Epidemics and Their Implications in Urban Environments: A Case Study on a National Scope

Robert Elsässer, Adrian Ogierman and Michael Meier

Abstract In times where urbanization becomes more important every day, epidemic outbreaks may be devastating. Powerful forecasting and analysis tools are of high importance for both, small and large scale examinations. Such tools provide valuable insight on different levels and help to establish and improve embankment mechanisms. Here, we present an agent-based algorithmic framework to simulate the spread of epidemic diseases on a national scope. Based on the population structure of Germany, we investigate parameters such as the impact of the number of agents, representing the population, on the quality of the simulation and evaluate them using real world data provided by the Robert Koch Institute [4, 22]. Furthermore, we empirically analyze the effects of certain non-pharmaceutical countermeasures as applied in the USA against the Influenza Pandemic in 1918–1919 [18]. Our simulation and evaluation tool partially relies on the probabilistic movement model presented in [8]. Our empirical tests show that the amount of agents in use may be crucial. Depending on the existing knowledge about the considered epidemic, this parameter alone may have a huge impact on the accuracy of the achieved simulation results. However, with the right choice of parameters—some of them being obtained from real world observations [10]—one can efficiently approximate the course of a disease in real world.

Keywords Epidemic algorithms · Power law distribution · Disease spreading · Public health

R. Elsässer
Department of Computer Sciences, University of Salzburg, 5020 Salzburg, Austria

A. Ogierman (✉) · M. Meier
Department for Computer Science, University of Paderborn, 33102 Paderborn, Germany
e-mail: adriano@upb.de

1 Introduction

In order to improve our chances to control an epidemic outbreak, we need proper models which describe the spread of a disease. Institutes, governments, and scientists all over the world work intensively on forecasting systems in order to be well prepared if an unknown disease appears.

In recent years a huge amount of theoretical and experimental study has been conducted on this topic. While theoretical analysis provides important and sometimes even counter intuitive insights into the behavior of an epidemic (e.g. [6, 8]), in an experimental study one can take many different settings and parameters [16, 17] into account. These usually cannot be considered simultaneously in a mathematical framework. A specific topology, for example, may have its own attributes that are completely different in other topological settings.

The goal of this paper is to present and empirically analyze a dynamic model for the spread of epidemics in an extended manner. One of our objectives is to find the right parameters, which lead to realistic settings. Therefore, we investigate a general simulation environment, in which the different parameters can easily be adjusted to real world observations. A second objective is to evaluate similarities between countermeasure approaches in our model and the real world. We use empirical data for the comparison. Our tool is agent-based, i.e., the individuals (or groups of such) are modeled by agents interacting with each other. The environment approximates the geography of Germany, in which agents may travel between cities. Within a city the agents interact according to the probabilistic model presented in [8] in a distributed manner. For a detailed description of the algorithmic framework see Sect. 2. In order to describe the problems and the related work, we often utilize the style and wording of [8].

1.1 Related Work

There is plenty of work considering epidemiological processes in different scenarios and on various networks. The simplest model of mathematical disease spreading is the so called SIR model (see e.g. [14, 20]). The population is divided into three categories: susceptible (S), i.e., all individuals which do not have the disease yet but can become infected, infective (I), i.e., the individuals which have the disease and can infect others, and recovered (R), i.e., all individuals that recovered and have permanent immunity (or have been removed from the system). Most papers model the spread of epidemics using a differential equation based on the assumption that any susceptible individual has uniform probability β to become infected from any infective individual. Furthermore, any infected player recovers at some stochastically constant rate γ.

This traditional (fully mixed) model can easily be generalized to a network. It has been observed that the corresponding process can be modeled by bond percolation on the underlying graph [13, 19]. Interestingly, for certain graphs with a

power law degree distribution, there is no constant threshold for the epidemic outbreak as long as the power law exponent is less than 3 [20] (which is the case in most real world networks, e.g. [1, 3, 11, 21]). If the network is embedded into a low dimensional space or has high transitivity, then there might exist a non-zero threshold for certain types of correlations between vertices [19]. However, none of the papers above considered the dynamic movement of individuals, which seems to be the main source of the spread of diseases in urban areas [10].

In [10] the physical contact patterns are modeled by a dynamic bipartite graph, which results from movement of individuals between specific locations. The graph is partitioned into two parts. The first part contains the people who carry out their daily activities moving between different locations. The other part represents the various locations in a certain city. There is an edge between two nodes, if the corresponding individual visits a certain location at a given time. Obviously, the graph changes dynamically at every time step.

In [7, 10] the authors analyzed the corresponding network for Portland, Oregon. According to their study, the degrees of the nodes describing different locations follow a power law distribution with exponent around 2.8.[1] For many epidemics, transmission occurs between individuals being simultaneously at the same place, and then people's movement is mainly responsible for the spread of the disease.

