Research Article

Research of adequacy of models and SIRS based on cellular automata

Dmytro Lande [a], *Leonard Strashnoy* [b] (✉)

[a] *Institute for Information Recording of NAS of Ukraine*
[b] *UCLA, Institute Infectious disease department, USA*

ABSTRACT:

This paper describes a comparison of the SIRS model, built on the basis of cellular automata, and the real statistics of daily mortality in the course of the COVID-19 pandemic in different countries. The SIRS model is used with different parameters of infectability and loss of susceptibility. The model calculates the cross-correlation of mortality dynamics with the real dynamics of mortality in different countries. A high correlation (more than 0.5) of mortality dynamics in all countries with the dynamics given by the model at the selected parameters is shown.

The research results allow us to fix the model parameters that can further perform forecasting. The advantage of the SIRS model based on cellular automata is the simplicity and clarity of a small number of parameters, and the ability to change them in accordance with epidemiological data. The model demonstrates the fact that the daily mortality rate is eventually reduced to zero, although under different parameters (primarily the probability of loss of immunity), the system can survive several waves of infection.

KEYWORDS:

COVID-19, pandemic, modeling, SIRS, cross-correlation, cellular automata, datasets

Introduction

Among the many methods of modeling the spread of infections and pandemics, the most well-known approaches are called SIS, SIR, and SIRS [1-3]. While most such models are currently modeled by analytical methods using differential equations, the most obvious implementations of these approaches are based on the theory of cellular automata [4], which are supplemented with analytical calculations if necessary. The SIS "susceptible — infected — susceptible" model (S for the number of susceptible, I for the number of infectious) is applicable to the analysis of the spread of diseases to

✉ Corresponding Author: E-mail: lstrashnoy@gmail.com

which immunity is not developed. It is described by the following system of equations:

$$\frac{dS(t)}{dt} = -\beta \frac{S(t)I(t)}{N} + \gamma I(t),$$
$$\frac{dI(t)}{dt} = \beta \frac{S(t)I(t)}{N} - \gamma I(t). \quad (1)$$

where

- $S(t)$ — the number of susceptible individuals at time *t*;
- $I(t)$ – the number of infected individuals at time *t*;
- *N* – population size;
- β — coefficient of the intensity of contacts of individuals with subsequent infection;
- γ — coefficient of the intensity of the recovery rate of infected individuals

In the second case, the SIR model consists of three compartments: S for the number of Susceptible, I for the number of Infectious, and R for the number of Recovered, (or deceased, or immune) individuals.

The SIR model has gained popularity due to its ease of construction, implementation, and use. Its application allows you to accurately model influenza and other diseases in large cities, enter new parameters, and analyze different scenarios.

In the third, even more realistic case, the probability of loss of immunity in previously infected individuals is assumed. The SIRS principle - "susceptible -- infected – recovered – susceptible", is a model for describing the dynamics of diseases with temporary immunity (recovered individuals become susceptible again over time).

Goals

The purpose of this paper is to compare the SIRS model, based on the concept of cellular automata discussed below, and the actual daily mortality statistics for various countries in the course of the COVID-19 pandemic. For this purpose, the SIRS model is built, which is used for various parameters of infectability and loss of susceptibility. Then the cross-correlation of the mortality dynamics that the model provides is calculated with the real mortality dynamics for different countries. For each country, the model parameters for which the cross-correlation values are sufficiently large are fixed, and the model parameters that allow for further forecasting are likewise fixed.

Cellular automata

A cellular automaton is a discrete dynamic system, a collection of identical cells that are interconnected in a certain way. All cells form a network of cellular automata. The state of each cell is determined by the state of cells in its local neighborhood or its "nearest neighbors". The state of the *j*-th cellular automaton at a time $t+1$ is thus determined as follows: $y_j(t+1)=F\left(y_j(t),O(j),t\right)$, where *F* is a rule that can be expressed, for example, in the language of Boolean algebra. In many problems, it is considered that the element relates to its closest neighbors, i.e. $y_j(t)\in O(j)$, in this case, the formula is simplified: $y_j(t+1)=F\left(O(j),t\right)$.

Cellular automata in the traditional sense satisfy such rules:

- the values of all cells are changed simultaneously (the unit of measurement is a clock cycle).
- the network of cellular automata is homogeneous, i.e. the rules for changing states for all cells are the same;
- the cell can only be affected by cells from its local neighborhood – the set of cell states is finite.

