

An Iterative Model of The COVID-19 Virus Spread Based on Population Density - A Case Study of Poland

Bernard TWARÓG (✉ btwarog@pk.edu.pl)

Cracow University of Technology

Research Article

Keywords: COVID-19, hazard development, Monte Carlo methods, spatial analyses

Posted Date: February 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-153522/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

An iterative model of the COVID-19 virus spread based on population density - a case study of Poland

Bernard Twaróg

Chair of Geoengineering and Water Management, Faculty of Environmental and Power Engineering, Cracow University of Technology, 31-155 Cracow, WarszawskaStr.24, Poland.
Email: btwarog@pk.edu.pl

Abstract

The correct recognition of spread patterns of the COVID-19 virus is crucial for effective decision-making on preventive measures to be taken and protection of society. A forecast of hazard development, substantiated by reliable virus propagation models, can support control of the infection risk and minimize its impact on the population. Decisions to close or limit social activities or to lift bans can then be made in the right area at the right time. This study proposes an iterative agent-based and compartmental model used to analyse the COVID-19 virus spread. The contagion mechanism is based on population density data and Monte Carlo methods. The model parameters were optimized and verified using data about incidence in Poland between March 3 and November 24, 2020. To obtain optimum values of model parameters, a simulation and analysis were completed for three vaccination strategies begun on December 27, 2020. This study contains numerical results and spatial analyses that represent and confirm the accuracy of assumptions of the proposed model. Certain limitations of the proposed approach and planned directions of future work are also indicated in the final section.

1. Introduction

A new variant of coronavirus named SARSCoV-2 (Severe Acute Respiratory Syndrome Coronavirus) ^{1,2} began to spread by the end of 2019 in the city of Wuhan ³, Hubei province, China ⁴. The World Health Organization (WHO) announced in January 2020 a new coronavirus epidemic posing a hazard to public health on an international scale ⁵. Considering the contagious disease caused by the new coronavirus, the WHO chose in February 2020 the official name - COVID-19 (an abbreviation of Coronavirus Disease 2019) ², and on March 11, 2020 announced the pandemic of COVID-19 ⁶. Three coronaviruses characterized by rapid transmission and pathogenic properties appeared over the last 20 years, causing the severe acute respiratory syndrome: SARS-CoV (Severe Acute Respiratory Syndrome) ⁷, MERS-CoV (Middle East Respiratory Syndrome Coronavirus) ⁸ and the recently identified new coronavirus, temporarily termed 2019-nCoV ¹, that on February 11, 2020 was finally named COVID-19 ², and since February 14, 2020 the pathogen causing that disease is known as the SARS-CoV-2 ⁹ virus.

Coronaviruses cause various diseases in mammals and birds, beginning with inflammatory bowel disease and conditions of the upper respiratory tract, ending with potentially fatal infections of the respiratory tract in humans ¹⁰. There are coronavirus variants that can affect animals, and in exceptional cases the coronavirus can cross the inter-species barrier to colonize the human population ⁸. It is highly probable that the new coronavirus could cross the barrier between an animal species and people to spread in the human population ¹¹.

2. Severity of the pandemic

The best-known historical pandemic was plague in the mid-14th century (Black Death, 1347-1351) that killed about 25% to 50% of the then population (about 200 million deaths) ^{12,13}. The dynamic development of medicine at the turn of the 20th and 21st centuries reduced the frequency of epidemics and pandemics. However, pandemics are not rare phenomena these days. Due to anthropogenic impact on the environment, that has intensified for decades, and the emergence of new conditions for pathogen transmission ⁹, several pandemics occurred over the last ten years: HIV/AIDS (Human Immunodeficiency Virus Infection and Acquired Immunodeficiency Syndrome) ¹⁴, A/H1N1 (species Influenza A virus of the genus Influenzavirus A) ¹⁵, SARS (Severe acute respiratory syndrome-related virus of the genus Betacoronavirus) ¹⁶, EBOLA (EBOV ⁷) (the hemorrhagic fever caused by the Ebola virus) ¹⁵. Now in 2020, the humankind faces the most severe crisis caused by COVID-19. The severity of a pandemic is indicated by the number of fatal cases. Despite varying numbers of fatal cases published by the health-monitoring institutions (CDC, WHO, Encyclopaedia Britannica, Johns Hopkins University et al.), the statistics of the pandemics remain horrifying. The pandemic of HIV/AIDS caused 25 to 35 million deaths, H1N1/09 caused 18 thousand deaths ¹⁶, (while the CDC Centers for Disease Control and Prevention estimate the number at 151 thousand to 575 thousand deaths ^{16,17}), SARS caused 774 deaths, EBOLA - 11.3 thousand deaths. The pandemic of COVID-19 coronavirus disease that began in December 2019 in Chinese Wuhan afflicted almost 80.78 million people and killed 1.76 million people (as of December 28, 2020) all over the world ¹⁸.

A pandemic is an outbreak of infectious disease that occurs over a wide geographical area and that is of high prevalence ¹⁹. An epidemic afflicts numerous people, transmitted simultaneously between them, but remains limited to a region. A pandemic exposes to the hazard of infection entire countries, continents or the world.

The correct recognition of spread patterns of the COVID-19 virus is crucial for effective decision-making on preventive measures to be taken and immediate protection of society. A forecast of hazard development, substantiated by reliable virus propagation models, can support control of the infection risk by minimizing its impact on the population. Decisions to close or limit social activities or to lift bans can then be made in the right area at the right time. Precise demographic data, good laboratory infrastructure (for performing tests) and logistics can contribute to the understanding of incidence, spread of the illness and its trends. This knowledge can support a rapid, precise and effective response of the public health sector to the COVID-19 pandemic ²⁰.

3. Models used to describe the pandemic

The contemporary modelling techniques and available computing power made a valuable contribution to the understanding of epidemic and pandemic spread phenomena ¹⁶. The models currently used are divided into two categories: agent-based models ²¹⁻²³ and compartmental models) ²⁴⁻²⁶.

The agent-based models (ABM) ²¹⁻²³ represent computational structures on a system (macro) scale described by behaviours of agents on a micro scale. "Agents" may include both individuals and collective entities, employees of services, units, particles or distinguished quantitative components such as social groups, institutions or organizations.

A typical ABM consists of three components: agents, the environment and the rules followed by an agent plus agent's mutual local interactions with other components of the system and environment. Due to the development of advanced computer techniques, agent-based modelling has become an attractive tool in research into complex and dynamic systems, especially in the life sciences ²².

The compartmental models ^{21,24} are characterized by a simple notation and effectiveness in the modelling of epidemiological hazards. They simplify mathematical modelling of contagious diseases. The population is assigned to compartments characterized by status labels, e.g. (susceptible, infectious, recovered) ²⁵. Groups of people can be transferred between compartments and exposed

to factors that assign them defined status labels. The order of labels determined in a pattern is essential. The typical patterns used include: SIR (Susceptible, Infected, Recovered), SEIS (Susceptible, Exposed, Infected, Susceptible) and SEIR (Susceptible, Exposed, Infected, Recovered), qualified as repeatedly susceptible ²⁷.

The models date back to the early 20th century and were developed in the study published by Kermack and McKendrick in 1927 ²⁸. The models are usually described using simple differential equations but can also be expanded by stochastic (random) components that are more realistic and far more difficult to analyse.

The models are built to predict the pattern of disease spread, define the current and total number of infected people, the number of fatal infections, the number of recovered patients and duration of the epidemic. The modelling process leads to estimates of various epidemiological parameters and of effects of preventive measures. It enables us to assess how various actions taken by the public health sector affect the outcome of the epidemic, what is the result of limited mobility, the result of direct protection measures and what is the most effective strategy of giving a limited number of vaccines in the population ^{24,29}.

The authors of agent-based models frequently criticize the compartmental models due to their difficulties in reflecting the real characteristics and relations between various areas, regions, in describing population characteristics, such as population distribution ^{30,31}.

This study proposes an iterative agent-based and compartmental model used to describe and analyse the spread of the COVID-19 virus. The contagion process has been expanded by Monte Carlo methods ³².

