 \ldots m, \\ &v\Gamma = 0, i = 1 \ldots 6, \\ &\Gamma \in \{V, E, S\}, \end{aligned} \tag{4}$$

and they are important as they offer useful information on the trend.

As specified in the previous section, this paper focused on the simplest model of pandemic or epidemic, namely the SI model. This model is characterized by the system of ordinary differential equations [35]

$$\begin{aligned} &\dot{S}(t) = -\beta SI(t)/N, \\ &\dot{I}(t) = \beta SI(t)/N - \gamma I(t), \\ &\dot{R}(t) = \gamma I(t), \end{aligned} \tag{5}$$

or the ordinary differential equation

$$\dot{V}(t) = aV(t)(N - V(t)), t \geq \tau, V(t) = 0.5N, \tag{6}$$

with the parameters: $a > 0$, $N > 0$ and $\tau > 0$, where $0.5N$ is the initial condition, and t is the independent time variable. Inserting the notation

$$k = 0.5aN, \tag{7}$$

the analytic solution to the ordinary differential equation given in (6), which is also the response of the dynamic system (or process) model given in (6), is

$$V(t) = 0.5N[1 + \tanh(k(t - \tau))] = Ne^{2k(t-\tau)}/[e^{2k(t-\tau)} + 1], t \geq \tau. \tag{8}$$

The model in (6) describes not only an epidemic, but also any other self-accelerating process in fields with limited resources. The term "self-accelerating process" means that each single event generates several other similar events, and the event replication rate may vary. The term "fields with limited resources" outlines that there are resource limitations which do not allow the infinite expansion of the process. Therefore, the process will stop, reaching a certain unknown size.

An example of epidemic modeled using the SI model in (6) is presented as follows. The values of the parameters are set as

$$a = 2 \cdot 10^{-7},\ N = 500\,000,\ \tau = 131 \text{ days},\ k = 0.5a\,N = 0.05. \tag{9}$$

The volume graphs of an epidemic according to the SI model in (6) are presented in Fig. 2, and they consist of the following system responses versus time: *V* – the number of cases, in the s-shaped blue curve, *E* – the number of people who finished getting sick, recovered or died, in the green curve, and *S* – the difference between *V* and *E*, indicating how many people are still sick, in the red curve. The following key events are marked on the curves: the maximum burden on the healthcare system, the last case, and the end of the pandemic or epidemic. Each key event defines a timestamp and a value for the corresponding variable *V, E* and *S*.

The velocity graphs of an epidemic according to the SI model in (6) are presented in Fig. 3, and they consist of the evolutions of the seven days average velocities versus time for *V, E* and *S* with the graphs presented in Fig. 2. Fig. 3 shows the corresponding key events: the maximum of each velocity, the point where *vV* becomes low enough, and the intersection point of *vV* and *vE*.

The example is continued with presenting the phase plane plot of the solution to the SI model in (6), namely the presentation in the $< V(t), \dot{V}(t) = \partial V/\partial t >$ plane, where *t* plays the role of parameter. This response is illustrated in Fig. 4, where the number of cases *V* is represented on the X-axis and the velocity *vV* of the number of cases is represented on the Y-axis.

The differential equation in (6) is a quadratic polynomial in *V*, *i.e.* $aV(N - V)$, in its right-hand term. This means that the phase plane plot is a parabola, which is shown in Fig. 4. The left root of the parabola (0, 0) describes the beginning of the epidemic, namely the situation when the number of cases was 0 and the velocity of the number of cases was also 0. The right root (*N*, 0) describes the end of the epidemic.

When the number of cases reaches a certain limit *N*, the epidemic will end, and the velocity of the number of cases will be 0. This value of *N* is unknown, but it exists and it can be calculated as the right root of the parabola.