The authors of [8] considered a dynamic epidemic process in a certain (idealistic) urban environment incorporating the idea of attractiveness based distributed locations. The epidemic is spread among n agents, which move from one location to another. In each step, an agent is assigned to a location with probability proportional to its attractiveness. The attractiveness' of the locations follow a power law distribution [10]. If two agents meet at some location, then a possible infection may be transmitted from an infected agent to an uninfected one. The authors obtained two results. First, if the exponent of the power law distribution is between 2 and 3, then at least a small (but still polynomial) number of agents remains uninfected and the epidemic is stopped after a logarithmic number of rounds. Secondly, if the power law exponent is increased to some large constant, then the epidemic will only affect a polylogarithmic number of agents and the disease is stopped after $(\log \log n)^{\mathcal{O}(1)}$ steps. In this case each agent is allowed to spread the disease for a number of time steps, which is bounded by some large constant.

In addition to the theoretical papers described above, plenty of simulation work exists. Two of the most popular directions are the so called agent-based and structured meta-population-based approach, respectively (cf. [2, 15]). Both models have their advantages and weaknesses. The main idea of the meta-population approach is to model whole regions, e.g. georeferenced census areas around airport hubs [5], and connect them by a mobility network. Then, within these regions the spread of epidemics is analyzed by using the well known mean field theory. The agent-based approach models individuals with agents in order to simulate their behavior. In this

---

[1] In [10] the degree represents the number of individuals visiting these places over a time period of 24 h.

context, the agents may be defined very precisely, including e.g. race, gender, educational level, etc. [16, 17], and thus provide a huge amount of detailed data conditioned on the agents setting. Furthermore, these kinds of models are also able to integrate different locations like schools, theaters, and so on. Thus, an agent may or may not be infected depending on his own choices and the ones made by agents in his vicinity. The main issue of the agent-based approach is the huge amount of computational capacity needed to simulate huge cities, continents or even the world itself [2]. This limitation can be attenuated by reducing the number of agents, which then entails a decreasing accuracy of the simulation. In the meta-population approach the simulation costs are lower, sacrificing accuracy and some kind of noncollectable data.

A specific field of application of such simulations is the investigation of the impact of (non-)pharmaceutical countermeasures on the behavior of epidemics. Germann et al. [12] investigated the spread of a pandemic strain of the influenza virus through the U.S. population. They used publicly available 2000 U.S. Census data to identify seven so-called mixing groups, in which each individual may interact with any other member. Each class of mixing group is characterized by its own set of age-dependent probabilities for person-to-person transmission of the disease. They considered different combinations of socially targeted antiviral prophylaxis, dynamic mass vaccination, closure of schools and social distancing as countermeasures in use, and simulated them with different basic reproductive numbers R_0. It turned out that specific combinations of the countermeasures have a different influence on the spreading process. For example, with $R_0 = 1.6$ social distancing and travel restrictions did not really seem to help, while vaccination limited the number of new symptomatic cases per 10,000 persons from ~ 100 to ~ 1. With $R_0 = 2.1$, such a significant impact could only be achieved with the combination of vaccination, school closure, social distancing and travel restrictions.

1.2 Our Results

The results of this paper are two-fold. First, we show that by increasing the number of agents we are able to significantly improve the accuracy of our results in the scenarios we have tested. This is due to different phenomena which are only visible if the amount of agents in use is large enough. For example, if the number of agents exceeds a certain value, then the epidemic manages to keep a specific (low) amount of infected individuals over a long time period. Furthermore, the number of agents has to be above some threshold to allow the epidemic to enter some specific areas/ cities in the environment we used. Obviously, a certain amount of agents is also needed to avoid significant fluctuations in our results.

The second main result of this paper is that by setting the parameters properly, one can approximate the effect of some non-pharmaceutical countermeasures, that are usually adopted if an epidemic outbreak occurs. This observation is supported by the empirical study of [18]. Interestingly, the right choices of parameters in our experiments seem to be in line with previous observations in the real world (e.g. the

right power law exponent seems to be in the range of 2.6–2.9, cf. [10]). To analyze the effect of the countermeasures mentioned above, we integrate the corresponding mechanisms on a smaller scale, and then verify their impact on a larger scale too.

2 Theoretical Models and Algorithmic Framework

In the following we provide an empirical analysis on a small as well as on a large scale. Hereby, the cities are chosen from a list in descending order of their population size. It is intuitively clear that large (and thus attractive) cities play a major role for the epidemic pace since a higher population density entails a potentially higher infection probability. Excluding such hotspots would of course slow down the infection spread. The problem is, this could only be achieved by putting the whole city into quarantine. However, isolating an entire city is not a trivial task. For example, people living in the suburbs but working in the city might not be willing to risk their job by obeying the quarantine. Therefore, we consider such strategies only on a smaller scale.

In our model the agents may not only move between locations within a city but between cities as well. Furthermore, due to simplicity, the agents are not categorized (i.e., they do not provide further properties like gender etc.). Since we are not interested in the evaluation of such details. In the following, we briefly introduce the model. Due to readability we present our model with respect to the following four categories: (1) The environment on a large scale (inter-city movement), (2) The environment on a small scale (intra-city movement), (3) The epidemic model, and (4) The countermeasure model.

