A case study of infectious disease epidemics using Agent-Based modeling

Sahand Hariri[a,*], Jamshid Ghaboussi[b], Michael F. Insana[c]

[a] *Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, 1206 W. Green St., Urbana, IL 61801, USA*
[b] *Civil and Environmental Engineering, University of Illinois at Urbana-Champaign Urbana, IL 61801, USA*
[c] *Bioengineering and Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

Abstract

We present an agent-based model for simulating the spread of an infectious disease through a population. The simulation allows for great amount of control over custom variables present in the system, such as the spatial distribution of people and their motion through space and time. We show the simulation is capable of reproducing the results of existing models in the literature such as SIR when stochastic effects of the model are ignored and the agents posses identical properties so they behave identically. However, the real usefulness of the model is in simulating complexity that is not possible to capture using conventional methods. We show our model produces a power-law probability distribution of epidemic sizes that is consistent with the distribution of infectious-disease epidemic sizes recorded at Faroe Island over a 100 year period. In an analysis similar to a parameter sweep, we show that starting with the simple case (approximating SIR), we can approach power law behavior as we add variability to the system.

Keywords: Self organized criticality, SIR Models, Role of Variability, Powerlaw Distributions.

1. Introduction

Many natural phenomena have been connected to self-organized criticality. The presence of power law distributions in natural processes is wide spread <3>. Various models have been developed in order to gain some understanding of this phenomenon. Ideas about characterization of self-organized critical systems were first put forth by Bak et. al. <1>. These ideas were proposed as an approach to describing dynamic phenomena such as sand pile avalanches, forest fires, earthquakes and biological processes. Rhodes et. al. extend this approach to study spread of communicable diseases in populations <9>. They draw an analogy between type III epidemics[1] and the forest fire model with a lightening event. Using this analogy, they argue that spread of disease also operates under a self-organized critical state. They show that the probability of occurrence of an epidemic larger than a certain size, in about 100 years worth of data from Faroe Islands, shows power law behavior. In this paper we present a multi-scale agent-based model that is capable of reproducing the power law behavior of this system. The model simulates the spread of an infectious disease in a population. It allows the modeler to control such variables as the spatial distribution of people and their motion through space and time. It is an individual based simulation where each "Person" in the simulation has a unique set of characteristics which determine their activities, responses and interactions with the environment and other individuals.

*Corresponding author: hariria2@illinois.edu
[1] These are epidemics where the source is a foreign agent.

The problem of spread of infectious disease has long been studied in many different contexts, using many different methods. The most basic and fundamental mathematical models for studying epidemiological problems are perhaps the compartmental models, first introduced by McKendrick and Kermack <2>, known as SIR, SEIR, SIS, SEIS, and so on. These models provide the foundation for mathematical epidemiology. In each of these models, the population is divided into different compartments: Susceptible (S), Exposed (E), Infected (I), and Recovered (R). The rates of change of the compartments, e.g. $\dot{S} = \frac{dS}{dt}$, are governed by Lotka-Voltera like equations,

$$\begin{aligned} \dot{S} &= -\beta SI \\ \dot{I} &= \beta IS - \gamma I \\ \dot{R} &= \gamma I \end{aligned} \tag{1}$$

where β and γ are positive constants.

Even though these models capture the qualitative behavior of an epidemic and have been very useful in helping to develop some understanding of these phenomena, they are too simplified to address the complexities involved in a real life situation. The shortcomings of these models can essentially be broken down to two categories: 1) They assume a limited number of homogeneous populations (groups with identical properties defined by model variables S, I, and R.) Homogeneity restricts the range of hierarchical variability that generates the emergent properties characteristic of complex systems. 2) The populations mix homogeneously at all times. Basically no spatial effects are taken into account, so there are no subpopulations with distinct properties.

