Tracing and Modeling of the COVID-19 Pandemic Infections in Poland Using Spatial Interactions Models

Piotr A. Werner(✉)

Faculty of Geography and Regional Studies, University of Warsaw, Krakowskie.Przedmiescie 30, 00-927 Warsaw, Poland
peter@uw.edu.pl

Abstract. The nexus of factors influencing the dissemination of the SARS-CoV-2 virus is so complex that identification of (some) determining factors of COVID-19 spatial diffusion is significantly hampered. COVID-19 characterize of specific dynamics and enormous volume of morbidity. The aim of the study is construction of the model of spatial dissemination of COVID-19 in Poland, identification of the main outbreak places and spatial heterogeneity of pandemic based on the spatial set of first twelve months morbidity data (in 2020 and 2021). The target (prototypical) model is intended rather as the supporting tool than replacement of the well-known and used SIR or SEIR (Susceptible – Exposed – Infected - Recovered) models in epidemiology. It also assumed that the target model could be used as a priori estimation tool of the spatial locations of infections outbreaks as well as evaluation of future volume of infections due to changing numbers of exposed and recovered persons related also to, newly, introduced and continuation coronavirus (COVID-19) vaccinations. One of the expected advantages of the construed model is its spatial aspect i.e. it will enable to evaluate the potential spatial differentiation of infected number of people within the set of observed spatial units i.e. counties in Poland.

Keywords: COVID-19 · Infections · Spatial interactions · Modeling · Poland

1 Introduction

Situation related to spread of pandemic, coronavirus SARS-COV-2, has been a global challenge, touching virtually every aspect of social life in every country of the world. Looking for the solutions of the current state aimed limiting spatial extent of pandemic and its medical, healthcare, social, economic and demographic consequences in social and individual scale, engaged numerous of specialists, based also on past plagues lessons [1] and exhaustive scientific research virtually in every domain of human life. "The spread of infectious disease is inherently a spatial process; therefore, geospatial data, technologies, and analytical methods play a critical role in understanding and responding to the coronavirus disease 2019 (COVID-19) pandemic" [2]. Policy makers, scientists and journalists agree that one of the key issues of effective combat against global pandemic

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12954, pp. 641–657, 2021.
https://doi.org/10.1007/978-3-030-86979-3_45

is tracing statistics, i.e., observations of the number, intensity, and spatial variation of disease status (infections). Construed spatial global, national and regional models of pandemic aim at [3]: (*) discussing a GIS technology for simulating disease outbreaks, (**) examining the relationship between numerous and different potential key variables (environmental, socio-economic, demographic etc.), (***) identifying and monitoring the spatial and temporal patterns and trends, and (****) recording and geocoding data (sometimes including symptoms, travel history, and dates of onset, admission and confirmation) to produce useful information able to support the emergency surveillance.

The nexus of factors influencing the dissemination of the SARS-CoV-2 pandemic is so complex that identification of (some) determining factors of COVID-19 spatial diffusion is significantly hampered. It shall be deemed, that COVID-19 characterize of specific dynamics and is far more transmissible [4]. Recent studies proved some specific characters of the diffusion processes in the early stages of virus outbreak among different areas (e.g. in Italy and in China) at the regional and international level [5], also correlated to individual regional physical conditions as well as social and demographic characteristics (e.g. in Spain) [6].

The aim of the study is construction of the model of spatial dissemination of COVID-19 in Poland, identification of the main outbreak places and spatial heterogeneity of pandemic based on the spatial set of first year morbidity data in 2020 in Poland. The target (prototypical) model in its core is intended rather as the supporting tool than replacement of the well-known and used SIR or SEIR[1] models in epidemiology. It also assumed that the target model could be used as *a priori* estimation tool of the spatial locations of infections outbreaks as well as evaluation of future volume of infections due to changing numbers of exposed and recovered persons related also to, newly, introduced and continuation coronavirus (COVID-19) vaccinations. One of the expected advantages of the construed model is its spatial aspect i.e. it will enable to evaluate the potential spatial differentiation of infected number of people within the set of observed spatial units (counties in Poland).

1.1 COVID-19 Disease Description and Pandemic Tracking

Usually three categories groups of characteristics of risk are enumerated: individual risk, natural environment (climatic) risk, social and demographic risk as the factors shaping the volume, intensity and spatial dispersion of diseases (epidemic). Individual risk is related to personal human physical and psychological characteristics in context of possible infection. Spatial natural conditions (e.g. temperature and humidity) are recognized as primary factors of environmental risk. The last category is related, inter alia, to mobility of people and number as well as density of population. Especially for COVID-19 pandemic which has been initially perceived as similar to epidemics in the near past (SARS, MERS) and sometimes to seasonal influenza [7–9]. However, at present, the possible seasonality of COVID-19 occurrence is not yet explained [10]. On the other side recent national analysis of the relationship between density of population and the COVID-19 infection and mortality rates for 913 metropolitan counties in the United States suggests unrelation of density to the infection rate and negativel relation

[1] Susceptible – Exposed – Infected – Recovered.

to the mortality rate, possibly due to greater adherence to social distancing policies and practices in denser areas and better quality of health care [11].

COVID-19 pandemic, despite the similarities to above mentioned diseases and epidemics, features the global scale, enormous dynamics and higher infection rate. There were also enormous number of particular social, economic and political global, national and local factors, which, in context of very hard-organized efforts within the national healthcare system, introduced lockdown[2] of economy and limited mobility of people shaped the spatial spread of COVID-19 pandemic. These all above-mentioned risks and factors made very difficult precisely to reflect the main determinants of current spatial trends of COVID-19 infections. Moreover, " there is much to suggest that the relatively low mortality associated with COVID-19 is the result of the introduction of severe restrictions on social life and a far-reaching freeze on the economy in Poland at a very early stage of the epidemic of this disease." [10].

Due to enormous efforts of national and local authorities related to lock-down, multiplicity and diversity of factors influencing the spread of COVID-19 as well as the psychological conditions of individuals' behavior, the complex nature of spatial aspects COVID-19 spread is much more difficult to model simultaneously in different spatial locations than at the general, i.e. at national scale. The obtained spatial model will be only the approximation at each level (global, national, regional and local).