2.1 The Environment on a Large Scale

Let $G(d) = (V, E)$ be a complete graph with m nodes $V := \{c_1, \ldots, c_m\}$ and parameter $d := \{d_{c_1}, \ldots, d_{c_m}\}$, whereas $c_i = G_{c_i}$ (see below) represents the corresponding city c_i. Let further p_{c_i} be the population size of c_i and $p := \{p_{c_i}, \ldots, p_{c_m}\}$. Then the attractiveness of c_i is given by $d_{c_i} = \dfrac{p_{c_i}}{\sum_{1 \leq i \leq m} p_{c_i}}$. Consequently, G represents the topology, which contains all cities, on a large scale. In other words, we model the inter-city movement using the complete graph $G(d)$. In this graph, each $c_i \in V$ corresponds to a city of Germany. However, depending on the size, not every city of Germany is represented by a node in V.[2] The population is represented by

---

[2] The amount of overall agents in use (n) determines how many cities are represented by V. Therefore we sort the list of all cities of Germany in descending order of their population size. Then, starting from the top, we include the currently considered city c_i to V if and only if the assigned amount of agents to said city is at least 1. The latter amount is given by $n \cdot d_{c_i}$.

$n = \sum_{1 \le i \le m} nd_{c_i}$ agents, with nd_{c_i} being the number of agents assigned to c_i. Note that said number is proportional to the cities real world population. Furthermore, each city $c_i \in V$ is assigned an attractiveness d_{c_i} proportional to its population size (w.r.t. the whole population). Let $A_{i,s,t}$ be the event that agent i travels from city s to t. Let further p' be the probability that an agent decides to travel at all, and let $dist(s,t)$ be the Euclidean distance between cities s and t. Then the probability that event $A_{i,s,t}$ occurs is given by

$$Pr(A_{i,s,t}) = p' \cdot \frac{d_t \cdot dist^{-1}(s,t)}{\sum_{(s,j) \in E} d_j \cdot dist^{-1}(s,j)}.$$

Thus, the probability for an agent entering a specific city depends on the distance to said city, its population size as well as the current position of the considered agent.

2.2 The Environment on a Small Scale

Let $G_{c_i}(d(v)) = \{V_{c_i}, E_{c_i}\}$ be a complete graph with $m_{c_i} = \lceil \kappa nd_{c_i} \rceil$ nodes (also called *cells*), with $d(c_i) := \{d_{v_1}, \ldots, d_{v_{m_c}}\}$. Here, $\kappa > 0$ is a constant, which will be specified in the upcoming experiments. Further, note that κ does not affect the amount of agents but the amount of cells only. Then the attractiveness of cell $v_i \in V_{c_i}$ within a city c_i is given by d_{v_i}. Said attractiveness is chosen randomly with probability proportional to $1/\beta^\alpha$ for a value $\beta > 1$, where $\alpha > 2$ is a constant depending on the simulation run. In other words, each $c_i \in V$, representing city c_i, is a clique of cells on its own, thus incorporating intra-city movement into our model. The cells represent locations within a city an agent can visit. Each cell may contain agents (individuals), depending on the cells so-called *attractiveness*. If an agent decides to stay within its current city, said agent moves to a randomly chosen cell according to the distribution of the attractiveness' among the cells. This also holds for the first cell an agent is accommodated in after entering a city.

2.3 Epidemic Model

We use three different states to model the distributed spreading process. These states partition the set of agents into three groups; $\mathcal{I}(j)$ contains the infected agents in step j, $\mathcal{U}(j)$ contains the uninfected (susceptible) agents in step j, and $\mathcal{R}(j)$ contains the resistant agents in step j. Whenever it is clear from the context, we simply write $\mathcal{I}, \mathcal{U}$, and $\mathcal{R}$, respectively. If an uninfected agent i visits a cell (within

a city) which also contains agents of $\mathcal{I}(j)$, then i becomes infected with probability $1 - (1 - \gamma)^{|\mathcal{I}'(j)|}$, where $\mathcal{I}'(j)$ represents the set of infected agents accommodated in the same cell as i. We refer to the concrete value of γ in the upcoming simulations.

2.4 Countermeasure Models

Our countermeasure models mainly take advantage of the parameters α and κ. That is, high values of these two parameters imply a high level of countermeasures and vice versa. With countermeasures applied, individuals avoid places with a large number of persons more often, waive needless tours, and are more careful when meeting other people. While α is mostly responsible for a decreasing number of visitors within a cell, and thus for the avoidance of crowded areas for example, κ determines the total space available for all individuals. As pointed out in [18], a single countermeasure alone is most likely not sufficient to stop an epidemic. Therefore, we assume a combination of them to be in place, which then would be able to sufficiently influence the parameters α and κ. Note, our countermeasure models apply to each city $c_i \in V$ individually.

We use two different types of countermeasure-models: a (multi-tier) *level based approach* considering the amount of infected agents in the current step, and a *ratio based approach* considering the amount of newly infected agents in the current step compared to the one in the step before. In the following we use α_0 and κ_0 as initial values for α and κ, respectively.

In the *level based model* we have one or more levels in which a certain pair of parameters α and κ is used. Let LM_m stand for the level based model with m levels $L = \{l_1, \ldots, l_m\} \cup l_0$. Further, let $T = \{l_0^d, l_0^u, l_1^d, l_1^u, \ldots, l_m^d, l_m^u\}$ be the set of transition points for all levels, i.e., l_i^d defines the transition point from level i to $i - 1$ whereas l_i^u defines the transition point from level i to $i + 1$. Note that, $l_i^u = l_{i+1}^d$ does not necessarily hold. Additionally, $\alpha_0, \alpha_1, \ldots, \alpha_m$ and $\kappa_0, \kappa_1, \ldots, \kappa_m$ define the parameters α and κ, which are applied in the corresponding levels $l_0, \ldots, l_m$. Figure 1a depicts an example situation.