Various approaches have been proposed to remedy these two shortcomings. Stochastic methods and individual based models are examples of attempts to represent the inhomogeneity present among realistic populations. <8; 7; 4; 6; 13; 15>. Furthermore, spatial effects and the contact pattern of people are very important in the spread of disease <8; 4>, and that almost all real life scenarios operate far from the states assumed in models describing homogeneously mixed population dynamics. Cellular automata, network or graph approaches are examples of models that address heterogeneity in populations.

The fact that there are so many different models, and more continue to be developed, is a testament to the complexity and importance of this problem. Short of simulating the full fledged system, we run into many limitations. Agent-based models are powerful computational tools for tackling this problem. They provide a means of adding important but non-parametric aspects into the model to add details of real-life responses. Greater complexity can be captured using agent-based models by defining simple rules for each agent and allowing large-scale dynamics to develop based on interactions of these agents.

This paper is divided up into five sections. In the second section, we present our model, and the various components that comprise it. Issues such as interaction among individuals and the spatial variations that those interactions generate are addressed in this section. Section 3 considers the long-term effects of infectious disease epidemics and discusses the results presented by Rhodes, et. al. <9>. We also discuss how our model is adapted to a long-time simulation where epidemics reoccur. In the fourth section we present modeling results. We first show that our model can be reduced to capture results of simpler models such as SIR. We show that the contact network that appears naturally as a product of running the simulation agrees with literature studies focused on contact patterns in communicable diseases. We then compute the probability of occurrence of epidemics larger than a given size, and show that it follows a power law, just as in <9>. In the last section, we offer some concluding remarks. The methodology and details of implementation are provided as supplementary material.

2. Model

We develop an agent-based model (ABM) that can capture the behavior of measured systems at several scales. The context, agents, and the function of the agents are described below.

2.1. Agents (People)

Each person is represented as an agent in the simulation. Agents are distinct because of their assigned properties, all of which can be independently assigned. Some of these properties are, age, gender, susceptibility level to catching the virus, how fast the virus can replicate within each host (person), how fast the individual can recover, and so on.

Along with these properties, agents are assigned actions that govern their behavior in the simulation environment, and how they react to certain events taking place among them. These functions include how they move, how their normal behavior changes in response to infection, the probability of becoming infected after contact with an infected person, and so on.

2.2. Domain (Location)

We need a spatial context for the simulation. We define island locations that describe this spatial information. Each individual has a location at any given time. This is the feature of the model that allows us to capture the contribution of spatial variation in agent properties and behavior in the problem. One or more regional environments can be defined for a given simulation each of which carry their own set of properties such as size, population, and type of place. The type of location might then be used in describing the motion of the individual. Individuals can enter or leave given location based on the rules that govern their motion.

2.3. Motion of Agents

The movement of agents through time and space is perhaps the most complex portion of the model. The agents can be assigned to move with purpose or randomly. They can be placed at different locations based on time, and their state of health. Once in a location, they can follow predefined deterministic paths or move about in a random walk. The nature of the motion in turn affects the contact network created among individuals as a result, which in turn affects the transmission of the disease, thus further affecting the agent movements.

For example, if one sets the model to simulate the daily activity of an individual, then the locations that individual might belong to are home and school or work. The daily motion of the individual can then be defined to place him/her in a specific location at a given time. If the individual is sick, then they might choose to stay home for the duration of the illness. Staying home limits the chances of spreading the disease and speeds up the recovery process.

In most cases in our simulations, once the individuals arrive at a new location, we set them to move about in a random walk within that location. There is also a preferential drift based on age groups, which is discussed later. In addition to local movements, there may be occasional long-range motions to more realistically approximate real-life population dynamics. As a result the network may be more intricate than just simple diffusion as described in <14>.