2 Pandemic COVID-19 in Poland in 2020

From the geographical point of view, there are several spatial factors influencing spread of COVID-19, which modify local transmission based on social contacts (through droplets transmission). These are natural conditions: climate, relief and hydrography, as well as settlement structure, density of population, human (social) development level (hygiene, awareness of epidemic), and life styles i.e. mobility of population [12]. Finally yet importantly, the key element is organization of national and local healthcare system. However, due to different social and psychological phenomena and natural disruptive events, some recognized, theoretical diffusion models of infectious diseases overworked within the framework of medical geography could differ from the empirical data because of their complex nature.

Source of number of registered infections, recoveries and deaths are three sets of data collected by Polish Ministry of Health. The first data set involve data from March to May, the second one from March to September and the third data set are continuous reports, which have been published daily since the last week of November. The cumulative numbers of infections (stacked) by counties in Poland from 11 week to 41 week 2020 are presented on chart and map (Fig. 1, Fig. 2). These figures provide the some spatial regularities of COVID-19 diffusion during the observed time. There are also visible general time fluctuations and spatial trends, which undoubtedly confirm the efforts of local and national authorities to introduce lock-down strategy and dissemination of DDM (distance, disinfection, masks) rules.

There are obvious outbreaks that spread epidemic into neighborhood as well as main, greatest cities touched (in some locations not the cities itself but surrounding suburban

[2] i.e. a state of isolation or restricted access instituted as a security measure.

areas). Concerns about the predicted (at the end of October) forthcoming 'second wave' of epidemic had been confirmed. During one, the 48th week 2020 there were greater number of infections than in the whole time before (presented on chart and map, Fig. 3 and Fig. 4). The largest number of infections concentrated in main populated cities and their surrounding counties. All capitals of voivodships as well as main (secondary) towns, capital of counties, characterize larger number of infections than peripherals.

3 Theoretical Foundations, Methods and Operationalization of Spatial Study in the Context of Pandemics

The literature of epidemic research present numerous methods and models concerning morbidity and spatial aspects of diseases. The specialists have looked for outbreaks, i.e. places of the first appearance as well as models of spatial dynamics and dissemination of diseases. The frequent approaches used cellular automata and multi-agent simulations ([multi] agent-based modeling), which appropriately reflect direct contacts of people as well as neural nets [13]. Approaches to spatial epidemiology involve some methods using GIS (geographic information systems) as well as analytical and mapping techniques like geocoding, distance estimation, assessment of residential mobility and commuting, spatial and space-time clustering and more sophisticated applications like Bayesian approaches to disease mapping, spatial regression models (including geographically weighted regression) using also Web-based mapping aimed dissemination of scientific studies results [14]. There are also non-spatial approaches (threads) of scientific studies of epidemics that aim "to understand the causes of a disease, then to predict its course, and finally to develop ways of controlling it, including comparisons of different possible approaches. The first step is obtaining and analyzing observed data." [15].

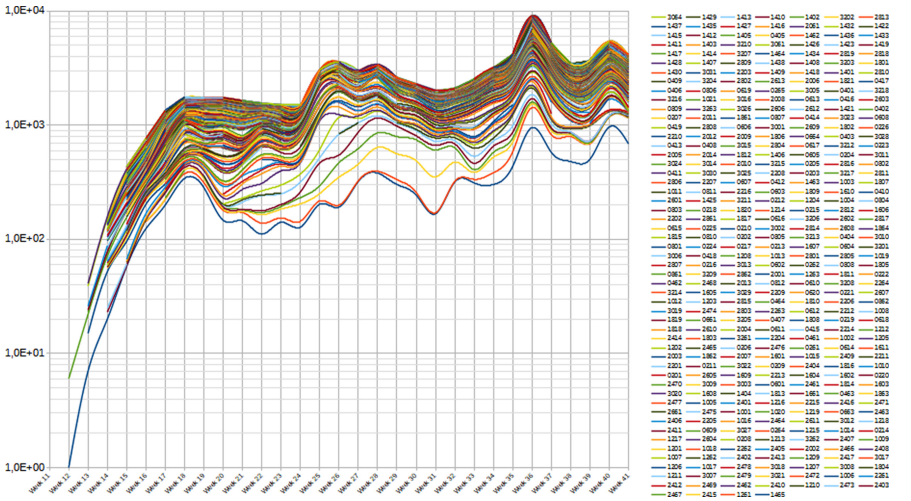


Fig. 1. Stacked, cumulative number of infections by counties in Poland by weeks (since 11 to 41 week 2020) and by counties; data source: Ministry of Health, Poland (counties identified by statistical numbers, four-digit TERYT string).

Researches concentrate on mathematical modelling for the spread of diseases using classic approaches, starting from Kermack - McKendrick epidemic (compartmental) model, and newer network models as well as the family of stochastic models, because the spread of infection is a random process [15–17].

It is used mainly for assess epidemic within geographic boundaries as entirety, treating it as the spatial homogenic unit without insight into its diversity. There must be appropriate number of different models construed for the set of different regions. The factors determining or influencing the spread of infection in different regions may differ and sometimes depend on specific set of complex conditions. There are set of different global and national models to be used to forecast the volume of infections [18–20].

The Interdisciplinary Centre for Mathematical and Computational Modelling UW (ICM UW) epidemiological model was used within the framework of proposed approach to trace the COVID-19 pandemic at the national level in Poland [20]. The results of simulation of ICM epidemiological model obtained at the national level were used, in turn, to assess the estimated number of infected population at the regional (county) level using spatial interaction models approach.

3.1 Spatial Interaction Models as a Tool for Tracing Epidemic at the Regional Level

Due to the defined above conditions of infections, which are the direct results of number of contacts of people, the proposed methodology use methods of spatial interactions modeling. A spatial interaction formula in a setting calibrated for diffusion can give the probability of contact from any location at any point in time, but this locus of greatest intensity of contact and change moves outwards from origins over time. Information requirements for the spatial interaction model calibrated for a diffusion problem include [18]: numbers and susceptibility of the population; spatial distribution of the population; and rate at which knowledge or opportunities to interact declines with distance.