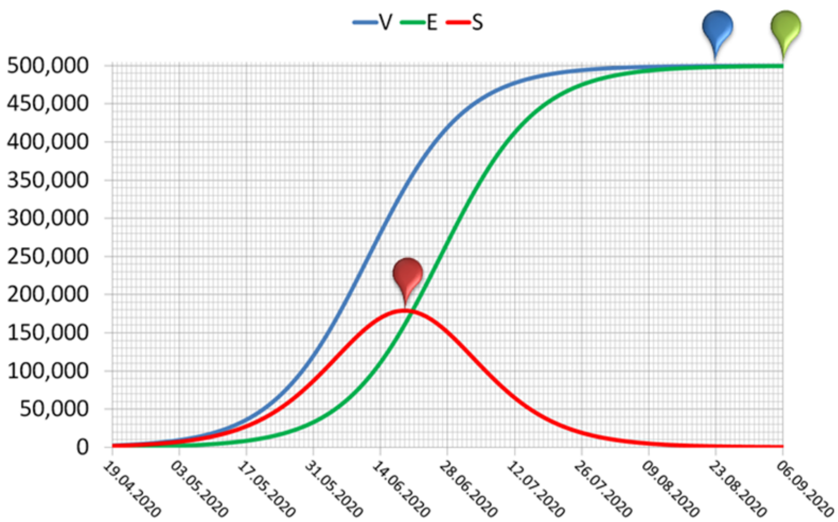

Fig. 2. Volume characteristics in the example of epidemic evolution characterized by the SI model in (6) and parameters in (9). Variables on axes: X-axis: time (days), Y-axis: V – number of new COVID-19 cases per day, E – number of people finished getting sick per day, $S = V - E$ – number of people still sick per day.

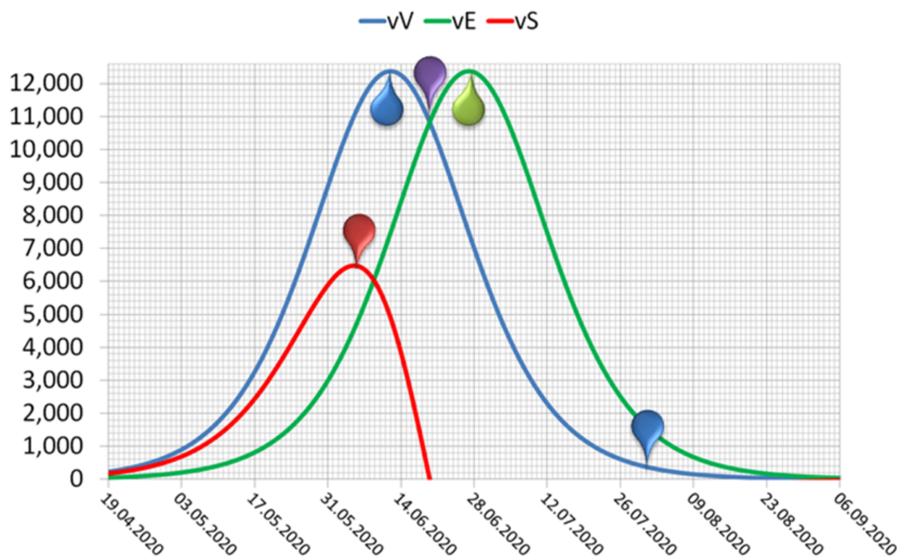

Fig. 3. Seven days average velocity graphs in the example of epidemic evolution characterized by the SI model in (6) and parameters in (9). Variables on axes: X-axis: time (days), Y-axis: vV – velocity of number of new COVID-19 cases per day, vE – velocity of number of people finished getting sick per day, vS – velocity of number of people still sick per day.

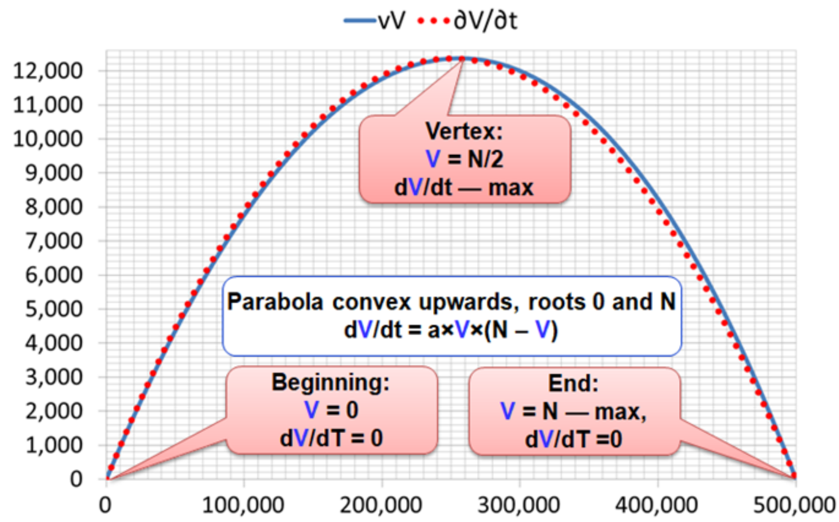

Fig. 4. Phase plane plot in the example of epidemic evolution characterized by the SI model in (6) and parameters in (9). Variables on axes: X-axis: V – number of COVID-19 cases, Y-axis: *vV* – velocity of number of COVID-19 cases.