2.4. Disease Spread Within Individuals

Once an individual is exposed, there is a set of model equations that governs the replication of virus within the host. How susceptible the individual is, how quickly they show symptoms, and how quickly they recover are all processes that the equations determine. We use a basic model for the within-host dynamics of viral infection presented in <11>. The equations are as follows:

$$\begin{aligned} \dot{T} &= -\beta TV \\ \dot{I} &= \beta TV - \delta I \\ \dot{V} &= PI - CV \end{aligned} \tag{2}$$

In these equations T is the number of susceptible cells in an individual, I is the number of the infected cells, and V is the viral load or the total number of virions. The parameters β, δ, P, and C are, respectively, the

rate at which susceptible cells become infected, infected cells recover, the virus replicates, and finally the rate at which the virus dies.

There are two separate time scales in the simulation. There is the overall time for the system t, which we call population time, and there is the time scale, $\tau = t/\epsilon$ $(0 < \epsilon << 1)$, on which individual dynamics, i.e., the system in equation (2) operate, which we call agent time. As the simulation steps forward in time, current values of T, I, and V are fed in as the initial values for computing values at the next time interval. The equations are then integrated forward on time scale τ until the next population time step. The new values of T, I, and V are then used to apply update rules on the population time scale, the results of which are taken to be the initial conditions for the next simulation on the agent time scale.

2.5. Disease Transmission Between Individuals

The agents are free to move about in a given location. Initially, every agent is considered susceptible to the disease under study. If a susceptible agent comes in contact with an infected agent, then there is a potential for virus transmission. In what follows, we describe what it means for two agents to come in contact, and how the virus transmission occurs.

There are two components that need to be considered here. The first case is connecting the development of the disease within an individual (host) to the large-scale population model. The second is the model that governs the transmission of the virus from one individual to another (between-host transmission). These two steps are not independent, they can affect each other in a feedback loop.

2.5.1. Coupling within-host model with population-level model

Most studies consider the dynamics at the scale of population decoupled from the dynamics within each host <5>. This simplifies the model significantly and allows for consideration of each model separately without worrying about the effects of processes happening at different scales. However, since we are building our model from the ground up, we need to nest the within-host model within our population level simulation. This is the mechanism through which the virus can be transmitted from one person to another, and so it is an important aspect of our model. Moreover, Feng et. al. <5> showed that by considering these two processes as coupled, more complex behavior can arise. When the two processes are considered decoupled, the intrinsic assumption is that every person carries a constant viral load and is equally infective. We therefore need to modify equation (2) to allow for this phenomenon.

We begin by noticing that an individual agent can be under the influence of more than one agent at a time. So for a susceptible agent, the amount of viral load that is picked up from the environment depends on the number of local agents that carry the virus. Note that an agent can carry the virus without showing any symptoms of the disease. It follows that the total viral load picked up by an agent at a given time is proportional to the aggregate of viral loads of the neighbors. Equation (3) shows the modified version of equation (2).

$$\begin{aligned} \dot{T}_i &= -\beta_i T_i V_i \\ \dot{I}_i &= \beta_i T V_i - \delta_i I_i \\ \dot{V}_i &= P_i I_i - C_i V_i + B_i Z \end{aligned} \tag{3}$$

where the term Z is the total viral load accumulated from all the neighboring agents. The constant B_i represents the receptivity of individuals to the viral load. If agent i has N neighbors, then:

$$Z = \sum_{j \neq i}^{N} p_j V_j \tag{4}$$

V_j is the viral load of each neighbor, and p_j is the associated probability of that agent transmitting the virus. Equations (3) and (4) provide coupling between agents for population level dynamics.

The addition of the extra term in equation (3) allows for more complex behavior than the standard model. In this case the level of infection of an individual can fluctuate based on his or her changing environment. This feature would not be captured by the model as presented in equation (2). Now we shift our attention to the between-host virus transmission, and calculation of p_j.