Stewart [21] defined potential as a measure of the influence of population (mass - m) at some point on the Earth's surface within the distance d from this point. The term is derived from social physics, and the concept is closely connected to that of the gravity model, in that it relates mass (population) to distance. This formula defined the beginning classical approach to potential models [22]. Whereas the gravity model deals with separate relationships between pairs of points, population potential encompasses the influence of all other points on a particular one [23]. Corresponding gravity model as well as its variants became also predictors in infectious disease epidemiology [24]. A new look, based on statistical mechanics, was formulated by Wilson [25], who formulated and elucidated a family of spatial interaction models in the context of urban and regional systems, concerning, inter alia, demographic accounts and projections models, migration estimation and forecasting [26, 27]. Suitability of use a variant of spatial interaction model as a 'model of epidemics in spatially disaggregated system with interactions' has been directly confirmed by Sir Alan Wilson himself in the recently published preprint [28]. The new appeared abstract has reassured us that the research thread of using at least two of the spatial interactions models family is reasonable and may be the base of the prediction of spatial differentiation dynamics of COVID-19 spread.

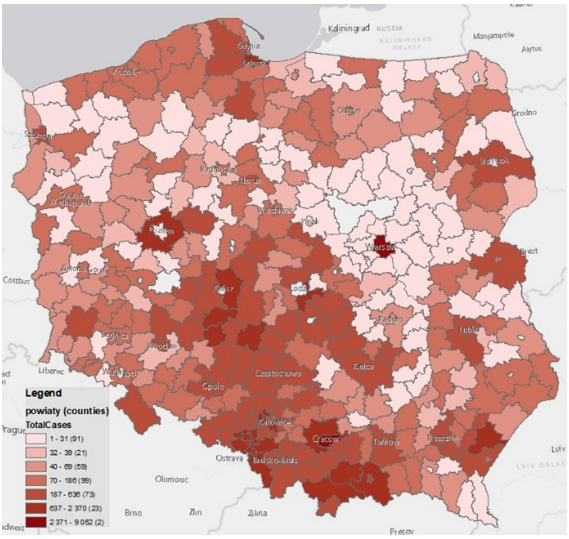


Fig. 2. Total number of infections by counties in Poland (since 11 to 42 week 2020); source Ministry of Health, Poland.

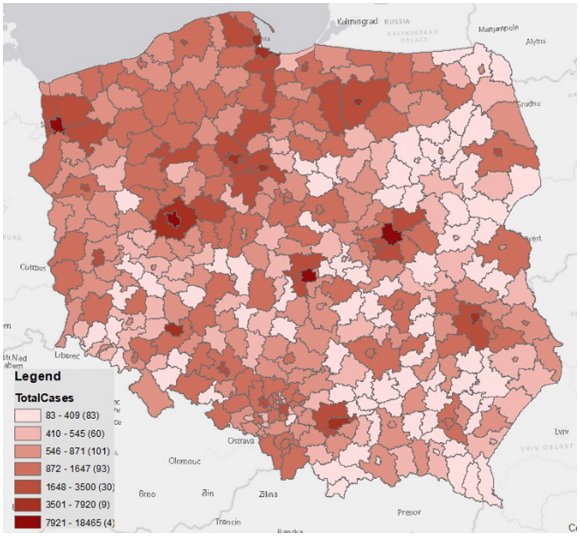


Fig. 3. Total number of infections by counties in Poland (since 48 to 52 week 2020); source Ministry of Health, Poland.

The roots of idea of use of population potential model as the measure of number of contacts between people are also derived from the earlier works related to spatial diffusion of cellular telephony [29–32] as well as taking into account 'first law of geography' [33]. Spatial interactions represent flows of activity between locations in geographical

space [34]. These may be flows of people, represented by population potential values. Population potential is a measure of nearness, or accessibility, of a certain aggregation of people to a given point. The family of spatial interactions models are well suited in the context of studies of direct contacts of people, which might result the infection of disease. In fact, the gravity model of two settlement points (characterized by appropriate number of people: *m* and *n* and omitting distance) is direct mathematical model of two separate groups of people welcoming themselves. Alan Wilson [35] introduced into the spatial interaction modelling the concept of probability of such contacts taking into account also the spatial dimension (distance or its functional measure, e.g. time) as well as the evaluation of the entropy of the studied system.

The starting point is an interaction matrix analogous to contingency table (graph notation) capturing the combinations of all pairs of distant sites (places) under consideration, and whose elements are the volumes of the interaction between the i^{th} and j^{th} source and destination points, and the origin *([O])* and destination *([D])* boundary vectors (rows and columns sums). The operationalization consisted in estimating two additional (edge) balancing vectors (A and B). This way the interaction matrix represent graph of all centroids (and edges connecting them) of counties in Poland. The main problem lies only how to treat cells on the diagonal i.e. when $i = j$, evaluation of the distances of contacts inside the county itself. Assumed finally, that the distance for contacts introduced into the model in these cases will be the radius of a circle of area equal to county area.

The general formula of population gravity between two certain points is described as $T_{ij} = f(O_i, D_j, d_{ij})$, *where: T – volume of interactions (contacts); O – origin point representing push factor (e.g. emissiveness); D – destination point representing pull factor (e.g. attraction); d – distance O-D representing spatial separation (e.g. Euclidean or road distance, time or cost function); i,j – table subscripts of points (regions).* The fifth element is matrix of distances between points (or regions, not presented in tab. 1). Summing up the individual gravity volumes for all other points or regions (including origin itself) gives the population potential for origin (or destination).

Operationalization of the above relationships using an entropy-maximizing framework seeks to reconstruct flows ([interactions, contacts]) between the set of origins and destinations by finding the most probable configurations. This is common optimization problem and including information about the total ([contacts]) inflows and outflows as well as parameters of modelling: distance decay function (d_{ij}^{β}, β – power exponent), origin *(A)* and destination *(B)* balancing factors ensuring the total (in- and out-flows) are preserved in the predicted flows. Final formulas for each of the model involving also independent variables influencing emissiveness *(V)* and attractiveness *(W)* take forms [36]:

$$\text{Unconstrained gravity model} : T_{ij} = V_i^{\mu} W_j^{\alpha} / d_{ij}^{\beta} \tag{1}$$

$$\text{Destination constrained model} : T_{ij} = B_j D_j V_j^{\mu} / d_{ij}^{\beta} \tag{2}$$

where:$B_j = 1/(\sum_i(V_i^{\mu})/(d_ij^{\beta}))$.