In contrast, the *ratio based model RM* uses a non static approach. Let the set of newly infected nodes of a city c_i in step j be denoted by $\mathcal{I}_{c_i}^*(j)$. Furthermore, let α_j and κ_j denote the corresponding parameters used in step j. If $\frac{|\mathcal{I}_{c_i}^*(j)|}{|\mathcal{I}_{c_i}^*(j-1)|} \geq a$, for some constant a, then we set $\alpha_{j+1} = \alpha_j + z_1$ and $\kappa_{j+1} = \kappa_j + z_2$, where z_1, z_2 are some small constants which will be specified later. Consequently, we set $\alpha_{j+1} = max\{\alpha_0, \alpha_j - z_1\}$ and $\kappa_{j+1} = max\{\kappa_0, \kappa_j - z_2\}$ whenever $\frac{|\mathcal{I}_{c_i}^*(j)|}{|\mathcal{I}_{c_i}^*(j-1)|} \leq 1/a$. The values applied in the various models are specified in Sect. 3.3. Figure 1b depicts an example situation.
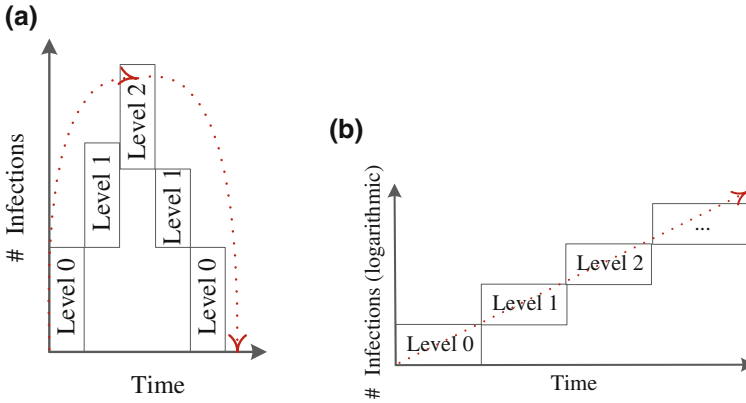
Fig. 1 Basic examples for the two countermeasure models. **a** An example demonstrating a possible configuration of LM_2. The *broken line* depicts the amount of total infections in the city while the scope of activity of the countermeasure levels is represented by the *rectangles*. **b** An example demonstrating a possible configuration of *RM*. The *broken line* depicts the amount of total infections in the city (on a logarithmic scale) while the scope of activity of the countermeasure levels is represented by the *rectangles*

3 Experimental Analysis

The environment in our simulations approximates the geography of Germany uti-lizing 10 million agents and more. Note that, the obtained results remain similar utilizing up to 100 million agents. Depending on the number of agents, our sim-ulations incorporate several hundred cities as visitable areas spread all over the country. Only cities with an initial agent amount of at least one are included in the simulation. Each city is assumed to be reachable from any other city. However, an agent may travel at most 1,000 km within a single round. Each round represents a whole day in the real world. Consequently, an agent moving from one city to another has to wait until its destination is reached before it can interact with other agents at said destination.

3.1 Simulations

In the upcoming sections we present and evaluate our results with the focus on: (1) the impact of the number of agents on the characteristics of the simulated epidemic compared to real world data, and (2) the impact of non-pharmaceutical counter-measures on the behavior of the epidemic (e.g. social distancing, school closures, and isolation [18]).

Furthermore, we also analyze our parameter settings. Although this is only a short part, our settings seem to coincide with the real world observations of [10],

and thus provide an additional valuable insight. Note that the figures presented in this section show values based on the real world population size and not on the number of agents.

3.2 Relevance of the Chosen Parameters

Based on real world observations (e.g. [10]), we chose $\alpha = 2.8$ and $\kappa = 1$ as a starting point for a series of simulations concerning α and κ, respectively. Each plot represents the average value of 50 different simulation runs for each parameter constellation utilizing 10 million agents. The parameter notation is as follows: γ is the probability for an agent $v \in \mathcal{U}$ to become infected independently by each $w \in \mathcal{I}$ occupying the same cell at the same time, ξ is the amount of rounds an agent $v \in \mathcal{I}$ is infectious, thus being able to infect others, $City_{init}$ is the initial amount of cities the infection is being placed in, $Agent_{init}$ is the amount of initially infected agents which are placed in $City_{init}$ different cities, α is the power law exponent used to individually compute the attractiveness of the cells within each city, and κ is a multiplicative factor to increase/decrease the amount of cells proportional to the initially assigned amount of agents.