The above example clearly shows that the modeling using the SI model works in a similar way to analyzing a ball flight. The modeling approach that produces the SI model is organized in terms of the steps SI1, SI2 and SI3.

Step SI1. The points of the real COVID-19 statistics are plotted in the phase plane.

Step SI2. The values of the parameters *N* and *a* are chosen so that the parabola in the right-hand term of (7) fits closely the COVID-19 statistics. The relationship (7) is next applied to obtain the value of the parameter *k*.

Step SI3. The value of the parameter τ is found from the condition that the function *V(t)* in (8) passes through the rightmost point of the COVID-19 statistics in the plane $< t, V >$, namely the point (t_m, V_m). Imposing $t = t_m$ in (8) and solving it with respect to the unknown τ, the solution is

$$\tau = t_m + 0.5/k \ln(N/V_m - 1). \tag{10}$$

This modeling approach leads to the expression of the function *V(t)* given in (8). The values of this function can be calculated in the future, thus obtaining the forecast based on the SI model.

A parabola $aV(N - V)$ that closely approximates the statistics points is computed at step SI2. This is a parabolic regression, and the left root of the parabola is set to zero.

This modeling approach is employed in modeling, analysis and forecasting. The monitoring is built upon the new statistics that come every day and are applied to the SI model in (5). The three steps of the modeling approach are applied repeatedly every day, and new updated values of the of the parabola parameters *N* and *a*, are computed, along with the calculation of the updated τ, the new forecast and the new values of key events.

The SI modeling approach was applied on 21.05.2020 to Russia using the data reported in [31]. The results of this SI model are presented in terms of the phase plane plot illustrated in Fig. 5. The plot shows that the parabola does not approximate the statistics with a reasonable accuracy.

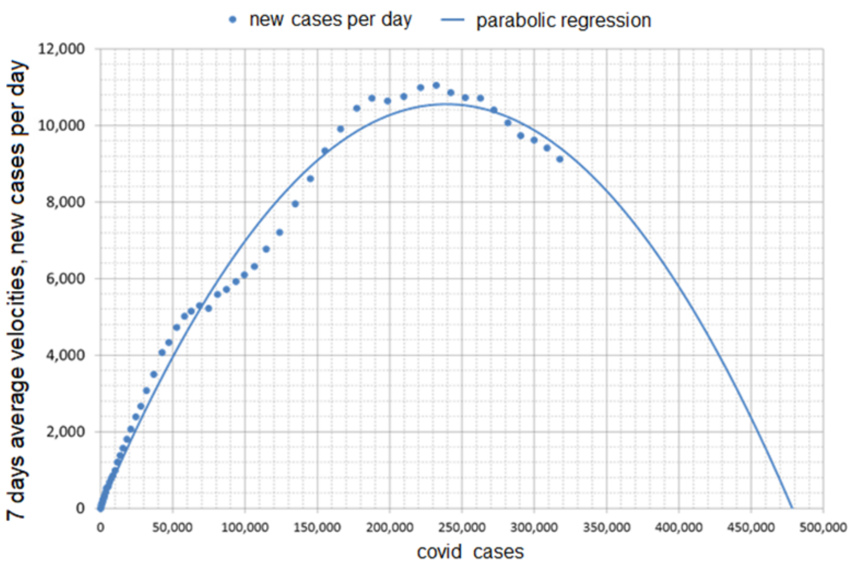

Fig. 5. Phase plane plot in the example of epidemic evolution in Russia at 21.05.2020 according to the SI model in (6) (parabolic regression). Variables on axes: X-axis: V – number of COVID-19 cases, Y-axis: vV – velocity of number of COVID-19 cases.

The results of this SI model are next presented after 50 days, at 13.07.2020, and are expressed as the phase plot presented in Fig. 6 and the system responses in Fig. 7. The plots in Fig. 6 and Fig. 7 clearly show that the results of the SI model are no longer good at all.

Obviously, the equation that holds true for all self-accelerating processes in fields with limited resources is violated. The epidemic leaves the parabola at point X illustrated in Fig. 1.

The epidemic was, is and will be a self accelerating process. Therefore, only the second condition can be violated, namely the amount of resources available for the epidemic has changed at point X, which restarts with a new resource limit. This restart occurs from a non-zero level, which means that it is necessary to build a new parabola with free, both left and right, roots. The epidemic has lost the opportunity to travel along the parabola A, ending with approximately 200 000 cases.