3.2 Operationalization in the Context of Pandemics

Looking for the real world independent variables, which influence the number of infected population, verified that the best results obtained using population number of counties, assuming that flows of spatial interactions models are represented by number of infected, which is the direct result of the number of social contacts. This way the population potential model is close to mathematical model counting the number of 'handshakes' taking into account distance measure. Two of above presented models seems adequate to apply in COVID-19 tracing. The unconstrained gravity model (1), when the only known is total number of infected people and the destination (attraction) constrained model (2), which make possible prediction of share of fixed (revealed) amount of infections at destinations that come from particular origins [26]. Two way of verification of forecast of number of infections within the counties is possible. Using collected data at national and county levels, and verifying the results using destination constrained spatial interaction model at the county level as well as unconstrained gravity model at the national level. The input in the beginning is the number of susceptible people. Later it will be possible to use updated numbers of susceptible subtracting the number of recovered (and vaccinated or planning vaccinations).

Having the collected data of number of infected and recovered (and failed) people it is possible to make the detailed calibrated local analysis of modeling spatial differentiation of infections at the county level using Python PySal SpInt library. The operationalization allows for any spatial resolution, temporal resolution, and demographic stratification. Unconstrained models are best purposed in an exploratory role, since the lack of balancing factors can result in predictions of more inflow or outflow than possible for a location [37].

4 Tracing the Pandemic COVID-19 in Poland

The record of tracing COVID-19 in Poland at the national level is presented in Fig. 4 and Fig. 6.

The data concerning infected population are only the fraction of the whole population infected. It is related to fact, that large number of infected people (especially youths) passes the disease asymptomatically and only the registered fraction of tested population is disclosed statistically. The vaccination against COVID-19 started in Poland since 27 Dec 2020. The complete collected data sets in Poland in 2020 can be treated as the state of outbreak of epidemic to the end of 2020. Only physicians and medical staff were vaccine since January 2021. The reproduction number [Rt] of outbreak over time shows differentiation (Fig. 5).

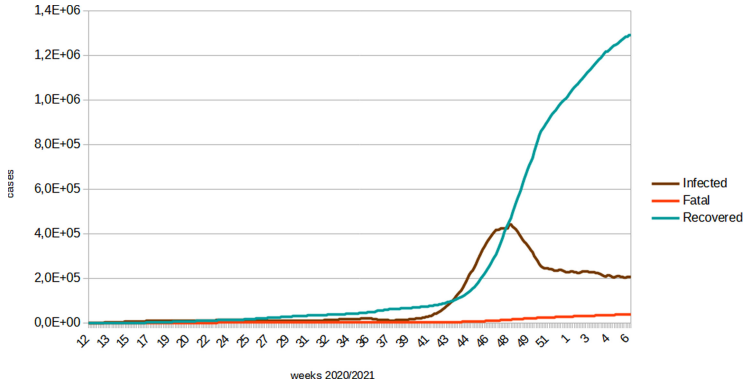


Fig. 4. Poland, COVID-19: Cases over time with monotonic increasing complemented recovered data (as Feb 07 2021, results of application: [18]).

4.1 Modeling and Simulation Infection by Counties

Preprocessing Phase

Spatial interaction models described above need calibration i.e. obtain accurate estimates of model's parameters in order to forecast interactions in system under study. As mentioned, two of them: unconstrained gravity (1) and destination (attraction) constrained (2) were chosen to simulate number of infections by counties in Poland. Preprocessing model calibration involved tracing and simulated past data of infections by counties for random as well as specific weeks during the last year and aimed confirmation that there is significant correlation between population of counties (treated as independent variable in context of possible contacts and explanation variables) and number of infections in these counties. Next, the set of parameters of spatial interactions models interpreted and compared to collected data modelling. Assumed that some of them describe the same processes and may explain similar relationships.

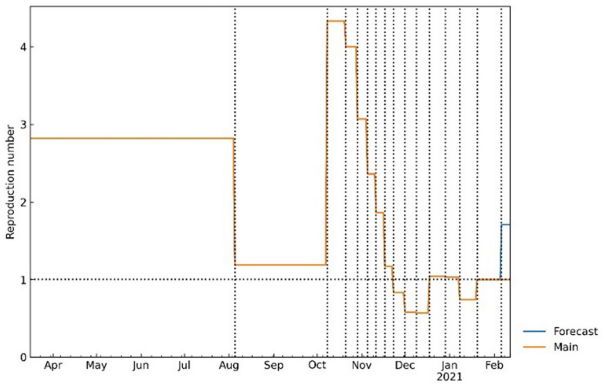


Fig. 5. Poland COVID-19: reproduction number over time (as Feb 13 2021, application forecast results [18]).

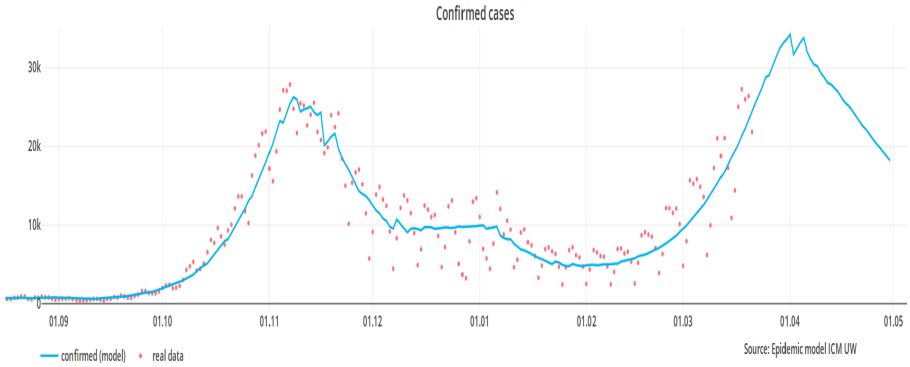


Fig. 6. Chart of identified cases of Covid-19 infection from 03/16/2020 with a prediction until 04/30/2021. Chart prepared based on the forecast from 03/19/2021 [20]

Spatial interaction models are calibrated via generalized linear modeling framework using flexibility and extendibility of regression framework, taking logarithm of both sides and obtaining Poisson log-linear model [36]. The input for model calibration were officially registered past data of infections by counties [38]. The regression made possible later to assess the regression model parameters: intercept (k, ensure the total number of infections is conserved for certain time), balancing factors (μ, origin propulsiveness or emissiveness parameter and α, destination attractiveness parameter) as well as exponential distance-decay parameter β, coefficient for the distance variable describing the diminishing effect of space [36, 37].