Figure 2a and b show the impact of α and κ on the behavior of the epidemic, and compare the results to the characteristics of a typical Influenza case reported by the RKI.[3] To increase the readability, we omit to add the RKI-plot as the 6th one. Instead we refer the reader to Fig. 3. Among all five α-values $\alpha = 2.8$, which has also been obtained in [10] in a different context, represents the best tradeoff between curve similarity and amount of infections. All remaining parameters were set to identical values as in Case 1 (cf. Sect. 3.3). For κ a similar phenomenon can be observed. With increasing κ (including fractional values), the characteristics of the curve (i.e., the amount of infected individuals at the peak vs. total number of infections and total duration of the outbreak) become less and less accurate. Even if we increase the value to 2, the obtained curve does not follow the characteristics of the real world observations reported by the RKI [see footnote 3] anymore.

In terms of the probability of infection γ we simply chose a reasonable value low enough to model an Influenza epidemic, but high enough to provoke an outbreak. This seemed reasonable due to the (at least to our knowledge) lack of concrete values deduced from real world observations.

---

[3] The Robert Koch Institute (RKI) is the central federal institution in Germany responsible for disease control and prevention and is therefore the central federal reference institution for both, applied and response-orientated research. (*Source* http://www.rki.de/EN/Home/homepage_node.html).
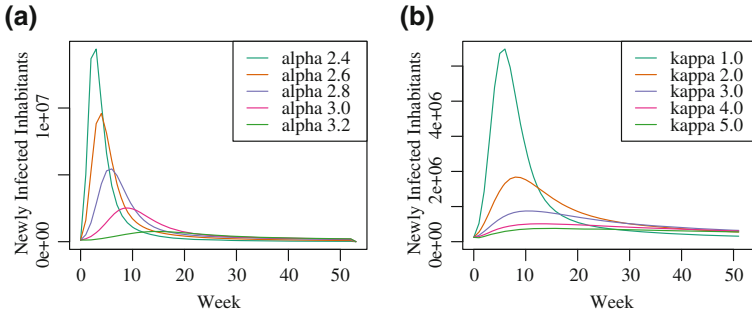
(a)

(b)

Fig. 2 A composition of multiple simulation runs concerning varying α (**a**) and κ (**b**) only. All other parameters are identical to Case 1 (cf. Sect. 3.3). Each result represents the average of 50 different simulation runs with 10 million agents for the topology of Germany. **a** $2.4 \le \alpha \le 3.2$. **b** $1 \le \kappa \le 5$

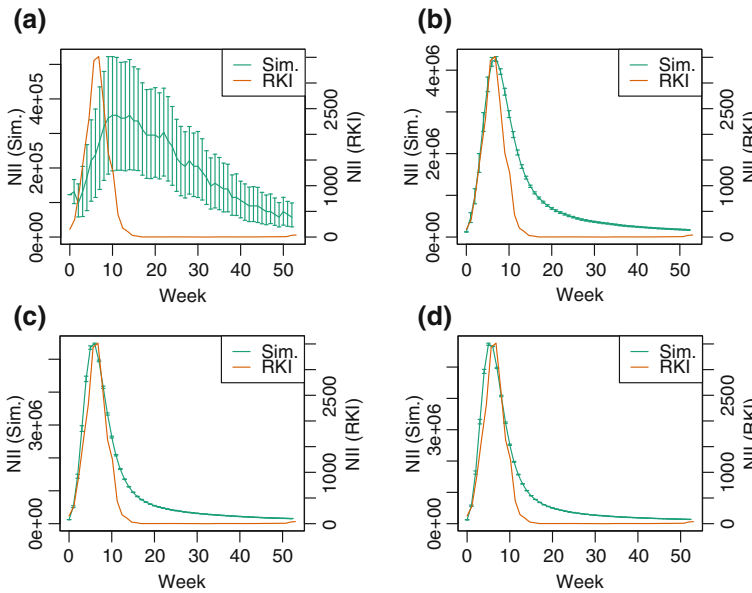**(a)**

(b)

(c)

(d)

Fig. 3 Simulation results for Case 1a) (*green*) in comparison to real world data (*red*) provided by the RKI for a varying amount of agents. The abbreviation *NII* stands for *Newly Infected Inhabitants*. All results up to and including 10 million agents represent the average values of 50 different simulation runs, whereas the result for 100 million agents averages over 20 different simulation runs. The reliability of the averaged value is indicated by the corresponding confidence interval. **a** 10.000 agents. **b** 500.000 agents. **c** 10.000.000 agents. **d** 100.000.000 agents

3.3 Case 1—Number of Agents

Before we present the results for this case, we first introduce the relevant sources for comparison. Here we compare our results to real world data provided by the RKI [22] SurvStat system for the year 2007. The parameter values were taken from reference data provided by the RKI [4] where possible, or set to reasonable ones otherwise (cf. Sect. 3.2 for more details).

3.3.1 RKI: Basis of Comparison

In the following we compare our results to the real world data provided by the RKI [see footnote 3]. For this purpose we use two different data sources: the official report of the Influenza epidemiology of Germany for 2010/2011 [4] and an online database containing obligatory reports called SurvStat [22].