That is the reason why the development of the new opportunities model is justified and motivated. It is considered that, as suggestively illustrated in Fig. 1 and Fig. 8, adding another parabolic regression at a certain moment will improve the model performance.

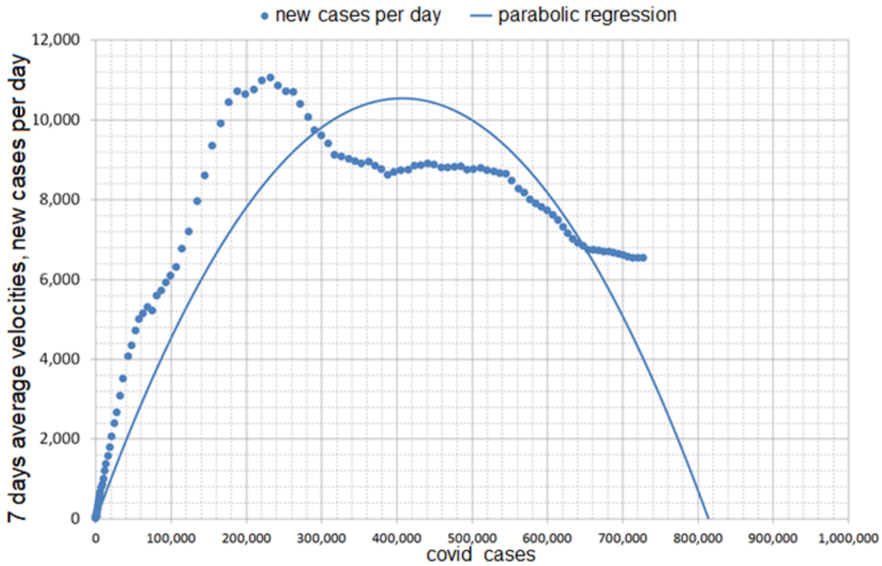

Fig. 6. Phase plane plot in the example of epidemic evolution in Russia at 13.07.2020 according to the SI model in (6) (parabolic regression). Variables on axes: X-axis: V – number of COVID-19 cases, Y-axis: vV – velocity of number of COVID-19 cases.

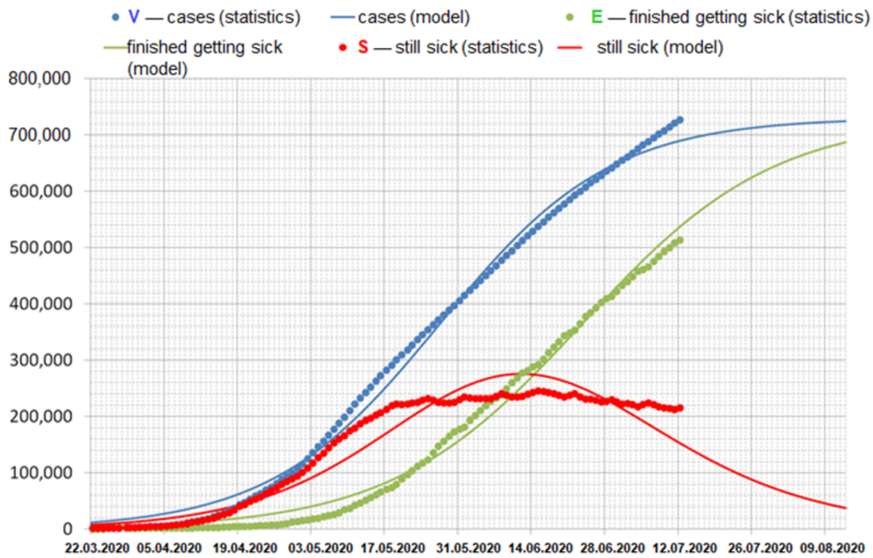

Fig. 7. System responses (epidemic volume characteristics) in the example of epidemic evolution in Russia at 13.07.2020 according to the SI model in (6) (parabolic regression). Variables on axes: X-axis: time (days), Y-axis: numbers of COVID-19 cases.

Let t_m be today, the official statistics saved in the dataset S defined in (1) and the seven days average velocities computed in accordance with (4). The value of the error threshold $\varepsilon > 1$ is set. The new opportunities modeling approach consists of the steps NO1 to NO4 described as follows.