The model is based on random infection and mobility. It represents an expanded classical SIRS epidemic model ²⁹. The model takes into account the characteristics of real world described by the distribution of population in the analysed region. Importantly, it also considers interrelations between various unit areas and those interrelations represent crucial factors that accelerate the process of transformation of an epidemic into a pandemic and pandemic development. The model describes a two-dimensional network of unit areas wherein cells, due to their mobility characterized in a random manner, are capable of infecting cells in other unit areas. The proposed model is based on transmission of the disease described by a contagion function. The pattern of statuses that may be assumed by the cells in the population includes the labels "healthy", "susceptible", "infected", "recovered", "dead" and "vaccinated". The healthy cells represent a set in the entire population that may be exposed to the risk of infection. The susceptible cells include the entire population of cells minus dead, recovered and vaccinated cells. It is assumed that the recovered and vaccinated cells will not join the group of susceptible cells. At present, there is no information about acquisition by people of perennial immunity to the SARS-CoV-2 virus, because it is known for a dozen of months only. The acquisition of a pool of antibodies following contact with the virus or effective vaccination gives immunity for at least several months ^{9,33,34}. The relationship between the time of acquired immunity and the dozen-month period of research into the SARS-CoV-2 makes the above assumption plausible. The infected cells can become recovered or dead. Vaccines can be received by healthy cells.

4. Data sets

A model of population density in Poland in 2020 was adopted on the basis of data obtained by the University of Louisville and Columbia University in the United States ³⁵ and verified using data gathered in the Demographic Yearbook of Poland ³⁶. The generally accessible data set can be collected in the GEOTIFF or ASCII XYZ format with a resolution of 30[arc]. That data resolution approximates population density per 1km². Data mapping is done in WGS84. One pixel represents the number of people on 1 km². In the described model, the analysed sample population was written as a matrix with dimensions 698x1202 elements, each representing a unit area of 1km². According to

the Demographic Yearbook of Poland ³⁶, the population of Poland stands at 38 382.7 thousand (Fig.1).

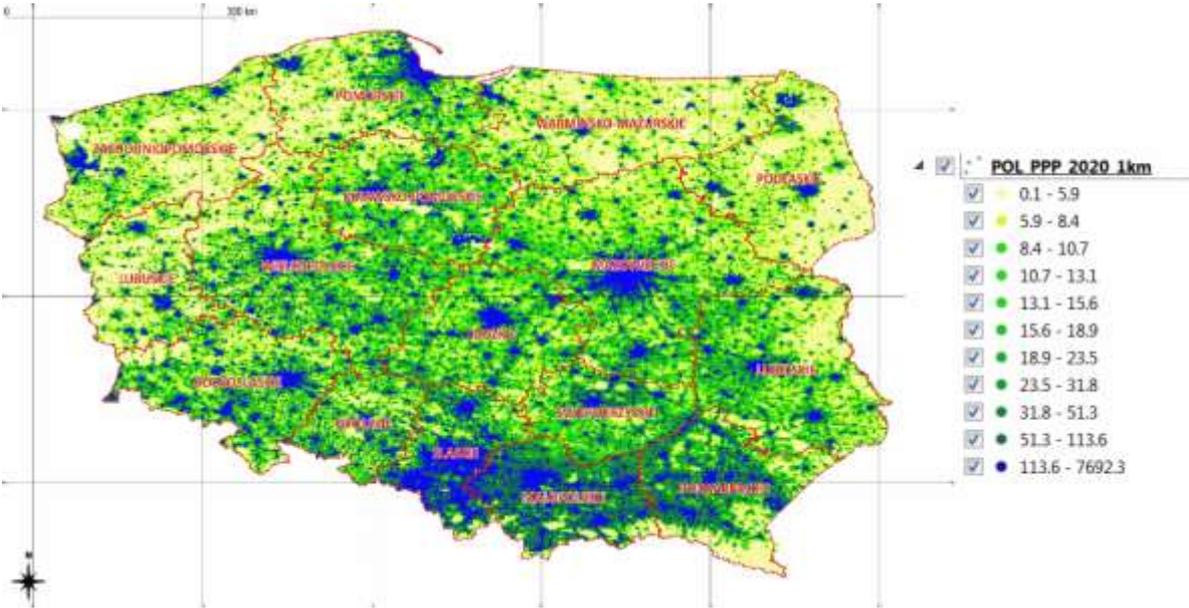


Figure 1. Population density per 1 km²

The data describing infection spread (Fig. 2) and the real status of exposure in the period between March 3, 2020 and November 24, 2020 were implemented based on daily updated sets published at gov.pl “Serwis Rzeczypospolitej Polskiej” ³⁷.

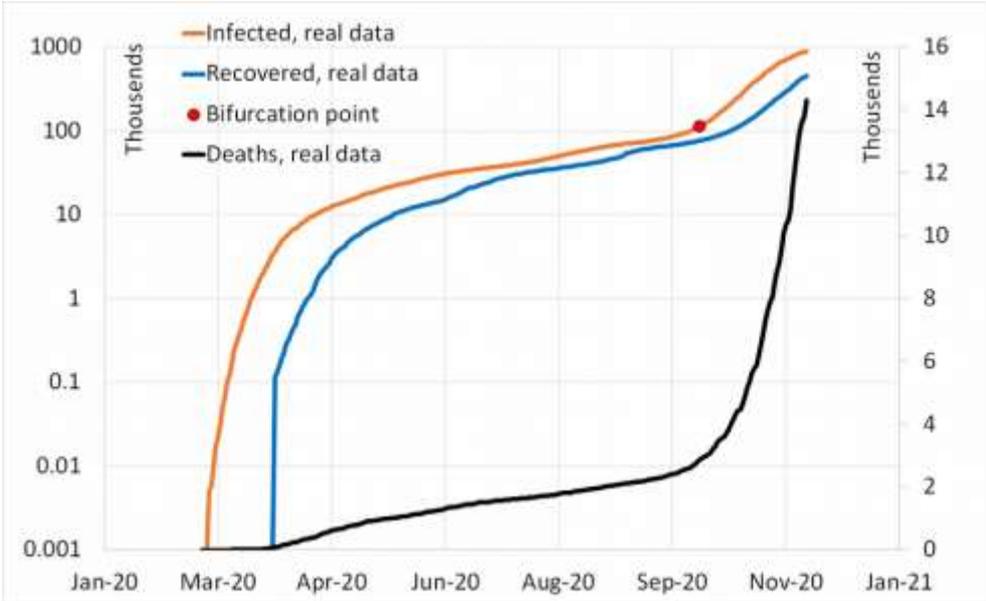


Figure 2. COVID-19 infections in the population of Poland between March 3, 2020 and November 24, 2020

5. Social and economic conditions during the pandemic

The social and economic conditions in Poland have been heavily influenced by pandemic pressure since March 2020. The most serious problems include the aversion to social contacts and

limited mobility as well as restrictions imposed on social activities and functions. These responses are dictated principally by the fear of infection and unknown development of the disease that may be mild and asymptomatic but also serious and fatal. At present, nobody can predict the variant of disease development in an individual.

The limitations imposed on activity of bars, restaurants, pubs, hotels and guest houses, restrictions affecting the gym and fitness industry and swimming pools arouse growing social discontent. On the other hand, the measures implemented by government, aimed to establish emergency health care centres, to limit social activities and functions, to provide businesses with funds, support continued existence of society.

Due to replacement of the sanitary regime in Poland by a more liberal policy (May / June 2020) ⁹ abolishing the obligation to wear masks and keep social distance, and introducing completely free market, the spread of disease, contained until May 2020, was released in society. This is clearly illustrated by a rapid increase in the number of infections in Fig. 2. The period between March 2020 and May 2020 can be characterized as a time of a controlled increase in infections. Lifted bans and the effects of a less restrictive sanitary regime are reflected in the curve of infections with a delay. The bifurcation point is observed in October 2020. Thereafter, the curve illustrating the number of new infections indicates a completely new, dynamically growing trend.

Considering potential threats, recent research projects demonstrate that the health care system may collapse four weeks after the outbreak of coronavirus epidemic, and this involves an uncontrolled increase in infections and a dramatic growth in mortality rate. An example is provided by the growing problems of the health care system in the United States ³⁸ and the collapse of the health care system in Italian Lombardy and in Spain ³⁹. At present, there is no reason to argue that without the measures aimed to maintain social distance (masks, physical distance, limited mobility) and without improved immunity acquired by vaccination the spread of disease will be contained and will not cause a dramatic growth in incidence and in fatal cases. The documented historical pandemics killed significant percentages of afflicted populations ¹⁹.