Relevance. The data for the SurvStat database and the report of 2010/2011 itself were obtained from more than 1 % of all primary care doctors spread all over Germany. This indicates the significance according to international standards. Unfortunately, there are some drawbacks resulting from the type of data ascertainment itself. Note that not every infected person consults a doctor, which implies that the data of the SurvStat system contains only the serious courses of the disease. Nonetheless, these sources provide a valuable tool to obtain insight into the spread and persistence of the Influenza virus in Germany. Further, due to the data's significance, it is possible to estimate the number of infections within certain areas as well. Since the spread of infections in the real world is influenced by factors like seasonal fluctuations or the virus' aggressiveness, the number of total infections may significantly differ from year to year, cf. [22] for different years. However, the course of the curve usually does not change. Consequently, we do not focus on absolute values in our simulations, but on the *characteristics* of our results. These characteristics remain, up to some scaling factor, identical over the whole data set provided by the RKI.

3.3.2 Case 1

The parameters for this subcase are as follows. We set $\gamma = 7\,\%$, ξ to 5 days, the amount of initially infected cities $City_{init}$ to 1 (namely Berlin), and the amount of initially infected agents $Agent_{init}$ to 0.0015 % of the overall agents used for these simulations. Furthermore, $\alpha = 2.8$ and $\kappa = 1$.

Figure 3 shows the results for this case. Here the green curve represents the real-world data provided by the RKI for the year 2007 while the red curve represents our simulation results. Note that both curves vary significantly in terms of absolute numbers. However, this is not our focus here. Due to the level of abstraction in our model and since the RKI data only contains reported cases (see above), the absolute

numbers do not coincide. Additionally, as stated above, the data provided by the RKI also differs significantly (in terms of absolute numbers but not the disease characteristics) from year to year (cf. [22]). Therefore, we focus on the course of the disease and the resulting characteristics of the plotted curves.

It is easy to see that the more agents are used, beginning from Fig. 3a up to 3d, the more the curve characteristics converge. Moreover, the accuracy of each simulation run increases as well (cf. the confidence intervals in Fig. 3). With at least $500{,}000$ agents in use, both curves become similar.

To obtain a more formal evaluation, we define three measures, which are used to compare our results to the data provided by the RKI. These are: the time to peak (TTP), the epidemic duration (ED), and the area of the curve (AC). The time to peak describes the week with the maximum amount of newly infected agents of the corresponding curve. The area of the curve is simply the summation of the area between the origin and the endpoint EP (defined by the epidemic duration). Finally, the endpoint of the epidemic duration is the week in which only a minor amount of new infections occur, and no significant new infections are observed anymore. Minor infections are defined to start in a round i and last till the last round j of the simulation while for all rounds $i \leq i' \leq j$ it holds that the amount of newly infected agents does not exceed 9 % of the maximum value.

In the following we consider the relative values of these measurements compared to the RKI data. That is, the numbers represent the ratio between the value obtained in our simulations and the value provided by the RKI. For example, a value of 4.00 for TTP in the case of $10{,}000$ agents implies that the time to peak in our case divided by the time to peak in the real world data is 4. Using the individual deviation measurements, we define a global deviation value by the formula $\frac{1}{3} \cdot \text{TTP} + \frac{1}{3} \cdot \text{AC} + \frac{1}{3} \cdot \text{ED}$. For simplicity we consider each individual measurement uniformly weighted. The results, which confirm our previous observations, are shown in Fig. 4. All obtained results and previous statements imply a fragile balance between the accuracy, the parameter setting, and the amount of agents in use.
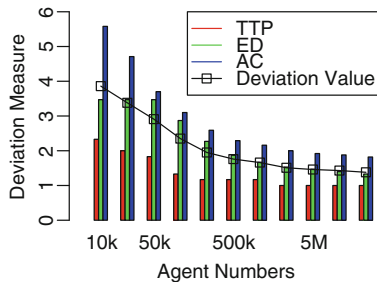

Fig. 4 A visual representation of the data computed using the deviation measurements for the experiments conducted in Case 1

3.3.3 Case 2—Non-pharmaceutical Countermeasures

Now we extend our analysis to incorporate non-pharmaceutical countermeasures such as school closures and social distancing. Here we stick to a fictional epidemic simply because it simplifies the presentation, i.e., due to the increase of γ to 12%, a faster spread is achieved and the impact of the used countermeasures is amplified. Similar to Case 1, we set ξ to 5 days, the amount of initially infected cities $City_{init}$ to 1 (namely Berlin), and the amount of initially infected agents $Agent_{init}$ to 0.00075% of the overall agents (to compensate the higher γ in the beginning). All relevant parameters regarding the countermeasure models can be found in the original paper [9].

We assume that non-pharmaceutical countermeasures basically affect the parameters α and κ, since the individuals will most likely avoid places with a large number of persons, waive needless tours, and be more careful when meeting other people. For obvious reasons, we cannot compare our simulation results to current real-world data containing results for different epidemics with varying (or no) countermeasures in use. Therefore, we use the work of Markel et al. [18] for this purpose. Especially Fig. 5 is of particular interest. There, the direct correlation between establishing countermeasures and a decreasing amount of new infections (and vice versa) is well highlighted. We observed an identical effect in our simulations (cf. Fig. 6). Note that different combinations of countermeasures used in [18] entail different kinds of impacts on the death rates. In contrast, we do not focus on specific combinations but on sufficient ones to actually achieve an immediate impact on the epidemic.