In the case of a two-dimensional lattice, whose elements are squares, the nearest neighbors entering the neighborhood of the element y_{ij} can be considered either only the elements located up-down and left-right from it (the so-called von Neumann neighborhood: $y_{i-1,j}, y_{i,j-1}, y_{i,j+1}, y_{i+1,j}$), or diagonal elements added to them (the so-called "Moore's neighborhood": $y_{i-1,j-1}, y_{i-1,j}, y_{i-1,j+1}$, $y_{i,j-1}, y_{i,j}, y_{i,j+1}$, $y_{i+1,j-1}, y_{i+1,j}, y_{i+1,j+1}$). This allows us to determine the overall ratio of the cell value at step t+1 compared to step t [4]:

$$y_{i,j}(t)=F(y_{i-1,j-1}(t),y_{i-1,j}(t),y_{i-1,j+1}(t),y_{i,j-1}(t),y_{i,j}(t),y_{i,j+1}(t),y_{i+1,j-1}(t),y_{i+1,j}(t),y_{i+1,j+1}(t)).$$

The SIRS Model

Such a model is constructed, and is given the following semantic meaning: a system of square cellular automata (Moore's model) with four cellular states is considered:

1. Non-infected state
2. Infected state
3. Affected state
4. Death

Rules:

1. A cell in an infected state infects its neighbors with a probability p_1 only those cells in a state of 1 (non-infected state).
2. A cell in an infected state goes to state 4 (death) with a probability of p_3. With the probability of $1-p_3$ it is likely to go into an immunized state.
3. The immunized cell loses its immunity and goes to state 1 with the probability of p_2.

The MATLAB system was used to implement the model [5].

In contrast to the standard SIRS model, where the value is *N=S+I+R*, in the model under consideration, *N= S+I+R+D*, where D is the number of deaths. This paper specifically examines the dynamics of this parameter.

In Figure 1. The States of the cellular automaton (a-d) are given. In particular in Fig. 1 (C) it is possible to observe the emergence of a second wave of the pandemic.

In particular, the results of the simulation show that rule 3 is important. When $p_2 \le 0.05$ the epidemic quickly subsides, otherwise – it can continue for quite a long time, leading to a large number of epidemic waves, and eventually leading to the destruction of a significant proportion of the population.

Sources of Information and Timelines

This work uses data from the World Health Organization (WHO) as a source of information https://www.who.int/emergencies/diseases/novel-coronavirus-2019/ situation- reports/ and in Our World In Data (is a project of the Global Chang Lab) [6]. Within this source (https://ourworldindata.org/coronavirus-source-data) data is daily renewed and presented in a “cleaned up” format, without the anomalous spikes that have been connected with technical failures, with a high level of integration, in the formats XLSX (https://covid.ourworldindata.org/data/owid-covid-data.xlsx), CSV, JSON.
Addresses of individual data arrays:

1. Total confirmed cases: https://covid.ourworldindata.org/data/ecdc/total_cases.csv
2. Total deaths: https://covid.ourworldindata.org/data/ecdc/total_deaths.csv
3. New confirmed cases: https://covid.ourworldindata.org/data/ecdc/new_cases.csv
4. New deaths: https://covid.ourworldindata.org/data/ecdc/new_deaths.csv

5. All four metrics:
   https://covid.ourworldindata.org/data/ecdc/full_data.csv
6. Population data:
   https://covid.ourworldindata.org/data/ecdc/locations.csv