2.5.2. Between-host transmission of virus

Intuitive understanding of viral disease transmission suggests the distance between an infected individual and a susceptible person plays a central role. At each time step, we compute the distance between any two individuals, which determines the value of p_j in equation (4). Standard assumptions for determining p_j from agent proximity include inverse or exponential distance weighting for contact probability. But in <14>, it was shown this model performed poorly for measles data. Instead, a gravity-like model was used that had the form $N_k N_j / d^\alpha$. In this particular study, the transmission is considered between populations of people, and N_k and N_j are two different populations and d is the distance between the two. In our case, N_k and N_j are replaced with T_j and T_k, the number of susceptible cells in each host. The value of $\alpha = 1.5$ was used as found in <14>.

Once the level of viral load within an individual reaches a predefined threshold, equations (3) become active and govern the infection process within the host. Conversely, when the value of viral load drops below the threshold, the effect it has on the individual is considered to be zero <12>.

3. Long-time simulation and self-organized criticality

3.1. The experiment

In their paper <9>, Rhodes et. al. draw an analogy between the problem of the spread of infectious diseases and forest fire models of Bak et al. <1>. They show that given a relatively isolated population, one can consider the introduction of a new infection as a "lightning" phenomenon, and the spread of infection is analogous to fire spreading to neighboring trees. Birth in the population is analogous to growth of new trees in place of burned ones. Rhodes <9> showed that when timing of the three events, birth rate, epidemic duration, and period of viral reintroduction are different, a power law emerges in the probability of occurrence of an epidemic of a certain size. The presence of a power law response indicates the existence of a self-organized critical state.

They studied recorded data for the Faroe Islands over a century, from 1870's to 1970's. Three viruses were considered: Measles, Mumps, and Whooping cough. For each epidemic, the investigators introduced a measure for the epidemic size given by,

$$s = \sum_{M_{start}}^{M_{end}} C(M), \tag{5}$$

where $C(M)$ is the number of recorded cases of a disease in month M. With $n(s)$ representing the number of epidemics of size s, the probability of an epidemic of greater than size s' is computed as,

$$P(s) = \text{prob}\{s \geq s'\} = \frac{\sum_{s'=s}^{\infty} s' n(s')}{\sum_{s'=1}^{\infty} s' n(s')}. \tag{6}$$

For each of the infections considered, this probability was shown to follow a power law. An attempt to reproduce this behavior using SEIR (Susceptible, Exposed, Infected, Recovered) model was unsuccessful.

3.2. Simulating long-time behavior

We set out to use our model to simulate the system described in <9>. We let the simulation run for 36500 time steps, where each time step represents one day, in order to simulate a 100 year run. We then compute the sizes of the epidemics that result as prescribed by equation (5), and then compute the probability via equation (6). There are several considerations that must be decided when modifying the model to simulate long-time behavior.

3.2.1. Birth and death of agents

Over many years, the birth and death rates in a population become important factors in modeling the dynamics of disease propagation and epidemics. Even though we model the agents as permanently immune to a second infection after recovery, the population as a whole is susceptible through introduction of new people (birth). Therefore, an epidemic can occur in a population repeatedly over many years. It is then important to incorporate birth and death into our model for long-time simulations. There are two forms of death that can occur: death due to old age and death caused by an infection.

We define a threshold for the number of infected cells a person can bear in equation (2). This threshold determines whether the infection is fatal or not. Notice that the initial susceptibility of the person (initial condition for the number of susceptible cells) may be smaller than this threshold. In that case, the person may become infected, but the infection is not fatal. Death by natural causes simply occurs when individuals reach their life expectancy.

3.2.2. Introduction of carrier agents

After an epidemic has occurred, it is likely that the virus will no longer exist in the population. Over time, the population becomes susceptible again through new births. However, in order to trigger another epidemic, an outside source of the virus is necessary. In our model, we capture this by introducing new carrier agents, who may or may not have symptoms of infection, but carry a certain amount of viral load. The frequency at which this happens is much smaller compared to the rate of birth, and inverse of a typical epidemic period.