As already described in Sect. 2, we use two different countermeasure approaches: the level based (LM_1, LM_2 and LM_3), and the ratio based (RM), respectively. Both use different mechanisms and parameters to achieve the embankment of the epidemic. Recall that all transition points in the level based model are chosen w.r.t.
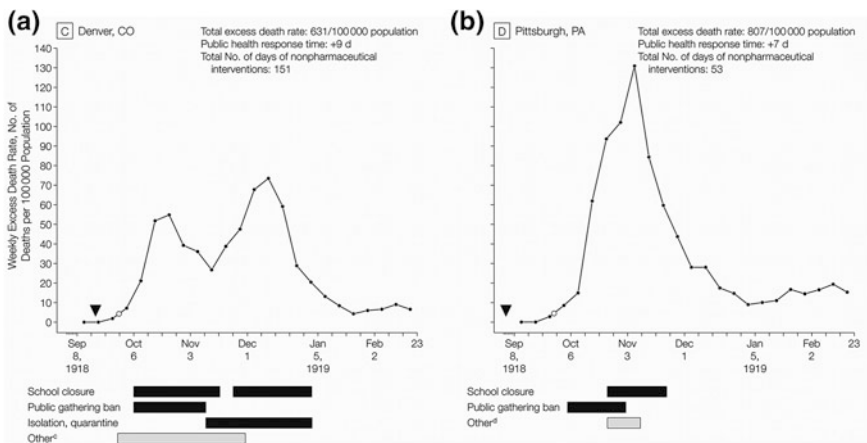

Fig. 5 Weekly excess death rates from September 8, 1918, through February 22, 1919 [18, Fig. 3]
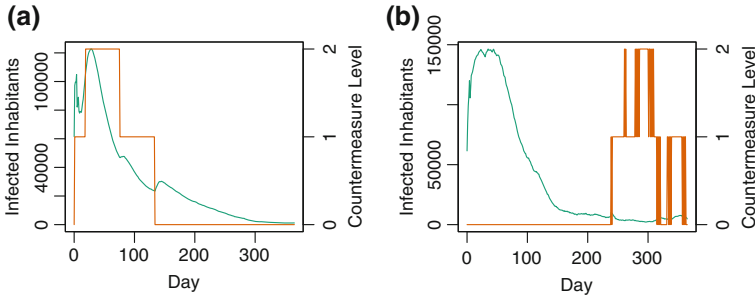
(a) **(b)**

Fig. 6 Example results for BER for Case 2. The *green curve* represents the amount of infected inhabitants in the corresponding city while the *red curve* indicates the activated countermeasure level with respect to the countermeasure model in use. The number of agents refers to the total amount of agents used for the simulation. **a** Model LM_3, 100 million agents. **b** Model RM, 10 million agent
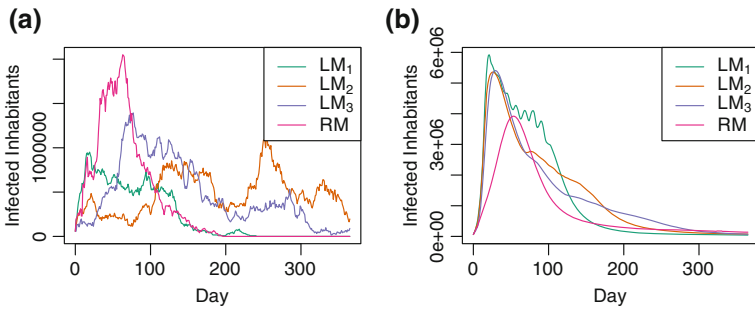
(a) **(b)**

Fig. 7 Comparison of different countermeasure models on a large scale, i.e., each value represents the situation for the corresponding countermeasure model in all cities combined. **a** 10.000 agents. **b** 100.000.000 agents

the ratio between the amount of currently infected individuals and the population size of the city. Figure 7 compares all models to each other on the national level.

Although the level based approach is completely different compared to the ratio based approach, the achieved results are similar. However, the overall increase of α and κ by the ratio based approach may be noticeably higher, especially if a large number of agents is used. That is, while all LM-models use $\alpha \leq 3.3$, the RM-model goes above 4. This implies that the LM-models are more cost efficient, since both α and κ are kept lower and therefore less effort is needed to achieve and maintain said values.

Additionally, to be able to compare our results to the findings in [18], we examine the following city in more detail: Berlin (BER) with a population size of ≈ 3.5 million. Figure 6 represents a composition of some interesting results. Note

that the green curve in these figures represents the countermeasure level for the *LM* models at the corresponding time, and indicates the number of times z_1, z_2 have increased α, κ in the *RM* model. Recall that, in the *RM*-model level j implies $\alpha_i = \alpha_0 + \sum_{k=1}^{j} z_1$ and $\kappa_i = \kappa_0 + \sum_{k=1}^{j} z_2$ for a step i.

Our results confirm the impact of different countermeasures observed in the real world [18]. Compared to Fig. 5, our simulations show a similar behavior (i.e., more than one peak during the epidemic). It is easy to see that the countermeasures presented in [18] directly influence the course of the epidemic. The same property can be observed in our results (cf. Fig. 5). One can see that depending on the countermeasure level (indicated by the activated/used level), the number of infections increases or decreases. Note that although our figures show the number of infected individuals and not the death rate as shown in Fig. 5, a comparison is still possible, since this deviation can be normalized using a scaling factor.