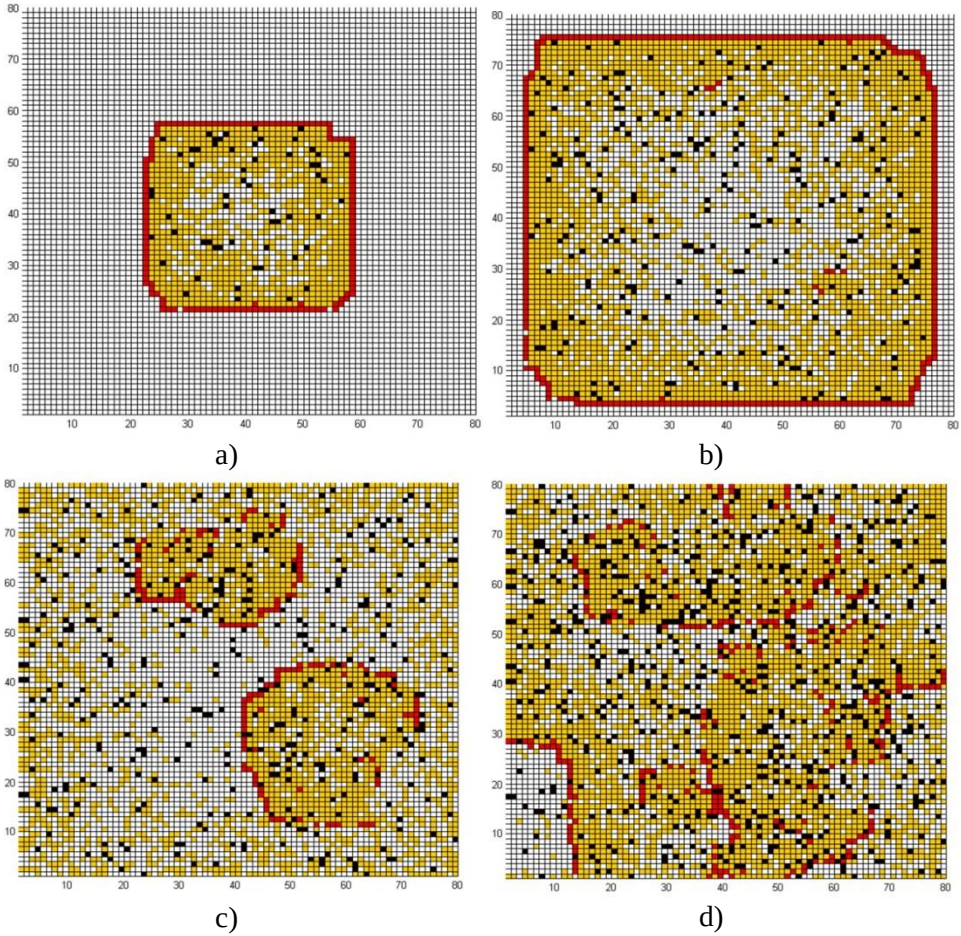


Figure 1: The intermediate states of cellular automata systems at various stages of the epidemic (light cells – susceptible (healthy), yellow – immunized, red – infected (sick), black (deceased)

In further studies for each country, the data downloaded from the aggregator source file determine the vectors of the dynamics of the pandemic process for a certain period (selected from 01/04/2020 to 06/07/2020), corresponding to mortality.

Correlation between the dynamics of the model and real data

To assess the adequacy of the proposed model, a cross-correlation of the dynamics of mortality in the model (the number of black cells per clock cycle) and real data on mortality in various countries is calculated [7]. In this case, the simulation results depend on the parameters p_1, p_2 and p_3. In the presented simulation results and the subsequent comparison with real data, the parameter p_3 (probability of death in case of infection) was fixed at the level of 5%. For each pair of parameter values from the range $(0.3 \le p_1 \le 1.0;\ 0.01 \le p_2 \le 0.1)$, the dynamics of the appearance of black cells was calculated. For each of these dynamics vectors, a cross-correlation was calculated with the dynamics of mortality in the selected country. The maximum cross-correlation and corresponding parameter values were selected for each country p_1 and p_2 (Table 1).

The set of maximal cross-correlations between the obtained vectors is calculated, and the corresponding correlation matrix is formed with the elements in the notation of the formula (1):

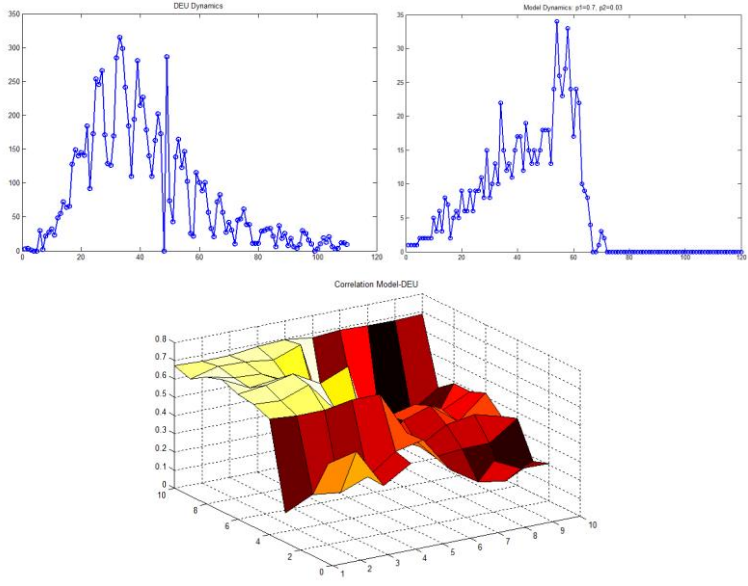
$$a_{ij}(m) = \max_m \frac{\sum_{k=1}^{n-m} w_{k+m}^i w_k^j}{\sqrt{\sum_{k=m+1}^{n} \left(w_k^i\right)^2} \sqrt{\sum_{k=1}^{n-m} \left(w_k^j\right)^2}}. \qquad (2)$$

The max function is used for the reasons that processes that are similar in nature may have similar dynamic behavior, but possibly with a time shift [8].
Figure 1 shows some examples of the dynamics of the appearance of black cells in the model under study, the dynamics of real mortality processes, and the surface of cross-correlations with different infection probabilities (p_1) and immune loss (p_2) with a mortality probability of 0.05.