3.2.3. Motion of the agents through space

The geometry of the simulation is set up to mimic that of the Faroe islands. That is, we create locations to represent each of the major islands as closely as possible, and distribute the agents in proportion to the demographics of Faroe islands. Figure 1 shows a screen shot of the simulation environment. The geographic and population data for the simulation was obtained from GADM [2].

The motion of agents is very important in determining the social contact network that is formed among people. There is evidence that the social network formed among people of different age groups has a non-negligible effect in spreading the infection <10>. We therefore consider the effect of agent age in deciding how different social groups are formed during the course of the simulation. Each agent is set to move in a random walk with a drift term. The random element of the motion represents the individual's decision making process independent of social influence. The magnitude of the random step is affected by the health level of the individual. Infected individuals are less mobile in an effort to minimize their contact with others. The drift term is the component of the motion determined by age group interactions. In other words, the individual agent reacts to the age distribution of its neighbors and purposefully moves in the direction of agents who are in the same age group. It is computed for each individual at each time step. Each person is assigned, as a property, an attractive (or repulsive) force with respect to every other individual in their vicinity. The vector sum of these forces determines the total drift the person experiences at each time step. In addition to a social attractive force, each individual is assigned a "sociability" constant. This value is independent of all the other factors and represents how introverted or extroverted a person might be. The net result of adding these drift terms to the motion of each individual is formation and dissolution of groups over time in the population. This of course has great implications on the propagation of disease.

4. Results and Discussion

4.1. Reproducing SIR

The agent-based model is expected to be a more general model than the SIR. That is, if we incorporate all the assumptions present in the SIR model into our agent-based model, we should be able to reproduce the

[2] www.gadm.org - "GADM is a spatial database of the location of the world's administrative areas (or administrative boundaries) for use in GIS and similar software."