Modeling Spatial Differentiation of COVID-19 Infections

Spatial interaction models fit statistics utilize replacement of coefficient of determination (pseudo R^2) based on the likelihood function and its counterpart, adjusted coefficient of determination (adjusted pseudo R^2) which can be interpreted similarly – maximum value close to 1 denote better model fit [36]. Model complexity is assessed using Akaike information criterion (AIC), derived for information theory. Lower values of AIC signalizes better model fit as evaluation of the volume of information lost [39]. Statistical parameter usually utilized is standardized root mean square error (SRMSE) of the observed and simulated volume of interactions (in these case infections). Higher value of SRMSE is the indicator of decrease model assessment (starting from zero as the best fit) [36].

The procedure of spatial modeling simulation of infections by counties in Poland was repeated for above chosen unconstrained gravity and destination (attraction) models using the population statistics as the explanatory variables for the sequence of randomly, and specific chosen weeks in 2020. Input values were empirical statistical data of population and collected registered data of infections aggregated weekly by counties. The results of fit statistics for each run (week) are summarized in table (Table 1) and small charts in last row of this table (sparklines).

Verification of Destination Constrained Spatial Interaction Model

Verification of destination constrained spatial interaction model was evaluated during sixth and seven week of 2021 (1 Feb – 7 Feb) using acquired empirical aggregated data of COVID-19 infections by counties.

Table 1. Spatial interactions models fit statistics (* spatial verification; ** spatial prediction – later completed and verified).

model	Unconstrained gravity model				Destination (attraction) constrained model			
week No 2020/ 2021	pseudo R^2	adj. pseudo R^2	AIC	SRMSE	pseudo R^2	adj. pseudo R^2	AIC	SRMSE
15 [2020]	0.21	0. 208	557871.6	2.69	0.74	0.74	182944.1	1.067
20	0.20	0.197	1.72e+06	2.14	0.82	0.82	3.89e+05	0.74
30	0.16	0.164	3.15e+06	3.66	0.88	0.88	4.44e+05	1.28
40	0.50	0.50	3.41e+06	2.08	0.89	0.89	7.31e+05	1.28
50	0.78	0.78	205945.5	26.42	0.85	0.848	140473.2	20.68
1 [2021]	0.77	0.77	178070.68	28.05	0.85	0.85	118350.50	22.97
5	0.78	0.776	107080.50	26.92	0.85	0.847	73356.91	21.15
6*	0.78	0.778	103959.48	27.32	0.85	0.847	71851.31	21.05
7**	0.77	0,774	112247.09	28.04	0.85	0.846	76803.60	21.47
Chart /sparklines/								

The general formula of destination-constrained model (2) took form for the sixth and seventh week of 2021 (Eq. 3, 4):

$$T_{ij} = exp(-9.6409 + 0.6234ln(V_i) - 0.1243 - (-0.1102)\ln(d_{ij})) \quad (3)$$

$$T_{ij} = exp(-9.7187 + 0.6363ln(V_i) - 0.1220 - (-0.1101)\ln(d_{ij})) \quad (4)$$

where: T_{ij}– modelled number of infections, V_i– population of county i, d_{ij}– distance between county i and j; for the county itself the distance taken into account was the radius of a circle of area equal to county area.

The calculated adjusted coefficient of determination *(R^2)* was equal *0.85* (may be interpreted as 85% of real number of infections). Total number of infected is the sum of all interactions between the each pair of counties. Evaluated modeled spatial differentiation of number of infections and residuals of regression are presented on maps (Fig. 7A and 7B). Assumption of proper assessment is based also on the observed trends of model fit statistics (see Table 2). Calculated distance decay parameter β equal *-0.11* reveals that in given county the average increase of infections will be 10.4% with the growth of certain population (treated as explanatory variable, i.e. contacts) per unit of distance (per 1 km). Maps on Figs. 7A (and 7B, 7C, 7D) prove that the model presents gap and underestimate number of infected. It is expected from the parameters of model fit statistics, especially adjusted coefficient of determination (85%). Taking into account for modelling (for above chosen unconstrained gravity and destination constrained spatial interaction models) different time intervals i.e. aggregated data of infections between

10th and 42nd week of 2020 and population of counties as explanatory variables the model fit statistics were similar (Table 2).

These models fit statistics for different time intervals prove that outbreak of COVID-19 pandemic targeted over time the most populated area. These values are more reliable comparing them to data concerning one-week intervals.

Modelled values of number of infections by counties let then estimate the numbers of failed (deceased) based on e.g. general parameters of previously collected pandemic data and at last calculate the number of recovered, subtracting from the number of susceptible the volume of failed cases. It seems that longer time interval is taken into account the more responsible inferring about the general trends can be drawled from models.