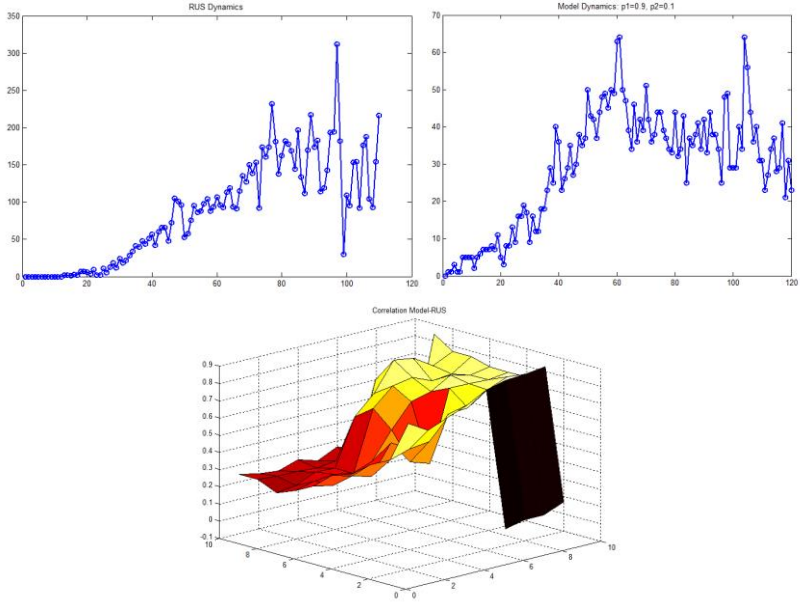
As can be seen, for almost all countries, the maximum cross-correlation value exceeds 0.5, which indicates the adequacy of the model and its predictive capabilities.

Table 1. Values of the maximum cross-correlation of model parameters for some countries (Country names in accordance with [9])

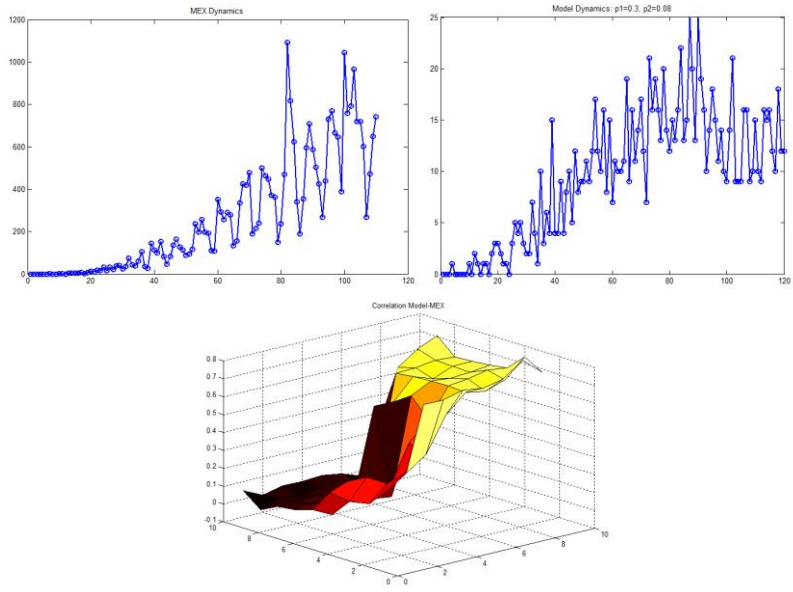
Country code	Parameter " Probability of infection»	Parameter " Probability of loss of immunity»	The calculated maximum cross-correlation
BGD	0.3	0.1	0.813
GTM	0.3	0.1	0.726
COL	0.3	0.1	0.651
BRA	0.3	0.1	0.812
EGY	0.3	0.07	0.701
IDN	0.3	0.08	0.637
PAK	0.3	0.08	0.752
MEX	0.3	0.08	0.747
AFG	0.3	0.09	0.625
IND	0.3	0.09	0.67
PER	0.3	0.09	0.802
SAU	0.3	0.09	0.798
TUR	0.5	0.01	0.764
ISR	0.5	0.01	0.629
ITA	0.5	0.03	0.687
NOR	0.5	0.03	0.666
SWE	0.5	0.03	0.54
ECU	0.5	0.05	0.543
JPN	0.6	0.01	0.591
ESP	0.6	0.01	0.543
GBR	0.6	0.02	0.674
AUS	0.6	0.03	0.64
AUT	0.6	0.03	0.675
IRL	0.6	0.03	0.604
DEU	0.7	0.03	0.727
BEL	0.7	0.04	0.761
USA	0.7	0.04	0.626
PRT	0.7	0.04	0.643
UKR	0.8	0.09	0.661
IRQ	0.9	0.1	0.729
RUS	0.9	0.1	0.765
NLD	1	0.03	0.707
DNK	1	0.08	0.696
GRC	1	0.08	0.515
FRA	1	0.09	0.675
IRN	1	0.09	0.545
KOR	1	0.09	0.645