Step LO1. *The current opportunity.* The dataset

$$\{(t_i, V_i, vV_i)|i = m - L + 1 \dots m\} \tag{11}$$

is collected for which there is an approximation

$$vV_i \approx P(V_i) = \dot{V}(t_i) = aV_i(N - V_i), i = m - L + 1 \dots m, \tag{12}$$

namely the following inequality is fulfilled:

$$|vV_i - aV_i(N - V_i)| < \varepsilon, i = m - L + 1 \dots m, \tag{13}$$

on an as long as possible time length L, indicating that the right-hand term is suitable for the pandemic model (6). The dataset in (11) and the parabola $P(V_i)$ are called current opportunity.

If the current opportunity is not found, then the nearest current opportunity in the past is used.

Step NO2. *The calculation of* τ. The relationship (7) is applied to obtain the value of the parameter k. The relationship (9) is next applied to obtain the value of the parameter τ.

Step NO3. *Forecasting the future.* The following computations are carried out for :

- the time stamp:

$$t_i = t_m + i - m; \tag{14}$$

- the estimated number of COVID-19 cases:

$$V_i = 0.5N[1 + \tanh(k(t_i - \tau))] = Ne^{2k(t_i-\tau)}/[e^{2k(t_i-\tau)} + 1], \tag{15}$$

  which is next rounded to the nearest integer;

- the end of the pandemic, when V_i does not change for seven days, and seven extra days are added;

- the number of people who have finished getting sick, $E_i = V_{i-c(i)}$, where $c(i)$ is the prognosis of the mean time of the disease progression;

- the number S_i of people who are still sick in terms of (3);

- velocities vV_i, vE_i and vS_i using (4) and key events.

Step NO4. *The forecast drift.* The parameters of key events calculated for t_m are saved.

The operations carried out in the framework of the new opportunities modeling approach are summarized and illustrated in Fig. 8. The daily modeling process is illustrated.

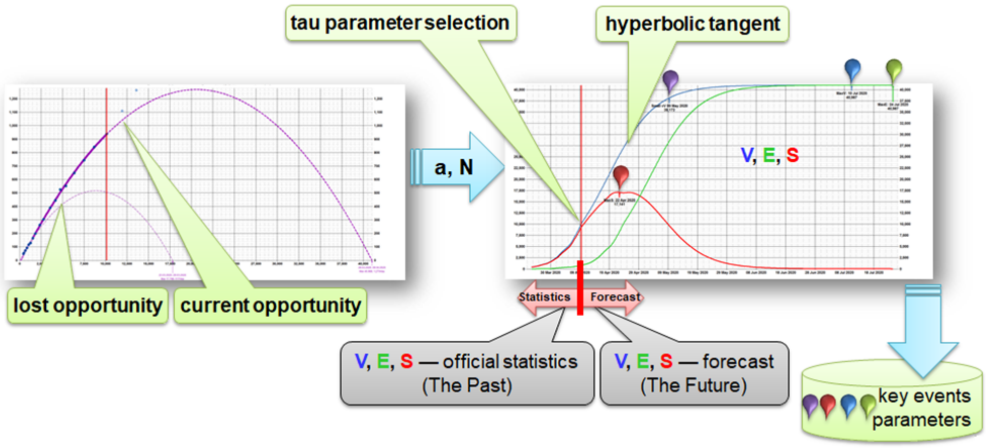

Fig. 8. Illustration of operations specific to the daily modeling process carried out in the framework of the new opportunity modeling approach.

The forecast drift carried out at step NO4 is a prediction of forecast changes. The forecast drift monitoring is an essential part of the new opportunities modeling approach as time series data is accumulated and charts are drawn in order to describe how the predictions change daily. The predictions of key events give information on expected values and timestamps of forecasted key events. This allows the prediction of changes in the forecasts regarding the volume, intensity and duration of the pandemic.

The model is built every day with the parameters computed for t_m, and the forecasts are calculated accordingly. All obtained values, namely $V_i, E_i, S_i, vV_i, vE_i$ and vS_i, can be used for operational forecast, to assess the near future for $i > m$.

The parameters of key events give a long-term estimate of the scale Γ, the velocities $v\Gamma$, with $\Gamma \in \{V, E, S\}$ in the context of (4), and the dates t of the key events, an understanding of what phase of the epidemic is in, which the maximum volumes and intensities will be. In other words, the forecast drift gives the information how the forecasted parameters of key events, namely t, Γ and $v\Gamma$, change with respect to the current date t_m.

3. Experimental Results

The new opportunities approach was applied to the official statistics of Russia, Romania, Moldova, and the Campania region in Italy. Several new opportunities models are built in this regard.