6. A description of components of the COVID19 spread model considering population density - a case study of Poland

The concept of contagion model is based on an iterative description of infection spread, incorporating a random process (Eq. 1). The model is used to process information contained in a matrix of population with a defined density per 1km². Only the susceptible cells can be infected. The number of infected cells among the susceptible cells is defined by the value of contagion function (Eq. 2) that depends on time. The contagion function has the form of two polynomials of degree 2. The bifurcation point ^{29,40}, 220 days that passed since the beginning of the pandemic between March 3 and October 8, 2020, is determined by the curve of infection cumulation in Fig. 2.

$$I_{(i+\Delta i, j+\Delta j), t+1} = I_{(i+\Delta i, j+\Delta j), t} + I_{ij, t} S_{(i+\Delta i, j+\Delta j), t} f(z, t) \quad (1)$$

$$f(z, t) = \begin{cases} z_1 t^2 + z_2 t + z_3, & \text{dla } t < t_1 \\ z_4 t^2 + z_5 t + z_6, & \text{dla } t \geq t_1 \end{cases} \quad (2)$$

$z(z_1, \dots, z_6, z_7)$ – decision variable vector,

z_1, \dots, z_6 – contagion function coefficients,

z_7 – variance, an element in the covariance matrix,

i, j – coordinates of the element in the population matrix,

$I_{ij, t}$ – the sum of cells infected by time t in element i, j ,

$I_{(i+\Delta i, j+\Delta j), t+1}$ – the sum of cells infected by time $t + 1$ in element $(i + \Delta i, j + \Delta j)$,

$(i + \Delta i, j + \Delta j)$ – coordinates of infected cells,

$(\Delta i, \Delta j)$ – coordinates of a random vector,

$S_{(i+\Delta i, j+\Delta j), t}$ – the number of susceptible cells,

$f(z, t)$ – the contagion / infection function,

t_1 – the bifurcation point, the time of change in the nature of contagion from controlled into uncontrolled (the 220th day of the pandemic, October 8, 2020 is assumed),

t – time.

The structure of the proposed iterative model of SARSCoV-2 contagion is suitable for calculating all components / cell statuses (healthy, infected, dead, recovered and vaccinated) for individual cases in unit areas at a moment.

6.1. The random model of infection

The infection process is described by a random vector of contagion, disease spread. The coordinates of infection spread are sampled using the Monte Carlo method and a two-dimensional normal distribution (Eqs. 3-8). For each unit area (an element of the matrix with the coordinates i, j), the process of sampling of the number of vectors of infection spread $(\Delta i, \Delta j)$ is repeated. Thus, each of the areas can affect itself and all other areas.

A multi-dimensional normal distribution can be described by:

$$f_{\mu, \Sigma}(x) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right) \quad (3)$$

$n = 2$

$x = (x_1, x_2)^T$ – a random vector created using the Monte Carlo method

$f_{\mu, \Sigma}(x)$ – the probability density function,

Σ – the matrix of covariance

$\mu \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ – the expected value vector, for the two-dimensional case.

The matrix of covariance:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix} \quad (4)$$

$$\rho = 0 \rightarrow \rho \sigma_1 \sigma_2 = 0 \quad (5)$$

$$\sigma_1^2 = \sigma_2^2 = z_7 \quad (6)$$

Final assumptions:

$$\Sigma = \begin{bmatrix} z_7 & 0 \\ 0 & z_7 \end{bmatrix} \quad (7)$$

$$\mu = 0 \quad (8)$$

The expected value vector was assumed a priori $\mu = 0$. This assumption results from the adopted contagion process originating in a region located symmetrically relative to the current element (i, j) with an area of 1km^2 . The model assumes mobility, but a cell returns to its original place of stay within the same time unit. Consequently, only propagation of the disease takes place without changes in the density of cells per area unit. This assumption is adopted considering the avoidance of travel by cells due to the pandemic and the restrictions imposed on mobility. The assumption of limited mobility excludes cross-border contagion.

Assuming that the direction of infection spread is not distinguished, the adopted value of Pearson correlation coefficient equals 0, and hence the elements outside the main diagonal of the matrix of covariance equal zero. The values on the main diagonal were initially defined based on ^{41,42} and then determined using an optimization procedure. Variances, Z_7 , were adopted at a level of 40000 km², and this in view of the restriction on mobility and the aversion to travel gives the same standard deviations in direction X and direction Y that equal 200 km.

6.2. A model of fatal cases and recoveries

The simplest available mortality model (Eq. 9) is applied in the described example. A constant, independent of time coefficient of fatal cases to infected cases is adopted. An analysis of the number of fatal cases shows that the coefficient equals 1.588% of the total number of infections in a 95% confidence interval (1.544% to 1.631%) ⁴³. Considering the iterative model of infection spread, the value of mortality coefficient was defined so that the total of fatal cases in the analysed time interval of 267 days equals the total of actual cases. The value adopted: 5.3494e-4.

$$M_{ij,t+1} = M_{ij,t} + I_{ij,t} * COEFF_{MORT} \quad (9)$$

$M_{ij,t+1}$ – fatal cases in element i, j in time $t + 1$,

$M_{ij,t}$ – fatal cases in element i, j in time t ,

$COEFF_{MORT}$ – fatal cases coefficient.

A similar method was used to estimate the number of recoveries (Eq. 10). A constant coefficient is adopted that describes the number of recoveries relative to the total of infected cases. An analysis of the number of recoveries shows that the coefficient equals 44.98% of the total number of infections in a 95% confidence interval (44.09% to 45.86%) ⁴³. In the model of infection spread, the value of recovery coefficient was defined so that the total of recoveries in the analysed time interval of 267 days equals the total of actual cases. The value adopted: 1.6994e-2.

$$R_{ij,t+1} = R_{ij,t} + I_{ij,t} * COEFF_{RECOV} \quad (10)$$

$R_{ij,t+1}$ – recovered cases in element i, j in time $t + 1$,

$R_{ij,t}$ – recovered cases in element i, j time t ,

$COEFF_{RECOV}$ – recovered cases coefficient.

7. Problem definition and optimization of parameters of the COVID19 spread model

Considering the large dimensions of the matrix describing the population (698x1202) and the hardly definable start values of the decision variable vector (an iterative model and the random nature of infections), the following steps were taken to define optimum values. In the first approximation, a matrix of population with smaller dimensions was assumed. In that initially limited population, optimization was performed for the periods between March 3 and October 8, 2020 (220 days) and between October 9 and November 24, 2020. The values of actual infection data were reduced in proportion to the population contained in the matrix with reduced dimensions. Assuming that the nature of population matrix described by the population density per 1km² only slightly differs from the entire matrix of population, the optimum values for the limited population matrix were determined as the initial values used in the proper optimization procedure.

The optimum values were determined for the entire matrix of population in the following steps, using the starting points. The procedure led to determination of values for 6 decision variables of the

contagion function and one parameter of the matrix of covariance. Thus, seven decision variables were defined. The procedure for optimizing the parameters of the contagion function was programmed to correspond to the sequence of actual events. The contagion process was initiated in Zielona Góra (Lubuskie province), on the first day, March 4, 2020 ⁹. The simulation model with the optimization and visualization module was developed in MATLAB by the author.

In the optimization procedure, a criterion function (Eq.11) was adopted defined as a minimum of the total sum of squares of the difference between actual values of infection cases and computational values. A vector z was sought that minimizes the value of the adopted objective function.

$$\hat{z} \rightarrow \min \sum_{i=1}^{267} (F_i - S_i(z))^2 \quad (11)$$

z – the decision variable vector, optimized values,

\hat{z} – the vector of optimum values of decision variables,

267 – the number of days between March 3, 2020 and November 24, 2020,

$i = 1, \dots, 267$ – subsequent days in the analysed time interval,

F_i – the number of registered cases of disease at individual moments,

$S_i(z)$ – the number of disease cases resulting from the model.

The results of optimization of the model parameters and the values of criterion function are given in Table 1.

Decision variable	Values
z_1	6.1690e-09
z_2	-1.7176e-06
z_3	1.3144e-04
z_4	3.5957e-09
z_5	-2.5232e-06
z_6	5.0576e-04
z_7	40000
Criterion value	3.5475e+10

Table 1. Optimum values of decision variables

8. Verification of the model parameters and precision of simulation for the entire territory of Poland

The goodness of fit was determined by the value of slope (Eq.12) that in the case of complete conformity of the results of simulation with the actual data should assume the value of 1 (Table 2). Also the extreme values for the 95% confidence interval and the value of the coefficient of determination are given.

$$y = \hat{a}x \quad (12)$$

y – values calculated in the model

x – actual values

Cases	Slope value \hat{a}	Limit values of slope for 95% confidence interval	Coefficient of determination R^2
Infected	0.981	(0.975, 0.987)	0.996
Fatal cases	1.062	(1.041, 1.083)	0.957
Recovered	1.193	(1.162, 1.224)	0.924

Table 2. Verification of the contagion model on a national scale during 267 days of infection between March 3, 2020 and November 24, 2020

9. Verification of calculation results on a regional scale

The simulation results were verified at a regional level separately for each province. The administrative criterion was adopted due to the nature of published information about the pandemic in Poland. Table 3 contains the values of slope (Eq.12) describing conformity of computational results with the actual data and the limits of confidence interval at the 95% level. Additionally, the coefficient of determination is given and simplified demographic data from 16 provinces with the population size and density per 1km² ³⁶.