a)



b)



c)

Figure 2: Graphs of the real dynamics of mortality, the dynamics obtained in the simulation, the values of cross-correlations for various parameters of the probability of infection (p_1) and the probability of loss of immunity (p_2) with the probability of mortality 0.05: a – Germany; b – Russia; c – Mexico

Conclusions

The purpose of this paper is to compare the SIRS model, based on the concept of cellular automata discussed above, and the actual daily mortality statistics for various countries in the course of the COVID-19 pandemic. For this purpose, the SIRS model is built, which is launched for various parameters of infectability and loss of susceptibility. Then we calculate the cross-correlation of the mortality rate dynamics that the model gives with the real mortality rate dynamics for different countries. For each country, the model parameters for which the cross-correlation values are sufficiently large are fixed, and likewise, the model parameters are fixed that allow for further forecasting. In addition, the correlation of the dynamics of different countries with a model with the same parameters can be considered as a basis for conducting cluster analysis.

The advantage of the SIRS model is the simplicity and clarity of its parameters, and the ability to change them. It is also possible to convert the model to an analytical one, which will make it possible to change the number of "neighbors" during local interactions. Also, there are the following advantages: a small number of parameters consistent with the parameters of

the theory of epidemiology, a relatively low dimension of the vectors of parameters, a reliable mathematical basis for correlation analysis, objectivity – a reliable aggregator of data answers for the "purity" of the data, the application of standard software tools, and its relative simplicity (ready-made software systems such as Matlab, Excel, R language, etc., can be applied).

The model demonstrates the obvious fact that mortality is eventually reduced to zero, although under various parameters (primarily the probability of loss of immunity), the system can survive several waves of infection, and, accordingly, either a significant epidemic time with relatively small mortality or high mortality with a relatively short period of the epidemic is possible.

The disadvantage of the model is that it only takes into account local interactions. At the same time, it can be expanded. The model also does not provide for the ranking of population members by infection/infectability.

References

1. Harko, Tiberiu; Lobo, Francisco S. N.; Mak, M. K. Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. Applied Mathematics and Computation (2014). 236: 184–194. DOI: 10.1016/j.amc.2014.03.030
2. Greenhalgh D. & Moneim I.A. SIRS Epidemic Model and Simulations Using Different Types of Seasonal Contact Rate. Systems Analysis Modeling Simulation (2003). Vol. 43. Iss. 5. PP. 573-600. DOI: 10.1080/0232902100000813
3. Minglei Fu, Hongbo Yang, Jun Feng, Wen Guo, Zichun Le, Dmytro Lande, Dmytro Manko. Preferential information dynamics model for online social networks (2019). Physica A: Statistical Mechanics and its Applications. Vol. 506. pp. 993-1005. DOI: 10.1016/j.physa.2018.05.017
4. S. Wolfram “New Kind of Science”. Wolfram Media (2002). – 1259 p. ISBN: 1579550088, 9781579550080
5. Alexander D. Poularikas “Understanding Digital Signal Processing with MATLAB® and Solutions”. CRC Press (2017). 472 p. ISBN: 9781315112855
6. Our World In Data Project. https://ourworldindata.org/coronavirus-source-data
7. Lande, Dmitry and Strashnoy, Leonard, Cross-Correlation of Publications Dynamics and Pandemic Statistics (June 12, 2020). Available at SSRN https://ssrn.com/abstract=3625725. DOI: 10.2139/ssrn.3625725
8. Strashnoy, Leonard and Lande, Dmitry, Networks of Countries Defined by the Dynamics of the COVID-19 Pandemic (July 8, 2020). Available at SSRN: https://ssrn.com/abstract=3647570. DOI: 10.2139/ssrn.3647570
9. ISO 3166-1 alpha-3. URL: https://www.iso.org/iso-3166-country-codes.html.