Table 1 given in [36] summarizes the main current forecasts of the pandemic in the first three countries at 28.06.2021. The "Percentage" row indicates that Romania is completing the pandemic in terms of forecasted cases, and the dates in the bottom two lines confirm this. Moldova is also completing the pandemic in terms of forecasted cases, but the completion dates are longer; that is due to the very low intensity at the end of the pandemic in Moldova. Russia is very far from the end of the pandemic.

Figure 9 given in [36] describes the current and new opportunities for Russia at 28.06.2021. It shows that Russia met a new opportunity three weeks before 28.06.2021 and it was in the third wave at that date.

The current calculation of epidemic volumes for Russia, the key events and the growing burden on the healthcare system are illustrated in Fig. 10 given in [36].

Figure 11 given in [36] displays the drift of the forecast of volumes for Russia. Figure 12 given in [36] shows that the forecast for the number of cases did not exhibit oscillations during the previous Winter. Rather big oscillations occurred at that time.

The drift of the forecast of dates for Russia at 28.06.2021 is presented in Fig. 12 given in [36]. Figure 12 shows that the forecast for the dates of many key events has been worsening last week prior to 28.06.2021.

Figure 13 given in [36] offers a synthetic view on the current and new opportunities for Romania at 28.06.2021. Figure 13 shows that Romania was firmly moving along a steep parabola towards the end of the pandemic.

The current calculation of epidemic volumes for Romania, the key events and the growing burden on the healthcare system are illustrated in Fig. 14 given in [36]. Figure 14 outlines that the key event "low enough velocity" has already happened, and the two remaining key events were expected to come soon.

Figure 15 given in [36] displays the drift of the forecast of volumes for Moldova. Figure 15 highlights that there have been no strong oscillations since the Spring of 2021.

The drift of the forecast of dates for Romania at 28.06.2021 is presented in Fig. 16 given in [36]. Figure 16 points out again that there have been no strong oscillations since Spring of 2021.

Figure 17 given in [36] shows the current and new opportunities for Moldova at 28.06.2021. Figure 17 indicates that Moldova has jumped at the very end of the steep parabola from it to a small and low parabola. This leads to an increase in the duration of the epidemic, but with a low intensity. The calculation of the epidemic volumes and their plots presented in Fig. 18 given in [36] illustrate this fact.

Figure 19 given in [36] points out the drift of the forecast of volumes for Moldova. Figure 19 points out that there have been no strong oscillations since the Spring of 2021.

The drift of the forecast of dates for Moldova at 28.06.2021 is presented in Fig. 20 given in [36]. Figure 20 highlights that the drifts has started to rise two weeks prior to 28.06.2021.

Figure 21 given in [36] outlines the plots of the epidemic volumes for the Campania region in Italy at 28.10.2020. Figure 21 shows that the predicted peak load on the healthcare system was S = 148 462 on 09.12.2020.

4. Conclusions

This paper proposed the new opportunities model and modeling approach to monitor, analyze and forecast the COVID-19 using official statistics. The model is built upon the classical SI model and it allows to automatically determine in real time when the epidemic is developing naturally according to current opportunity and when there is a transition from one opportunity to another one.

The experimental results obtained for three different countries in terms of area, economy and population, namely Russia, Romania and Moldova, plus the Campania region in Italy, prove the superior performance of the suggested model in comparison with other similar models including the classical SI model. Although the data is rather old and related to 2021, the evolution of the pandemic afterwards till nowadays further confirms the very good performance of the model proposed in this paper.

The limitation of the model is the need to set the values of one of its parameters. This will be mitigated in the future research by its automatic computation as the solution to an appropriately defined optimization problem. Another direction of future research is the development of the karass model, which, unlike all other known models, allows to describe the multi-wave behavior of the pandemic, and it will replace the SI model in the current new opportunities model.

Acknowledgements. This work was supported by the grant of the Romanian Ministry of Education and Research, CNCS - UEFISCDI, project number PN-III-P4-ID-PCE-2020-0269, within PNCDI III, and by the national project of Republic of Moldova number 20.80009.5007.27.

References

[1] C. C. CHRISTARA and N. C.-H. LEUNG, *Analysis of quantization error in financial pricing via finite difference methods*, SIAM Journal of Numerical Analysis **56**(3), 2018, pp. 1731–1757.

[2] W. TIKJHA and L. GARDINI, *Bifurcation sequences and multistability in a two-dimensional piecewise linear map*, International Journal of Bifurcation and Chaos **30**(6), 2020, pp. 2030014:1–2030014:21.