	Province	Slope value \hat{a}	Limit values of slope for 95% confidence interval	Coefficient of determination R^2	Population total / density per 1km ² as of December 31, 2019 ³⁶
1	Lubuskie	0.488	(0.475, 0.502)	0.920	1011.6/72
2	Wielkopolskie	0.786	(0.776, 0.795)	0.987	3498.7/117
3	Dolnośląskie	0.825	(0.813, 0.837)	0.980	2900.2/145
4	Kujawsko-Pomorskie	0.953	(0.934, 0.972)	0.964	2072.4/115
5	Lubelskie	0.712	(0.700, 0.724)	0.975	2108.3/84
6	Łódzkie	1.245	(1.235, 1.256)	0.973	2454.8/135
7	Małopolskie	1.007	(0.997, 1.017)	0.991	3410.9/225
8	Mazowieckie	1.278	(1.267, 1.288)	0.991	5423.2/153
9	Opolskie	0.953	(0.947, 0.959)	0.997	982.6/104
10	Podkarpackie	0.720	(0.709, 0.732)	0.9450	2127.2/119
11	Podlaskie	0.509	(0.502, 0.515)	0.986	1178.4/58
12	Pomorskie	0.683	(0.672, 0.694)	0.976	2343.9/128
13	Śląskie	1.034	(1.006, 1.062)	0.937	4517.6/366
14	Świętokrzyskie	1.339	(1.330, 1.348)	0.996	1234.0/105
15	Warmińsko-Mazurskie	0.906	(0.884, 0.927)	0.949	1422.7/59
16	Zachodniopomorskie	0.357	(0.346, 0.367)	0.910	1696.2/74

Table 3. Verification of the contagion model on a national scale

10. An analysis of simulation results

The structure of agent-based and compartmental model gives access to the results of calculation of all adopted cell statuses (labels) in unit areas, at all time intervals. Due to the high resolution (the number of population elements) of information about the population density, visualization of each analysed status is simple and immediate. The results of various computational cases are spatially depicted below.

Fig. 3 shows the increase in infection of cells with COVID-19 after 267 days, i.e. between March 4, 2020 and November 24, 2020 in a verified model with optimum parameters. An increased density of infections is clearly seen in large urban centres: Silesia, Gdańsk, Kraków, Warsaw, Wrocław and Poznań. Contagion is particularly dense in Małopolska, Śląsk and Mazowsze. The nature of virus spread is confirmed by actual cases. Actual data and calculations as of November 24, 2020 show the number of infected people: 937 thousand, the number of fatal cases: 14 thousand, the number of recoveries: 454 thousand.

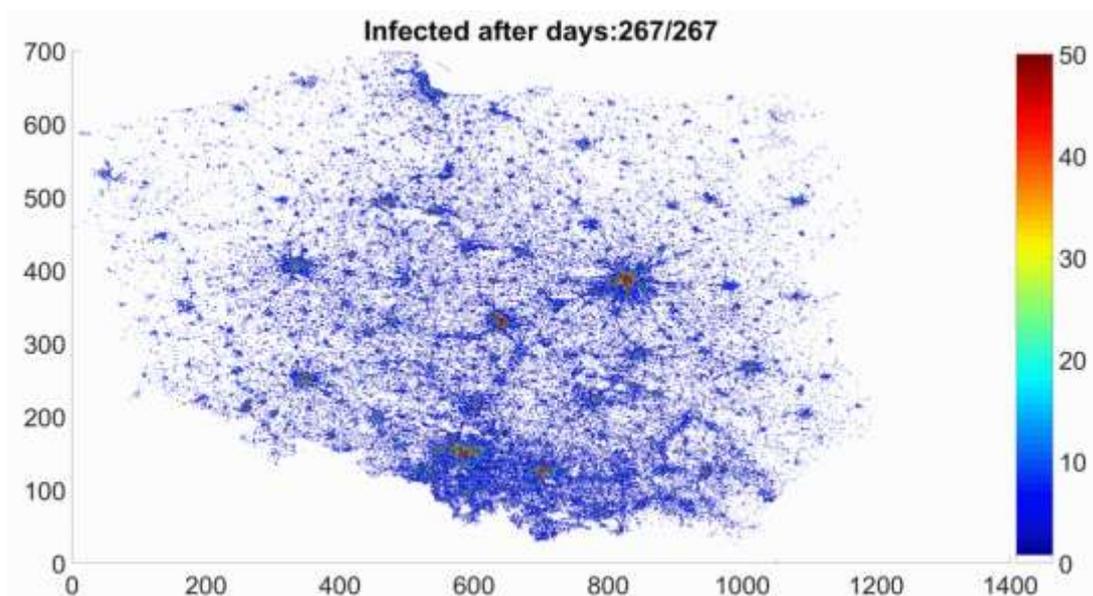


Figure 3. A spatial picture of COVID19 contagion, optimal parameters, after 267 days (November 2020)

In subsequent analysed cases, it was assumed in simulations that the parameters of the contagion function will not change after day 267 (November 24, 2020). At the turn of November and December, the Polish government announced that vaccination will begin in February 2021. The question arises about the picture of infections in the population of Poland in early March 2021, i.e. after one year. Fig. 4 shows the results of calculation of the infection status after a year, as of March 2, 2021. Due to the dramatic increase in contagion beginning on October 8, 2020 (the bifurcation point), the entire country population would be almost uniformly infected in the period of 99 subsequent days, with concentrations in large urban centres. Actual data and calculations as of March 3, 2021 show the number of infected: 5,252 thousand, the number of fatal cases: 164 thousand, the number of recoveries: 5,215 thousand.

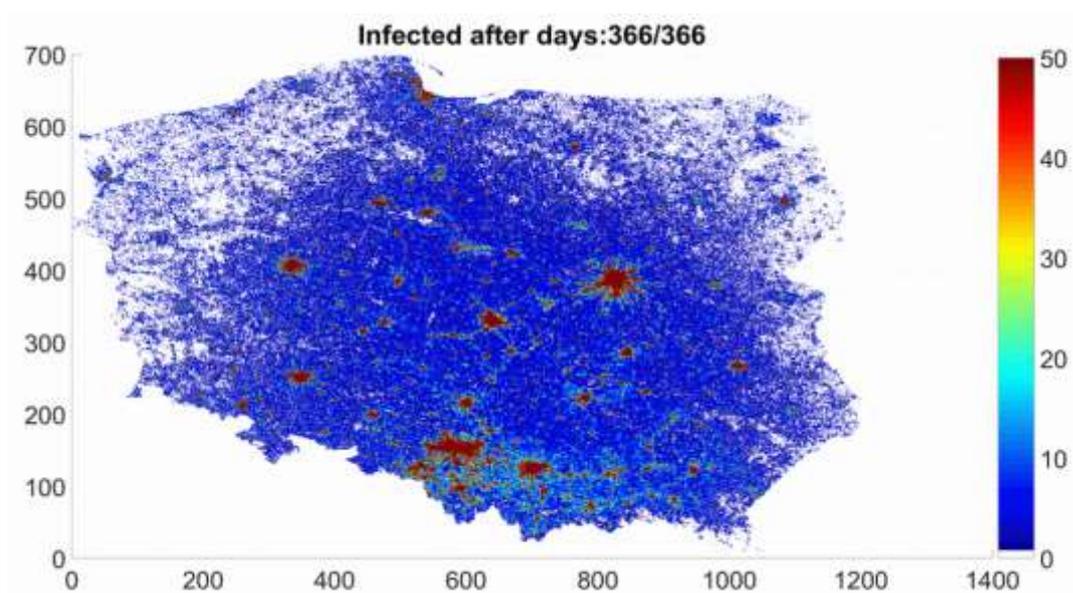


Figure 4. COVID19 infections, optimal parameters, after 366 days (forecast)

11. A model of vaccination and adopted vaccination strategies

The model of vaccination (Eq.13) like the model of fatal cases and recoveries is built assuming a constant factor of proportionality. A constant, independent of time coefficient of vaccines given, defining the number of vaccinated cells relative to the number of healthy cells is adopted. In its initial assumptions, the Polish government planned to vaccinate 1 million people monthly. On average, 33 thousand people can be vaccinated daily. The value of vaccination coefficient is determined to obtain a total number of vaccinated people similar to that planned.

$$V_{ij,t+1} = V_{ij,t} + H_{ij,t} * COEFF_{VACC} \quad (13)$$

$V_{ij,t+1}$ – vaccinated cases in element i, j at time $t + 1$,

$V_{ij,t}$ - vaccinated cases in element i, j at time t ,

$H_{ij,t}$ - healthy cases in element i, j at time t ,

$COEFF_{VACC}$ – vaccination coefficient.