Furthermore, we observe that small adjustments of the two parameters α and κ entail a significant impact on the number of overall infections. Among others we already proved that if the power law exponent (and κ as well) is assumed to be some large constant, then even a very aggressive epidemic with $\gamma = 100\,\%$ will affect no more than a polylogarithmic number of the population. Our findings now back up these observations.

In conclusion, we showed the impact of different countermeasures on the behavior of a population w.r.t. our model. Although some complexity of the real world is lacking, the similarities to real-world observations are still present. Starting with settings for the environment, and therefore implicitly the individuals' behavior, based on real-world observations, relatively low level countermeasures were sufficient to embank or at least significantly suppress an outbreak. Essentially the same properties were already observed in reality (cf. [18]). This underlines the importance of behavioral and environmental models based on power law distributions.

4 Conclusions

Agent based simulators offer various possibilities to perform very detailed experiments. However, the parameters used in these experiments highly influence the results one might obtain. As we have seen, even the number of agents has a significant impact on the quality of the results. This includes the reliability of different simulation runs with an identical parameter setting. By using the right parameter settings and a proper number of agents, it is possible to approximate the course of a disease as observed in the real world. Furthermore, our experiments indicate that the algorithmic framework presented in this paper is able to describe, to some extent, the impact of certain non-pharmaceutical countermeasures on the behavior of an epidemic.

References

 1. Adamic LA, Huberman BA (2000) Power-law distribution of the world wide web. Science 287(5461):2115
 2. Ajelli M, Goncalves B, Balcan D, Colizza V, Hu H, Ramasco J, Merler S, Vespignani A (2010). Comparing large-scale computational approaches to epidemic modeling: agent-based versus structured metapopulation models. BMC Infect Dis 10:190
 3. Amaral LA, Scala A, Barthelemy M, Stanley HE (2000) Classes of small-world networks. PNAS 97(21):11149–11152
 4. Arbeitsgemeinschaft Influenza (2011). Bericht zur Epidemiologie der Influenza in Deutschland Saison 2010/2011
 5. Balcan D, Hu H, Goncalves B, Bajardi P, Poletto C, Ramasco JJ, Paolotti D, Perra N, Tizzoni M, den Broeck WV, Colizza V, Vespignani A (2009) Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. BMC Med 7:45
 6. Borgs C, Chayes J, Ganesh A, Saberi A (2010) How to distribute antidote to control epidemics. Random Struct Algorithms 37:204–222
 7. Chowell G, Hyman JM, Eubank S, Castillo-Chavez C (2003) Scaling laws for the movement of people between locations in a large city. Phys Rev E 68(6):661021–661027
 8. Elsässer R, Ogierman A (2012) The impact of the power law exponent on the behavior of a dynamic epidemic type process. In: SPAA'12, pp 131–139
 9. Elsässer R, Ogierman A, Meier M (2013) Agent based simulations of epidemics on a large scale. In: SIMULTECH'13, pp 263–274
10. Eubank S, Guclu H, Kumar V, Marathe M, Srinivasan A, Toroczkai Z, Wang N (2004) Modelling disease outbreaks in realistic urban social networks. Nature 429(6988):180–184
11. Faloutsos M, Faloutsos P, Faloutsos C (1999) On power-law relationships of the internet topology. In: SIGCOMM'99, pp 251–262
12. Germann TC, Kadau K, Longini IM, Macken CA (2006) Mitigation strategies for pandemic influenza in the United States. PNAS, 103(15):5935–5940
13. Grassberger P (1983) On the critical behavior of the general epidemic process and dynamical percolation. Math Biosci 63(2):157–172
14. Hethcote HW (2000) The mathematics of infectious diseases. SIAM Rev 42(4):599–653
15. Jaffry SW, Treur J (2008) Agent-based and population-based simulation: a comparative case study for epidemics. In: Louca LS, Chrysanthou Y, Oplatkova Z, Al-Begain K (eds) ECMS'08, pp 123–130
16. Lee BY, Bedford VL, Roberts MS, Carley KM (2008) Virtual epidemic in a virtual city: simulating the spread of influenza in a us metropolitan area. Transl Res 151(6):275–287
17. Lee BY, Brown ST, Cooley PC, Zimmerman RK, Wheaton WD, Zimmer SM, Grefenstette JJ, Assi T-M, Furphy TJ, Wagener DK, Burke DS (2010) A computer simulation of employee vaccination to mitigate an influenza epidemic. Am J Prev Med 38(3):247–257
18. Markel H, Lipman HB, Navarro JA, Sloan A, Michalsen JR, Stern AM, Cetron MS (2007) Nonpharmaceutical interventions implemented by US cities during the 1918–1919 Influenza Pandemic. JAMA 298(6):644–654
19. Newman MEJ (2002) Spread of epidemic disease on networks. Phys Rev E 66(1):016128
20. Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45 (2):167–256
21. Ripeanu M, Foster I, Iamnitchi A (2002) Mapping the Gnutella network: properties of large-scale peer-to-peer systems and implications for system. IEEE Internet Comput J 6(1):50–57
22. Robert Koch Institute (2012). SurvStat@RKI. A web-based solution to query surveillance data in Germany