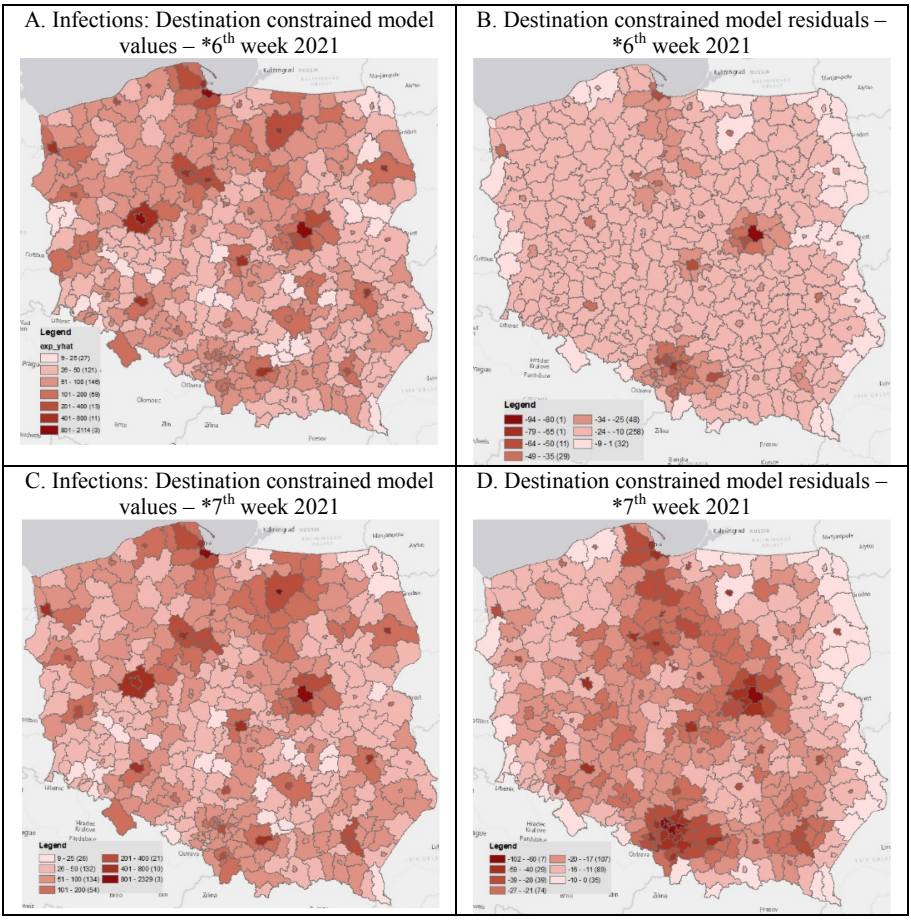


Fig. 7. Destination constrained model: COVID-19 infections in Poland by counties; sixth week 2021 - map of modeled values (A) and residuals (B, modelled - real value); seventh week 2021 - map of modeled values (C) and residuals (D)

Table 2. Spatial interactions models fit statistics (for different time intervals in early stages of outbreak)

model	Unconstrained gravity model				Destination (attraction) constrained model			
week No 2020	pseudo R^2	adj. pseudo R^2	AIC	SRMSE	pseudo R^2	adj. pseudo R^2	AIC	SRMSE
10–22/sum/	0.122	0.122	1.75e+07	2.203	0.609	0.609	7.777e+06	2.018
10–42/sum/	0.484	0.484	3.23e+07	1.607	0.947	0.947	3.33e+06	0.822
1–7 [2021] /sum/	0.779	0.779	925496.67	27.18	0.851	0.851	624561.11	21.49

4.2 Modeling COVID-19 Based on Spatial Interactions

Verification of chosen spatial interactions models for aggregated data of infections by counties in sixth week 2021 (8 Feb - 14 Feb 2021) was made based on empirical acquired. This time it was not possible to use destination-constrained model due to unknown spatial differentiation of infections by counties. The only input value was the (past) total number of infections and use the unconstrained gravity model.

The approaches of use unconstrained gravity model aimed detailed simulation of spatial differentiation of infected by counties during seven weeks of 2021. First, application of the model; second construction of new model with aggregated data since the beginning of the 2021 year. These models reflect (seventh) one-week and seven weeks modeling.

Unconstrained gravity models took different forms for one week (sixth and seventh weeks of 2021, Eq. 4 and 5) and for seven weeks (Eq. 6). The results of modelling are presented on maps (Fig. 8A, 8B, 8C).

$$T_{ij} = exp(-9.6409 + 0.6226\ ln\ (V_i) + 0.6226\ ln\ \left(W_j\right) - (-0.1102)\ ln\ \left(d_{ij}\right)) \quad (5)$$

$$T_{ij} = exp(-9.7187 + 0.6281\ ln\ (V_i) + 0.6281\ ln\ \left(W_j\right) - (-0.1101)\ ln\ \left(d_{ij}\right)) \quad (6)$$

$$T_{ij} = exp(-6.8791 + 0.6023\ ln\ (V_i) + 0.6023\ ln\ \left(W_j\right) - (-0.1153)\ ln\ \left(d_{ij}\right)) \quad (7)$$

where: T_{ij} – modelled number of infections, V_i – population of county i, W_j – population of county j, d_{ij} – distance between county i and j; for the county itself the distance taken into account was the radius of a circle of area equal to county area.

Application of unconstrained gravity model for seventh week 2021 may be treated as revealing current level of infections. The cumulative number of modeled infection cases, using number of population of counties as explanatory variables within the unconstrained gravity model shows significantly higher values. It is much more than officially reported data (based on reported COVID-19 tests and treatments). "Analysis by the Interdisciplinary Centre for Mathematical and Computational Modeling University of

Warsaw (ICM UW) shows that the actual numbers of daily coronavirus infections in recent days may be about ten times higher than the numbers of confirmed infections reported by the government. That is, 20,000 confirmed infections could mean as many as ca. 200,000 actual infections that day." [20, 40]. That is why the model under research has been calibrated proportionally to the estimated values obtained using external data from SIR model [18], retaining the spatial differentiation provided by unconstrained gravity model. The predicted total number of infections in Poland showed over 40 thousands infections (42206 cases, Fig. 8D) during the studied seventh week. However, again the SIR-F model used also the data from the Ministry of Health, taken from the Report published on coronavirus infections Sars-CoV-2 [18, 19]. Much less reliable is map