Three vaccination strategies were defined. The value of coefficient equals $5e-4$ in Strategy I, $1e-3$ in Strategy II and $2e-3$ in Strategy III. Table no. 4 contains the results of model simulation, considering adopted vaccination policies.

Coefficient of vaccination		Strategy I	Strategy II	Strategy III
Data of maximum of infected		08/12/2021	12/04/2021	03/08/2022
Maximum infection	Cases	5 779	4 358	2 697
Population structure after 1 year (March 3, 2021) from the time of outbreak of the pandemic on March 3, 2020; beginning of the vaccination process on December 27, 2020				
Total of deaths	cases	139	122	101
Total infected	cases	3 958	2 875	1 718
Total recovered	cases	4 420	3 873	3 194
Total vaccinated	cases	1 070	2 152	4 282
Vaccination intensity	cases/day	16.976	32.124	63.905
Population status after 3 years (March 3, 2023) from the time of outbreak of the pandemic on March 3, 2020; beginning of the vaccination process on December 27, 2020				
Total of deaths	cases	1 044	887	580
Total infected	cases	46	66	345
Total recovered	cases	33 153	28 192	15 486
Total vaccinated	cases	3 467	8 442	21 480
Vaccination intensity	cases/day	4.341	10.568	26.886

Table 4. Population structure in various vaccination strategies (in thousands) – forecasting

An analysis of the calculation results obtained for the 3 adopted vaccination strategies (Table 4, Fig. 5) demonstrates how the structure of Poland's population changes in a 3-year pandemic time horizon. It also shows the scope of involvement in the vaccination process. Strategy I is the worst-case scenario for the population, considering such criteria as the number of deaths and infected people. Strategy I requires minimum involvement of health service personnel and resources and thus represents the cheapest solution.

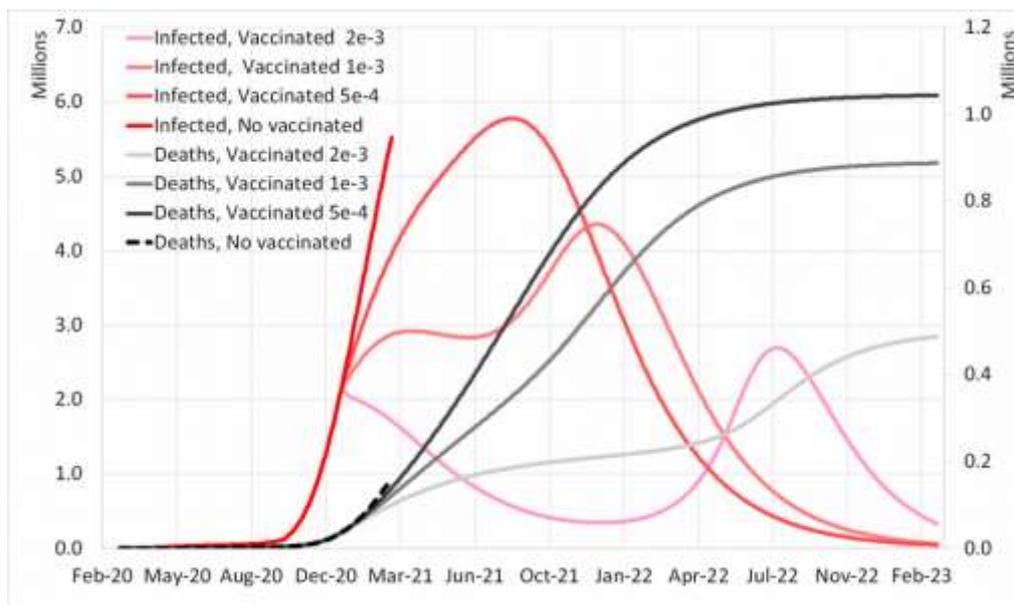


Figure 5. Numbers of infected cells and the number of fatal cases depending on vaccination strategy over 3 years (forecast); vaccination initiated on December 27, 2020

Strategy		Strategy I	Strategy II	Strategy III
Data of maximum of infected		05/05/2021	05/07/2021	05/14/2021
Maximum infection	%		24.6%	53.3%
Population structure after 1 year (March 3, 2021) from the time of outbreak of the pandemic on March 3, 2020; beginning of the vaccination process on December 27, 2020 – a forecast				
Total of deaths	%		12.2%	27.3%
Total infected	%		27.4%	56.6%
Total recovered	%		12.4%	27.7%
Population status after 3 years (March 3, 2023) from the time of outbreak of the pandemic on March 3, 2020; beginning of the vaccination process on December 27, 2020 – a forecast				
Total of deaths	%		15.0%	44.4%

Table 5. A comparison of the strategies considering the adopted statuses / criteria (various vaccination policies)

An analysis of the results of strategies (the value of vaccination coefficient) from Tables 4, 5 demonstrates that an increase in the number of vaccinations by a factor of 4 ($2e-3/5e-4$) between December 27, 2020 and March 3, 2021 entails a decrease in the number of fatal cases by 8.1% and by 21.6% compared to Strategy no. I. The number of infected people falls by 15.4% and 39.8% compared to Strategy no. I. In a 3-year time horizon, the number of fatal cases is reduced by 13.6% and by 31.5% compared to Strategy no. I. The number of infected people is reduced by 16.0% and 40.8% compared to Strategy no. I.

12. Conclusion

The above description of the COVID-19 virus spread is obtained using an iterative agent-based and compartmental model with the population density as its key parameter. The contagion process is expanded by a random component based on Monte Carlo methods. Seven decision variables (model

parameters) were defined, and their values calculated in an optimization procedure. The model parameters were optimized for data from the period between March 3, 2020 and November 24, 2020. The model was assessed and verified using data about contagion in the entire country and at a regional level of 16 provinces. The calculations and forecast analyses were done entirely for the territory of Poland, considering 3 vaccination strategies. The model represents an expanded classical SIR epidemic model. It may be described using the acronym SIR-V (Susceptible, Infected, Recovered - Vaccinated). The model is dynamic and does not include time delays. The model incorporates the properties of real world such as population distribution and describes a two-dimensional network of unit areas wherein cells by their mobility characterized in a random manner are capable of infecting cells in other unit areas. The proposed model is based on transmission of the disease described by a contagion function. The structure of the model is capable of grasping the effects of real measures, such as restrictions imposed or lifted bans, by defining a suitable form of the contagion function. The model assumes mobility, but a cell returns to its original place of stay within the same time unit. Consequently, only propagation of the disease takes place without changes in the density of cells per area unit.

The adopted contagion function is independent of the coordinates of element in the population matrix. Even with this simplification, the results of modelling are confirmed both on a national and regional scale. The results of model verification on a national scale demonstrate that computational results are consistent with real data at a level of coefficient of determination R^2 between 0.924 and 0.996. The value of coefficient of determination for regional data is contained between 0.910 (zachodniopomorskie) and 0.995 (opolskie).

The iterative model with the proposed, unique for the territory of Poland, forms of contagion, recovery and mortality functions, based on population density data, represents a flexible approach that can be used in any country or territory.

The model is limited in that it ignores time delays in the process of infection, recovery, and acquisition of immunity due to contact with the virus or vaccination. If the relationship between delay times and the simulation period is negligible, the delays have no material effect on the model usefulness.

The structure of the model can be expanded to incorporate delays characteristic of various model components. Conformity of computational results with the real data on a regional scale confirms the material effect of the population density on the virus spread. A change in the form of criterion function from the national into the regional level (with the normalization procedure applied) can improve conformity of calculation results with reality. Another step includes transition from a "national" contagion function to "regional" contagion functions with contagion process expansion by regionally defined random processes. These notes also indicate the directions of future development work.