[3] C. GOMES BEZERRA, B. S. J. COSTA, L. A. GUEDES and P. ANGELOV, *An evolving approach to data streams clustering based on typicality and eccentricity data analytics*, Information Sciences **518**, 2020, pp. 13–28.

[4] M. PRATAMA, CHOIRU ZA'IN, EDWIN LUGHOFER, ERIC PARDEDE and D. A. P. RAHAYU, *Scalable teacher forcing network for semi-supervised large scale data stream*s, Information Sciences **576**, 2021, pp. 407–431.

[5] F. G. FILIP, *Automation and computers and their contribution to human well-being and resilience*, Studies in Informatics and Control **30**(4), 2021, pp. 5–18.

[6] Z. C. JOHANYÁK, *Fuzzy rule interpolation based model for student result prediction*, Journal of Intelligent & Fuzzy Systems **36**(2), 2019, pp. 999–1008.

[7] J. VAŽČÁK, J. HVIZDOŠ and M. PUHEIM, *Agent-based cloud computing systems for traffic management*, Proceedings of 2016 International Conference on Intelligent Networking and Collaborative Systems, Ostrava, Czech Republic, 2016, pp. 73–79.

[8] S. BLAŽIČ, D. DOVŽAN and I. ŠKRJANC, *Cloud-based identification of an evolving system with supervisory mechanisms*, Proceedings of 2014 IEEE International Symposium on Intelligent Control, Antibes, France, 2014., pp. 1906–1911

[9] R.-E. PRECUP, C.-A. BOJAN-DRAGOS, E.-L. HEDREA, R.-C. ROMAN and E. M. PETRIU, *Evolving fuzzy models of shape memory alloy wire actuators*, Romanian Journal of Information Science and Technology **24**(4), 2021, pp. 353–365.

[10] I. ŠKRJANC, S. BLAŽIČ and P. ANGELOV, *Robust evolving cloud-based PID control adjusted by gradient learning method*, Proceedings of 2014 IEEE Conference on Evolving and Adaptive Intelligent Systems, Linz, Austria, 2014, pp. 1–6.

[11] P. BARANYI, Y. YAM and P. VARLAKI, *TP Model Transformation in Polytopic Model-Based Control*, Taylor & Francis, Boca Raton, FL, 2013.

[12] P. BARANYI, *How to vary the input space of a T-S fuzzy model: a TP model transformation-based approach*, IEEE Transactions on Fuzzy Systems **30**(2), 2022, pp. 345–356.

[13] E.-L. HEDREA, R.-E. PRECUP, R.-C. ROMAN and E. M. PETRIU, *Tensor product-based model transformation approach to tower crane systems modeling*, Asian Journal of Control **23**(3), 2021, pp. 1313–1323.

[14] C. POZNA, N. MINCULETE, R.-E. PRECUP, L. T. KÓCZY and Á. BALLAGI, *Signatures: Definitions, operators and applications to fuzzy modelling*, Fuzzy Sets and Systems **201**, 2012, pp. 86–104.

[15] C. POZNA and R.-E. PRECUP, *Aspects concerning the observation process modelling in the framework of cognition processes*, Acta Polytechnica Hungarica **9**(1), 2012, pp. 203–223.

[16] L. T. KÓCZY, M. E. CORNEJO and J. MEDINA, *Algebraic structure of fuzzy signatures*, Fuzzy Sets and Systems **418**, 2021, pp. 25–50.

[17] R.-E. PRECUP, G. DUCA, S. TRAVIN and I. ZINICOVSCAIA, *Processing, neural network-based modeling of biomonitoring studies data and validation on Republic of Moldova data*, Proceedings of the Romanian Academy, Series A: Mathematics, Physics, Technical Sciences, Information Science **23**(4), 2022, pp. 403–410.

[18] Y. YU and X. ZHANG, *Relation patterns extraction to climate data using long short-term memory networks*, Proceedings of the Romanian Academy, Series A: Mathematics, Physics, Technical Sciences, Information Science **23**(3), 2022, pp. 307–314.

[19] G. DUCA and F. MACAEV, *Compounds and Materials for Drug Development and Biomedical Applications* (in Romanian), Editura Academiei Romane, Bucharest, 2018.

[20] E. OSABA, J. DEL SER, A. D. MARTINEZ, J. L. LOBO and F. HERRERA, *AT-MFCGA: An adaptive transfer-guided multifactorial cellular genetic algorithm for evolutionary multitasking*, Information Sciences **570**, 2021, pp. 577–598.