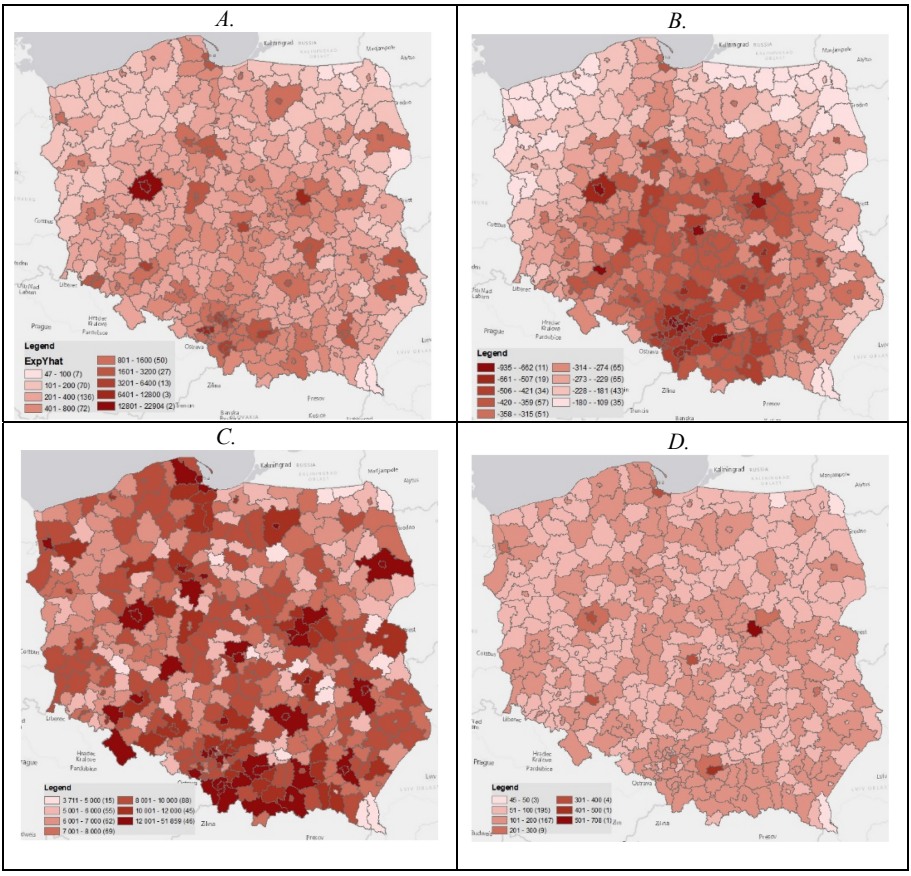


Fig. 8. Modeling of COVID-19 infected cases in Poland; A. Modeled infected cases – seven weeks 2021; aggregated unconstrained gravity model; B. Residuals of infected cases – seven weeks 2021; aggregated unconstrained gravity model; C. Modeling of infected cases – based on seven weeks 2021; aggregated unconstrained gravity and calibrated using data from external SIR models; D. Modeling of infected cases – seventh week of 2021; aggregated unconstrained gravity and calibrated using data from external SIR models

simulating infected cases for the seven weeks (Eq. 6, Fig. 8A, 8B) because the model is based on assumption that the situation of outbreak was not changing rapidly like in the past weeks. Estimated number of infected cases reached over one and half million people (modelled total value of infected cases equal for seven weeks using unconstrained gravity model, Fig. 8C).

Discussion and Conclusions

Both spatial interaction and SIR-F model were evaluated for constant explanatory variables i.e. the number of possible contacts were evaluated using as the constants the officially published number of inhabitants (by counties and total in Poland) and only changing parameters estimated during modelling were the distance-decay parameters, limiting the possible number of infected cases. The modelling just aimed verification of the matter of distance. That is why no data of predicted failed cases (deceased) were evaluated. The promising results of modeling with both models lead us to consider introducing as variables the number of recovered cases and vaccinations, which can be successively subtracted from the initial explanatory variables, i.e. the total population and population by counties. Unfortunately, up to now only total number of people vaccinated in Poland is available (starting since the beginning of 2021). A demographic stratification of contacts can be introduced into the gravity model described earlier taking age cohorts (young, working age, post-working age, or more specifically) - the interacting masses then change, but the distances remain the same.

As dependent variables, one can choose, for example, the average number of infected per week, the maximum number of infected per week instead of (as before) the sums. The input parameters can also be manipulated: either by changing the distance function or by changing the exponent of the power of the distance estimated in the models (e.g. as the mean, minimum or maximum) from a number of time periods under study, assuming a constant maximum (potential) density of contacts as interacting masses.

The number of contacts can be modified by subtotaling the number of deaths and vaccinations. Visitors (in tourist areas) and foreign visitors to the area can be added as components of the sum of contacts. Simulated results can be aggregated, e.g., to provinces and for the whole country. The paper presents the intermediate results of the ongoing research in recent state.

Acknowledgment. *This research was partially funded by IDUB against COVID-19 project granted by Warsaw University of Technology under the program Excellence Initiative: Research University (IDUB).*

References

1. Past plagues offer lessons for society after the coronavirus pandemic. https://www.sciencenews.org/article/coronavirus-covid-19-ancient-plagues-pandemics-lessons-society. Accessed on 20 Jan 2021
2. Smith, C.D., Mennis, J.: Incorporating geographic information science and technology in response to the COVID-19 pandemic. Prev. Chronic. Dis. **17** (2020). https://doi.org/10.5888/pcd17.200246.