References

1. Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020).
2. J., Z. *et al.* Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *Lancet Infect. Dis.* **20**, 793–802 (2020).
3. Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
4. Wu, H. liang liang, Huang, J., Zhang, C. J. P. J. P., He, Z. & Ming, W. K. K. Facemask shortage and the novel coronavirus disease (COVID-19) outbreak: Reflections on public health measures. *EClinicalMedicine* **21**, (2020).
5. Pourhomayoun, M., Shakibi, M., Angeles, L., Shakibi, M. & Angeles, L. Predicting Mortality Risk in Patients with COVID-19 Using Artificial Intelligence to Help Medical Decision-Making. **19**,

- (2020).
6. OMS. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. *WHO Dir. Gen. speeches* 4 (2020).
 7. Wong, G. *et al.* MERS, SARS, and Ebola: The Role of Super-Spreaders in Infectious Disease. *Cell Host Microbe* **18**, 398–401 (2015).
 8. Azhar, E. I. *et al.* Evidence for Camel-to-Human Transmission of MERS Coronavirus. *N. Engl. J. Med.* **370**, 2499–2505 (2014).
 9. Duszyński, J. *et al.* Zrozumieć COVID-19 OPRACOWANIE ZESPOŁU DS. COVID-19 PRZY PREZESIE POLSKIEJ AKADEMII NAUK. https://informacje.pan.pl/images/2020/opracowanie-covid19-14-09-2020/ZrozumiecCovid19_opracowanie_PAN.pdf (2020).
 10. Fehr, A. R. & Perlman, S. Coronaviruses: An Overview of Their Replication and Pathogenesis. in *Coronaviruses: Methods and Protocols* (eds. Maier, H. J., Bickerton, E. & Britton, P.) 1–23 (Springer New York, 2015). doi:10.1007/978-1-4939-2438-7_1.
 11. Car, Z., Baressi Šegota, S., Anđelić, N., Lorencin, I. & Mrzljak, V. Modeling the Spread of COVID-19 Infection Using a Multilayer Perceptron. *Comput. Math. Methods Med.* **2020**, (2020).
 12. Lazzari, G. *et al.* A digital reconstruction of the 1630–1631 large plague outbreak in Venice. *Sci. Rep.* **10**, 1–7 (2020).
 13. Spyrou, M. A. *et al.* Phylogeography of the second plague pandemic revealed through analysis of historical *Yersinia pestis* genomes. *Nat. Commun.* **10**, (2019).
 14. Cohen, M. S., Hellmann, N., Levy, J. A., Decock, K. & Lange, J. Review series The spread, treatment, and prevention of HIV-1: evolution of a global pandemic. *J. Clin. Invest.* **118**, (2008).
 15. Preiss, S., Garçon, N., Cunningham, A. L., Strugnell, R. & Friedland, L. R. Vaccine provision: Delivering sustained & widespread use. *Vaccine* **34**, 6665–6671 (2016).
 16. Goel, R. & Sharma, R. Mobility based SIR model for pandemics – with case study of COVID-19. *arXiv* (2020).
 17. Hasell, J. *et al.* A cross-country database of COVID-19 testing. *Sci. Data* **7**, 1–7 (2020).
 18. Total confirmed COVID-19 deaths. <https://ourworldindata.org/grapher/total-covid-deaths-region?tab=table&time=2020-01-11..latest>.
 19. pandemic | Description, Preparedness, & Historical Pandemics | Britannica. <https://www.britannica.com/science/pandemic>.
 20. Mollalo, A., Rivera, K. M. & Vahedi, B. Artificial-neural-network-modeling-of-novel-coronavirus-COVID19-incidence-rates-across-the-continental-United-States2020International-Journal-of-Environmental-Research-and-Public-HealthOpen-Access.pdf. (2020).
 21. Hunter, E., MacNamee, B. & Kelleher, J. A comparison of agent-based models and equation based models for infectious disease epidemiology. *CEUR Workshop Proc.* **2259**, 33–44 (2018).
 22. Bodine, E. N., Panoff, R. M., Voit, E. O. & Weisstein, A. E. Agent-Based Modeling and Simulation in Mathematics and Biology Education. *Bull. Math. Biol.* **82**, 1–19 (2020).
 23. Perez, L. & Dragicevic, S. An agent-based approach for modeling dynamics of contagious disease spread. *Int. J. Health Geogr.* **8**, 1–17 (2009).
 24. Khanna, A. S., Dimitrov, D. T. & Goodreau, S. M. What can mathematical models tell us about the relationship between circular migrations and hiv transmission dynamics? *Math. Biosci. Eng.* **11**, 1065–1090 (2014).
 25. Sallah, K. *et al.* Mathematical models for predicting human mobility in the context of infectious disease spread: Introducing the impedance model. *Int. J. Health Geogr.* **16**, 1–11 (2017).
 26. Rocklöv, J. & Sjödin, H. High population densities catalyse the spread of COVID-19. *J. Travel Med.* **2020**, 1–2 (2020).
 27. Widyaningsih, P., Saputro, D. R. S. & Nugroho, A. W. Susceptible exposed infected recovery (SEIR) model with immigration: Equilibria points and its application. *AIP Conf. Proc.* **2014**, (2018).
 28. Petard, H. A Contribution to the Mathematical Theory of Big Game Hunting. *Am. Math. Mon.*

- 45**, 446 (1938).
29. Jin, Y., Wang, W. & Xiao, S. An SIRS model with a nonlinear incidence rate. *Chaos, Solitons & Fractals* **34**, 1482–1497 (2007).
 30. Chinazzi, M. *et al.* The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. <http://science.sciencemag.org/>.
 31. Eubank, S. *et al.* Modelling disease outbreaks in realistic urban social networks. *Nature* **429**, 180–184 (2004).
 32. Xie, G. OPEN A novel Monte Carlo simulation procedure for modelling COVID - 19 spread over time. *Sci. Rep.* 1–9 (2020) doi:10.1038/s41598-020-70091-1.
 33. CDC. Duration of Isolation Precautions for Adults with Duration of Isolation & Precautions for Adults. *Centers Dis. Control Prev.* **2019**, 1–7 (2020).
 34. Long, Q. X. *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* **26**, 1200–1204 (2020).
 35. School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Science Information Network (CI, C. U. (2018)). Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076). vol. 2507 1–9 <https://www.worldpop.org/geodata/summary?id=33283> (2020).
 36. Yearbook, D. Demographic Yearbook of Poland. (2020).
 37. Gov.pl. Serwis Rzeczypospolitej Polskiej. Koronawirus : Pliki archiwalne - dane historyczne Materiały. 2–3 (2020).
 38. Blumenthal, D., Fowler, E. J., Abrams, M. & Collins, S. R. *Sounding Board Covid-19- Implications for the Health Care System.* (2020).
 39. Aguiar, M., Ortuondo, E. M., Bidaurrezaga Van-Dierdonck, J., Mar, J. & Stollenwerk, N. Modelling COVID 19 in the Basque Country from introduction to control measure response. *Sci. Rep.* **10**, (2020).
 40. Shen, J. A recursive bifurcation model for early forecasting of COVID-19 virus spread in South Korea and Germany. *Sci. Rep.* **10**, 1–10 (2020).
 41. Trip modelling and travel forecasting. in *Trip modelling and travel forecasting* (ed. Żurowska, J.) vol. 1 (Research and technical papers of polish association for transportation engineers in Cracow, 2014).
 42. European Commission. *Statistical Pocketbook 2019: EU transport in figures.* (2019). doi:10.2832/017172.
 43. Barda, N. *et al.* Developing a COVID-19 mortality risk prediction model when individual-level data are not available. *Nat. Commun.* **11**, 1–9 (2020).