[21] Z.-M. GU and G.-G. WANG, *Improving NSGA-III algorithms with information feedback models for large-scale many-objective optimization*, Future Generation Computer Systems **107**, 2020, pp. 49–69.

[22] J.-H. YI, L.-N. XING, G.-G. WANG, J.-Y. DONG, A. V. VASILAKOS, A. HOSSEIN ALAVI and L. WANG, *Behavior of crossover operators in NSGA-III for large-scale optimization problems*, Information Sciences **509**, 2020, pp. 470–487.

[23] E. OSABA, E. VILLAR-RODRIGUEZ, I. OREGI and A. MORENO-FERNANDEZ-DE-LECETA, *Hybrid quantum computing-tabu search algorithm for partitioning problems: preliminary study on the traveling salesman problem*, Proceedings of 2021 IEEE Congress on Evolutionary Computation, Kraków, Poland, 2021, pp. 351–358.

[24] Y.-H. FENG, S. DEB, G.-G. WANG and A. HOSSEIN ALAVI, *Monarch butterfly optimization: A comprehensive review*, Expert Systems with Applications **168**, paper 114418, 2021.

[25] R.-E. PRECUP, R.-C. DAVID, R.-C. ROMAN, E. M. PETRIU and A.-I. SZEDLAK-STINEAN, *Slime mould algorithm-based tuning of cost-effective fuzzy controllers for servo systems*, International Journal of Computational Intelligence Systems **14**(1), 2021, pp. 1042–1052.

[26] R.-E. PRECUP, R.-C. DAVID, R.-C. ROMAN, A.-I. SZEDLAK-STINEAN and E. M. PETRIU, *Optimal tuning of interval type-2 fuzzy controllers for nonlinear servo systems using slime mould algorithm*, International Journal of Systems Science, DOI: 10.1080/00207721.2021.1927236, 2021.

[27] Z. C. JOHANYÁK, *A modified particle swarm optimization algorithm for the optimization of a fuzzy classification subsystem in a series hybrid electric vehicle*, Technicki Vjesnik - Technical Gazette **24**(2), 2017., pp. 295–301.

[28] C. POZNA, R.-E. PRECUP, E. HORVATH and E. M. PETRIU, *Hybrid particle filter-particle swarm optimization algorithm and application to fuzzy controlled servo systems*, IEEE Transactions on Fuzzy Systems **30**(10), 2022, pp. 4286–4297.

[29] L. RODRÍGUEZ, O. CASTILLO, M. GARCÍA VALDEZ and J. SORIA, *A new meta-heuristic optimization algorithm based on a paradigm from physics: string theory*, Journal of Intelligent & Fuzzy Systems **41**(1), 2021, pp. 1657–1675.

[30] F. VALDEZ, O. CASTILLO and C. PERAZA, *Fuzzy logic in dynamic parameter adaptation of harmony search optimization for benchmark functions and fuzzy controllers*, International Journal of Fuzzy Systems **22**(4), 2020, pp. 1198–1211.

[31] Official Coronavirus Cases in Russia. Accessed: Sep. 15, 2021. [Online]. Available: https://www.worldometers.info/coronavirus/country/russia.

[32] Official Coronavirus Cases in Romania. Accessed: Sep. 15, 2021. [Online]. Available: https://www.worldometers.info/coronavirus/country/romania/.

[33] Official Coronavirus Cases in Moldova. Accessed: Sep. 15, 2021. [Online]. Available: https://www.worldometers.info/coronavirus/country/moldova/.

[34] Official Coronavirus Cases in Italy. Accessed: Sep. 15, 2021. [Online]. Available: https://www.worldometers.info/coronavirus/country/italy/.

[35] S. ABRAMOV and S. TRAVIN, *Modeling and forecast of the coronavirus epidemic statistics in Russia*, Digital Economy **2**(10), 2020, pp. 5–18.

[36] S. M. ABRAMOV, S. TRAVIN, G. DUCA and R.-E. PRECUP, Supplementary material of the paper Sergey M. Abramov, Sergey Travin, Gheorghe Duca, Radu-Emil Precup, *New Opportunities Model for Monitoring, Analyzing and Forecasting the Official Statistics on Coronavirus Disease Pandemic*, Romanian Journal of Information Science and Technology, 2023. Accessed: Feb. 1, 2022. [Online]. Available: http://www.aut.upt.ro/∼rprecup/Supplementary_material_ROMJIST.doc.