3. Dangermond, J., Vito, C.D., Pesaresi, C.: Using GIS in the Time of the COVID-19 Crisis, casting a glance at the future. A joint discussion. J-READING - J. Res. Didactics Geography. **1**, 195–205 (2020).
4. Petersen, E., et al.: Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. Lancet. Infect. Dis **20**, e238–e244 (2020). https://doi.org/10.1016/S1473-3099(20)30484-9
5. Murgante, B., Borruso, G., Balletto, G., Castiglia, P., Dettori, M.: Why Italy first? health, geographical and planning as-pects of the COVID-19 Outbreak. Sustainability. **12**, 5064 (2020). https://doi.org/10.3390/su12125064
6. Paez, A., Lopez, F.A., Menezes, T., Cavalcanti, R., da Rocha Pitta, M.G.: A spatio-temporal analysis of the environmental correlates of COVID-19 incidence in Spain. Geograph. Anal. **53**(3), 397–421 (2020)
7. Confalonieri, U., et al.: Human health. Climate change 2007: impacts, adaptation and vulnerability: contribution of Working Group II to the fourth assessment report of the Intergovernmental Panel on Climate Change (2007)
8. Kimura, Y., et al.: Geodemographics profiling of influenza A and B virus infections in community neighborhoods in Japan. BMC Infect. Dis. **11**, 1–12 (2011)
9. Merler, S., Ajelli, M.: The role of population heterogeneity and human mobility in the spread of pandemic influenza. Proc. Royal Soc. B: Biol. Sci. **277**, 557–565 (2010)
10. Duszyński, J., et al.: Zrozumieć COVID-19. Opracowanie zespołu ds. COVID-19 przy Prezesie PAN (2020)
11. Hamidi, S., Sabouri, S., Ewing, R.: Does density aggravate the COVID-19 pandemic?: early findings and lessons for planners. J. Am. Plann. Assoc. **86**, 495–509 (2020). https://doi.org/10.1080/01944363.2020.1777891
12. Śleszyński, P.: Prawidłowości przebiegu dyfuzji przestrzennej rejestrowanych zakażeń koronawirusem SARS-COV-2 w Polsce w pierwszych 100 dniach epidemii. Czasopismo Geograficzne**91**, 5–19 (2020)
13. Iwańczak, B.: Zróżnicowanie przestrzenne zagrożenia grypą w Polsce. Prace i Studia Geograficzne **57**, 127–144 (2015)
14. Kirby, R.S., Delmelle, E., Eberth, J.M.: Advances in spatial epidemiology and geographic information systems. Ann. Epidemiol. **27**, 1–9 (2017). https://doi.org/10.1016/j.annepidem.2016.12.001
15. Brauer, F.: Mathematical epidemiology: past, present, and future. Infectious Disease Modelling. **2**, 113–127 (2017). https://doi.org/10.1016/j.idm.2017.02.001
16. Kiss, I.Z., Miller, J.C., Simon, P.L.: Mathematics of Epidemics on Networks, vol. 598. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-50806-1
17. Pourbohloul, B., Miller, J.: Network Theory and the Spread of Communicable Diseases (2021)
18. CovsirPhy Development Team: CovsirPhy, Python package for COVID-19 analysis with SIR-derived ODE models. https://github.com/lisphilar/covid19-sir. Accessed on 04 Feb 2021
19. Guidotti, E., Ardia, D.: COVID-19 Data Hub. JOSS. **5**, 2376 (2020). https://doi.org/10.21105/joss.02376.
20. ICM UW Model description – COVID. https://covid-19.icm.edu.pl/en/model-description/. Accessed on 12 Feb 2021
21. Stewart, J.Q.: Empirical mathematical rules concerning the distribution and equilibrium of population. Geogr. Rev. **37**, 461–485 (1947)
22. Pomianowski, W.: Transportation network structure and spatial accessibility. Dynamic graph approach (2018)
23. Johnston, R.: The Dictionary of Human Geography. Wiley-Blackwell, Oxford (1985)
24. Yan, X.-Y., Zhou, T.: Destination choice game: a spatial interaction theory on human mobility. Sci. Rep. **9**, 9466 (2019). https://doi.org/10.1038/s41598-019-46026-w
25. Wilson, A.G.: Entropy in Urban and Regional Modelling. Pion, London (1970)

26. O'Kelly, M.E.: Spatial interaction models. In: International Encyclopedia of Human Geography. pp. 365–368. Elsevier (2009). https://doi.org/10.1016/B978-008044910-4.00529-0.
27. Rees, P., Dennett, A.: Alan Wilson – contributions to research on population and migration. Interdisc. Sci. Rev. **44**, 249–271 (2019). https://doi.org/10.1080/03080188.2019.1670428
28. Wilson, A.: Epidemic models with geography. https://arxiv.org/abs/2005.07673v1 (2020)
29. Werner, P.: Simulating of accessibility of ICT infrastructure in Poland using geographical potential models. In: Proceedings of the 16th International Conference on System Science. pp. 261–269. Wroclaw University of Technology (2007)
30. Werner, P.: Symulacja Monte Carlo rozwoju sieci teleinformatycznych, radiokomunikacji i radiodyfuzji w Polsce (na podstawie modeli potencjału geograficznego). Archiwum Fotogrametrii, Kartografii i Teledetekcji. vol. 13 (2003)
31. Werner, P.: Geograficzne uwarunkowania rozwoju infrastruktury spoleczenstwa informacyjnego w Polsce. Uniwersytet Warszawski, Wydz. Geografii i Studiów Regionalnych, Warszawa (2003)
32. Werner, P.: Konstrukcja i interpretacja modeli potencjału geograficznego w badaniach sieci teleinformatycznych. In: Problemy interpretacji wyników metod badawczych stosowanych w geografii społeczno-ekonomicznej i gospodarce przestrzennej: praca zbiorowa. Bogucki Wydaw. Naukowe (2003)
33. Tobler, W.R.: Cellular geography. In: Gale, S. and Olsson, G. (eds.) Philosophy in Geography. pp. 379–386. Springer Netherlands, Dordrecht (1979). https://doi.org/10.1007/978-94-009-9394-5_18.
34. Batty, M.: Fifty years of urban modeling: Macro-statics to micro-dynamics. In: The dynamics of complex urban systems. pp. 1–20. Springer (2008). https://doi.org/10.1007/978-3-7908-1937-3_1
35. Wilson, A.G.: Complex spatial systems: the modelling foundations of urban and regional analysis. Pearson Education (2000)
36. Oshan, T.M.: A primer for working with the Spatial Interaction modeling (SpInt) module in the python spatial analysis library (PySAL). REGION. **3**, 11 (2016). https://doi.org/10.18335/region.v3i2.175.
37. Fotheringham, A.S., O'Kelly, M.E.: Spatial interaction models: formulations and applications. Kluwer Academic Publishers, Dordrecht, Boston (1989)
38. Raport zakażeń koronawirusem (SARS-CoV-2) - Koronawirus: informacje i zalecenia - Portal Gov.pl. https://www.gov.pl/web/koronawirus/wykaz-zarazen-koronawirusem-sars-cov-2. Accessed on 07 Feb 2021
39. Akaike, H.: A new look at the statistical model identification. IEEE Trans. Automat. Contr. **19**, 716–723 (1974). https://doi.org/10.1109/TAC.1974.1100705
40. 20 tys. zakażeń dziennie? Eksperci mówią, że tę liczbę należy pomnożyć przez 10. https://biqdata.wyborcza.pl/biqdata/7,159116,26507630,20-tys-zakazen-dziennie-eksperci-mowia-ze-te-liczbe-nalezy.html. Accessed on 12 Feb 2021