Figures

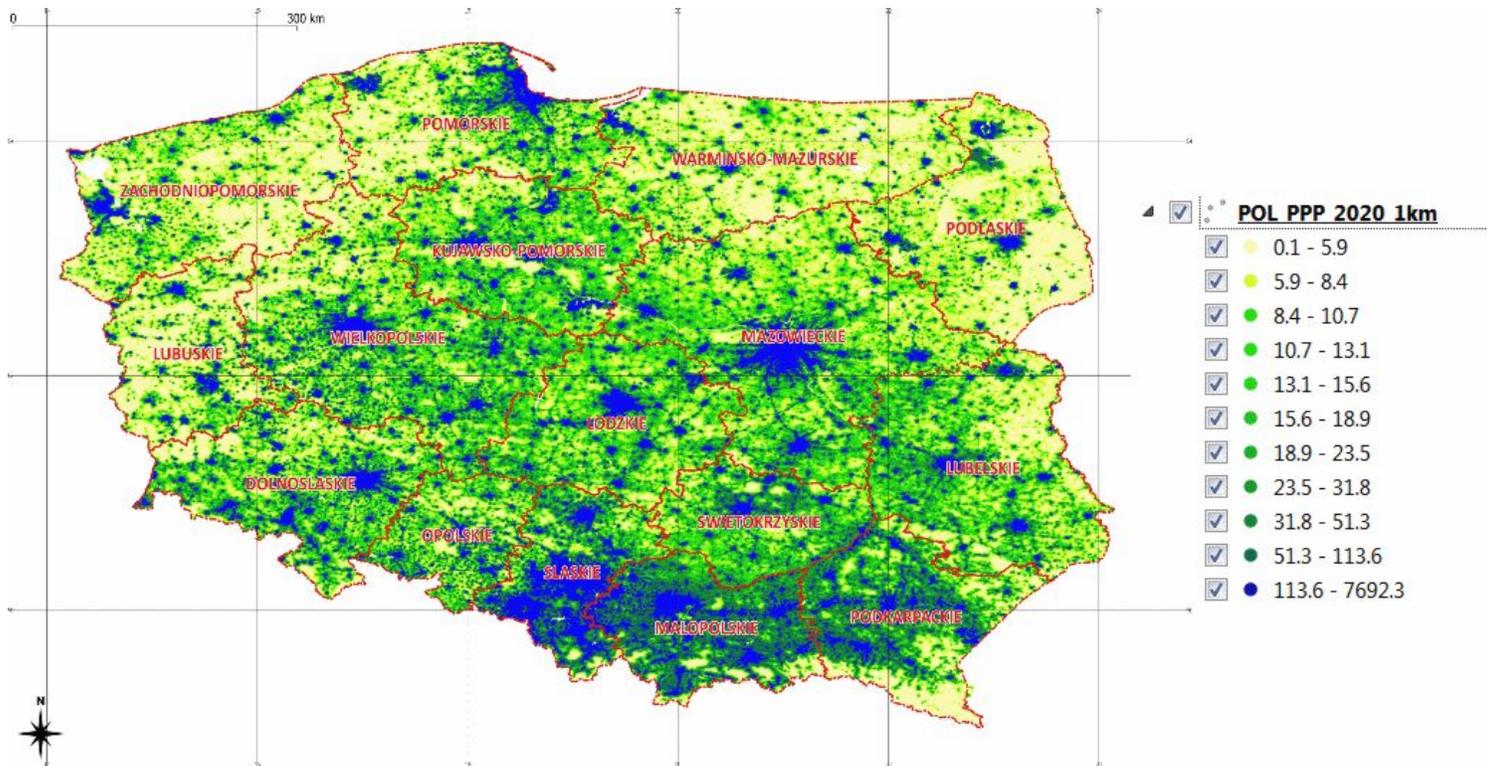


Figure 1

Population density per 1 km² Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

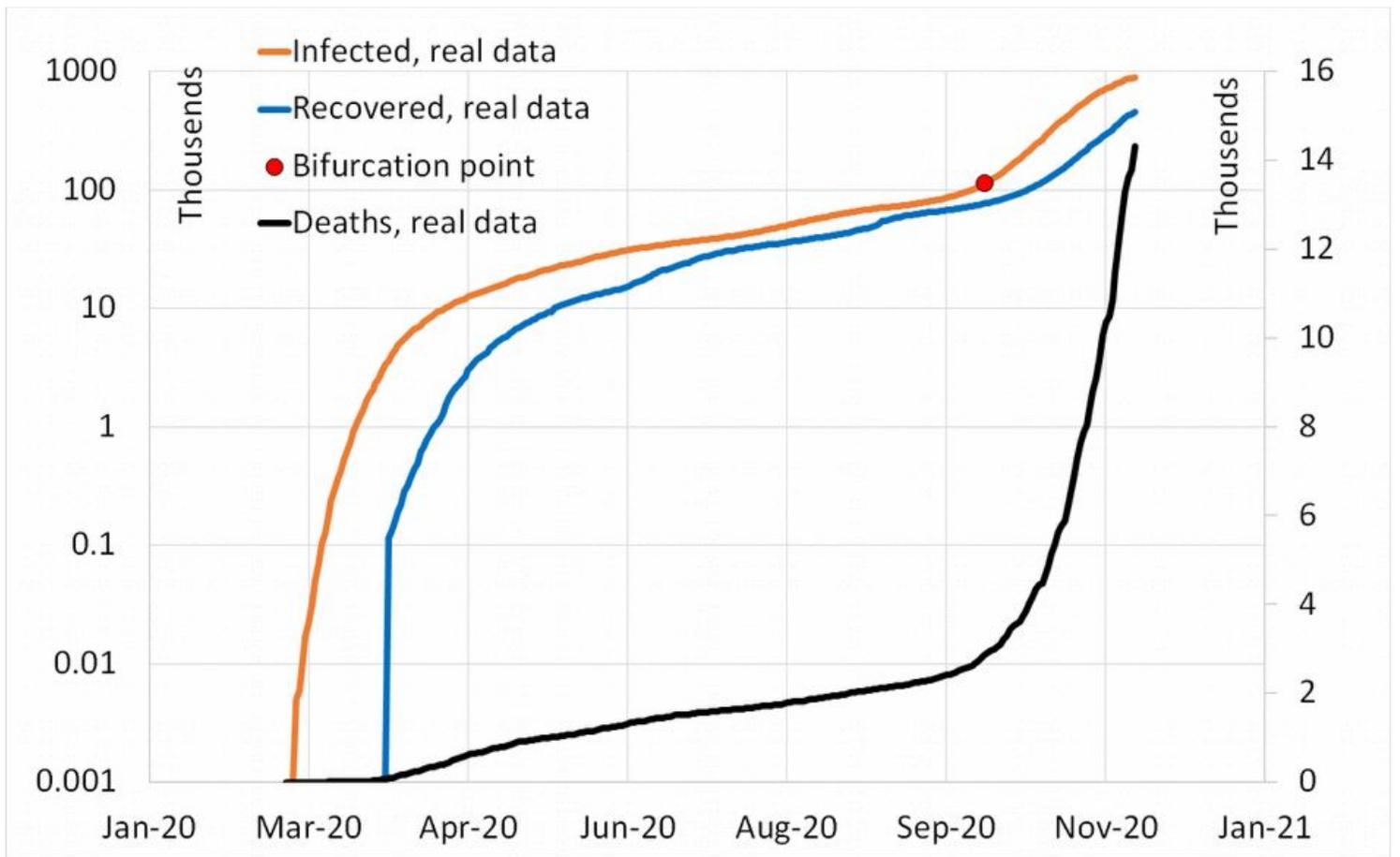


Figure 2

COVID-19 infections in the population of Poland between March 3, 2020 and November 24, 2020

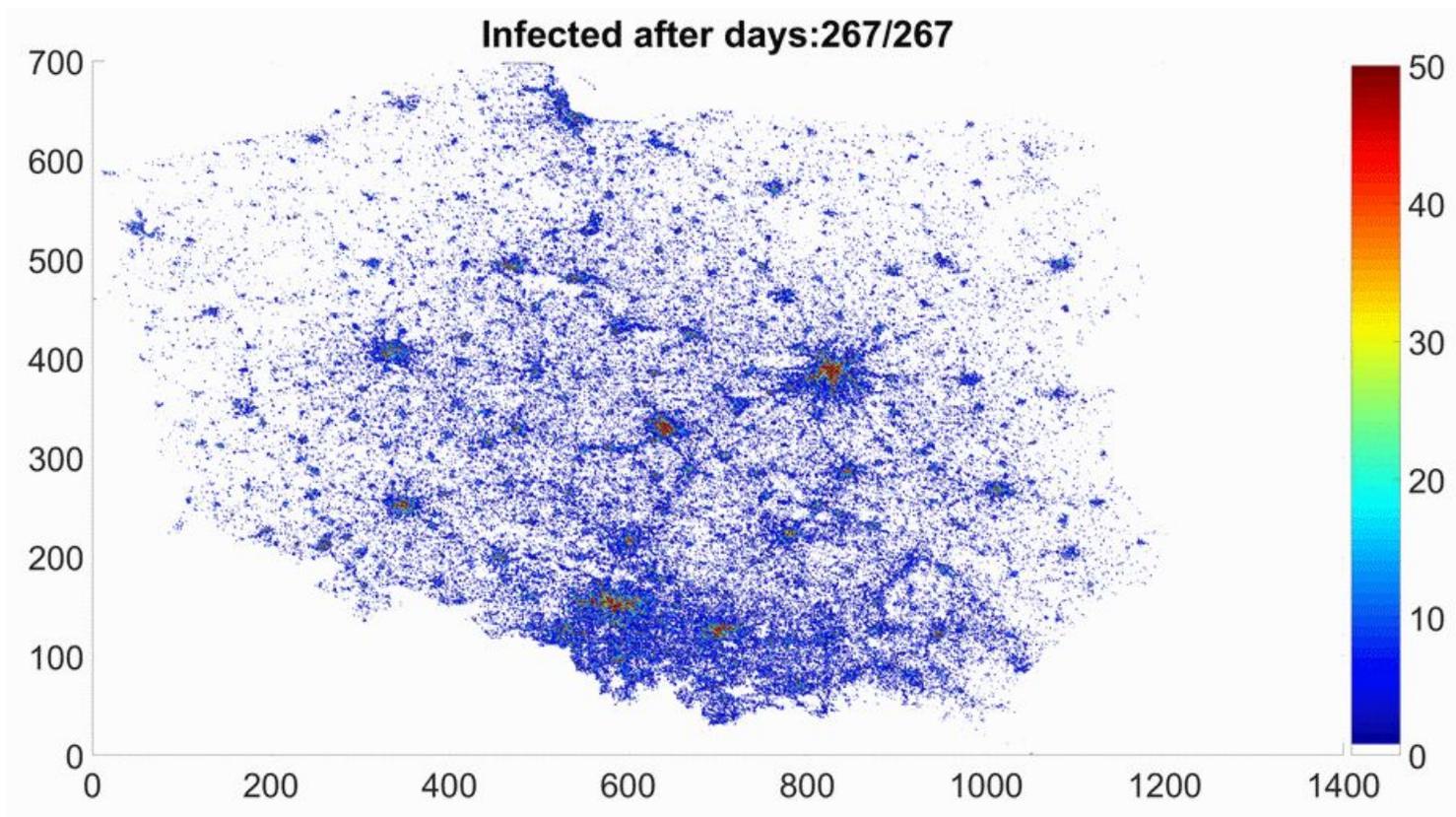


Figure 3

A spatial picture of COVID19 contagion, optimal parameters, after 267 days (November 2020) Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

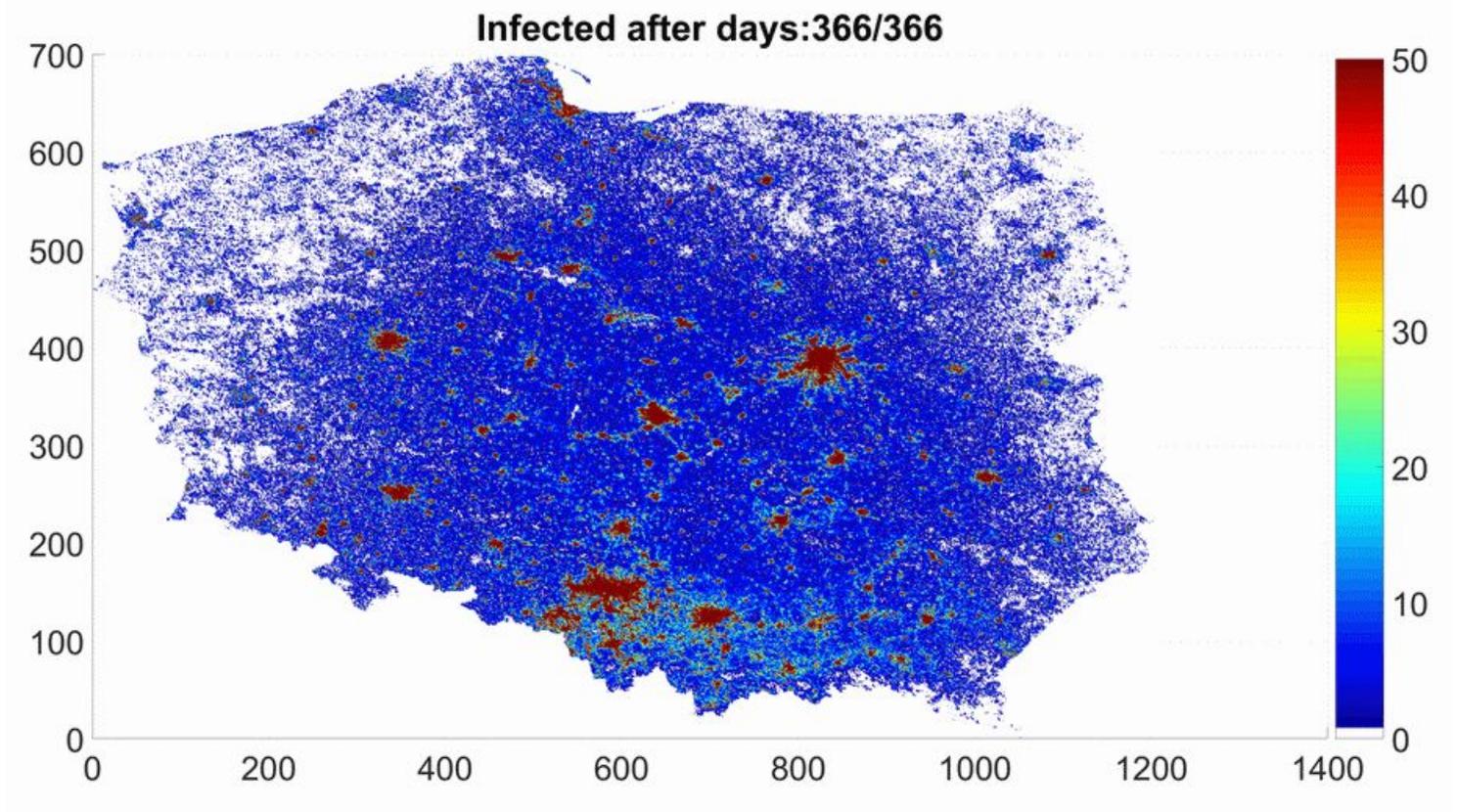


Figure 4

COVID19 infections, optimal parameters, after 366 days (forecast) Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

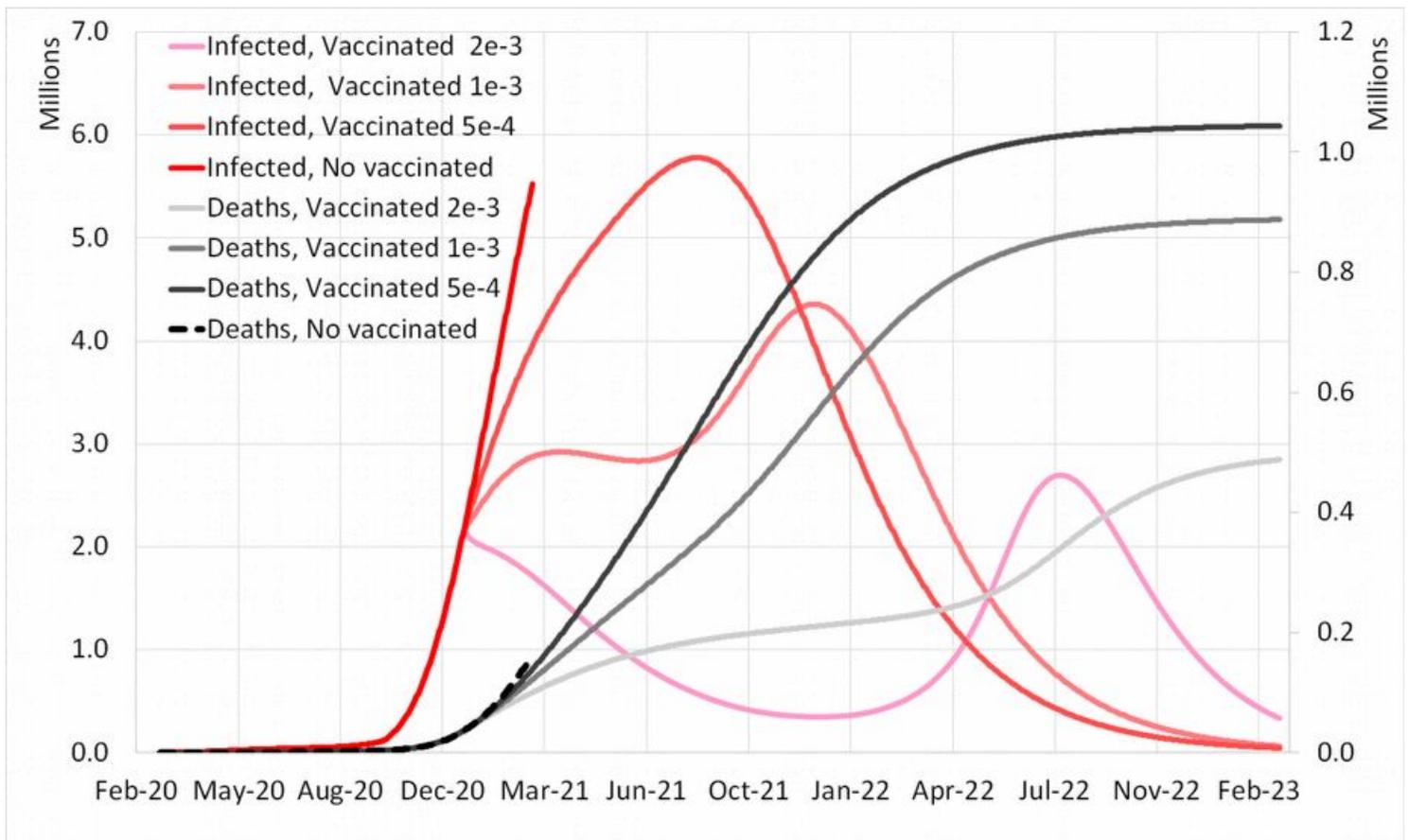


Figure 5

Numbers of infected cells and the number of fatal cases depending on vaccination strategy over 3 years (forecast); vaccination initiated on December 27, 2020