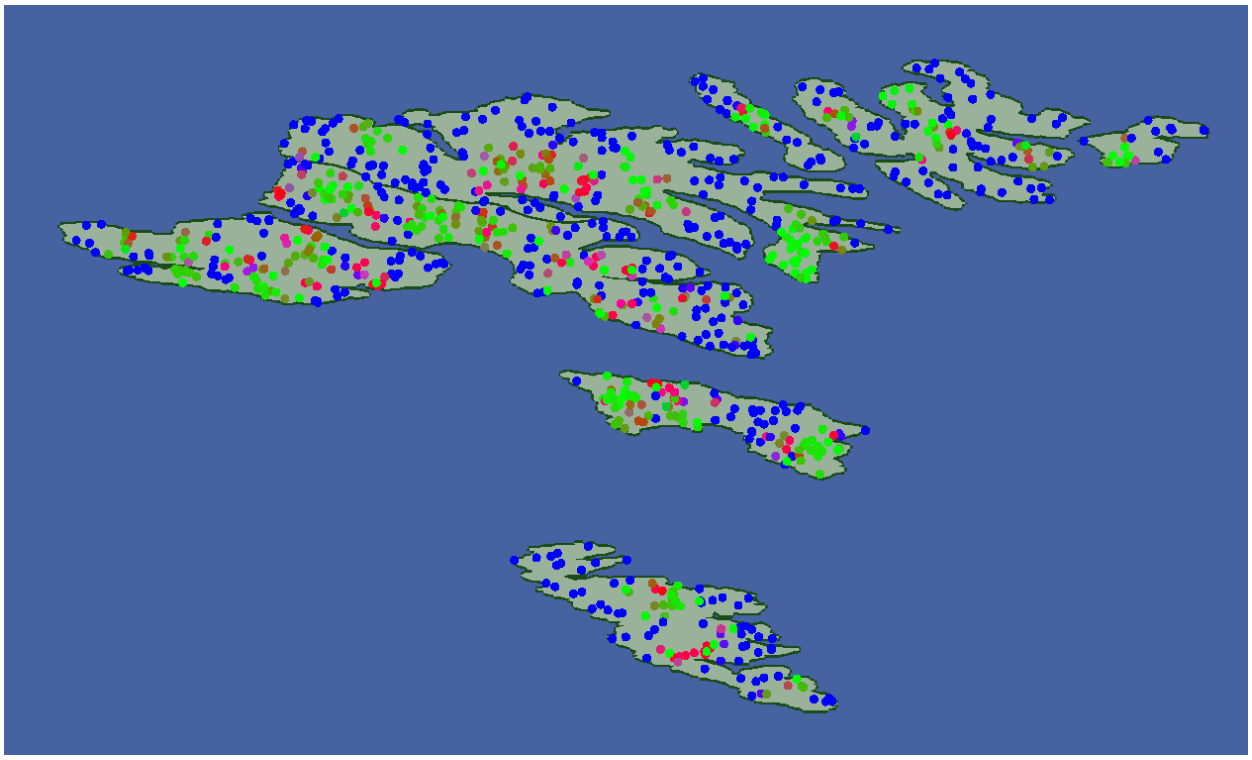
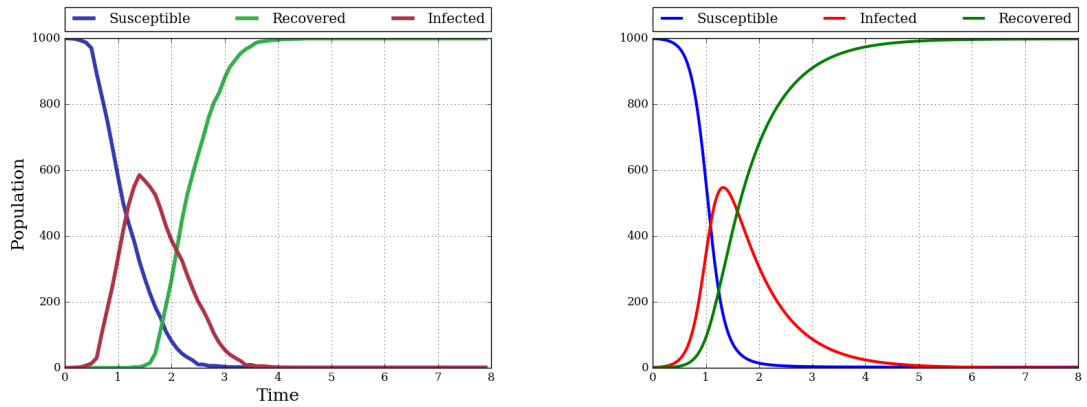


Figure 1: Visualization of the simulation environment. Blue, red, and green dots represent susceptible, infected, and recovered agents, respectively.

behavior of the SIR model, at least qualitatively. The SIR model assumes people have identical properties. It also assumes a homogeneous mixing of the population. We can certainly make all agent properties identical. To achieve homogeneous mixing, we create only one location and restrict people's motion to a random walk with large step sizes. Figure 2a shows a population plot using a sample simulation of our agent-based model. As we can see, the agent-based simulation is able to qualitatively capture the trajectory generated by the SIR model (equation (1)) shown in 2b.



(a) ABM simulation.

(b) SIR model

Figure 2: The signature sigmoid behavior of the SIR is obtained using the agent based model.

The SIR model depends on two parameters β and γ. On the other hand, there are many more parameters in the agent-based model that can affect the behavior of the system. Many of these parameters add complexity to the system by means of spatial effects. In what follows, we will examine these effects.

4.2. Spatial effects and contact network

Spatial effects play a crucial role in the mixing of population. One aspect important in determining the motion of agents in the simulation domain is the age of agents in their contact network. It is known that the effects of age in the formation of the contact networks and subsequently on the spread of communicable disease is not negligible <10>. Here we present the age contact network produced by the simulation when the motions of individuals are simulated as described in 3.2.3.

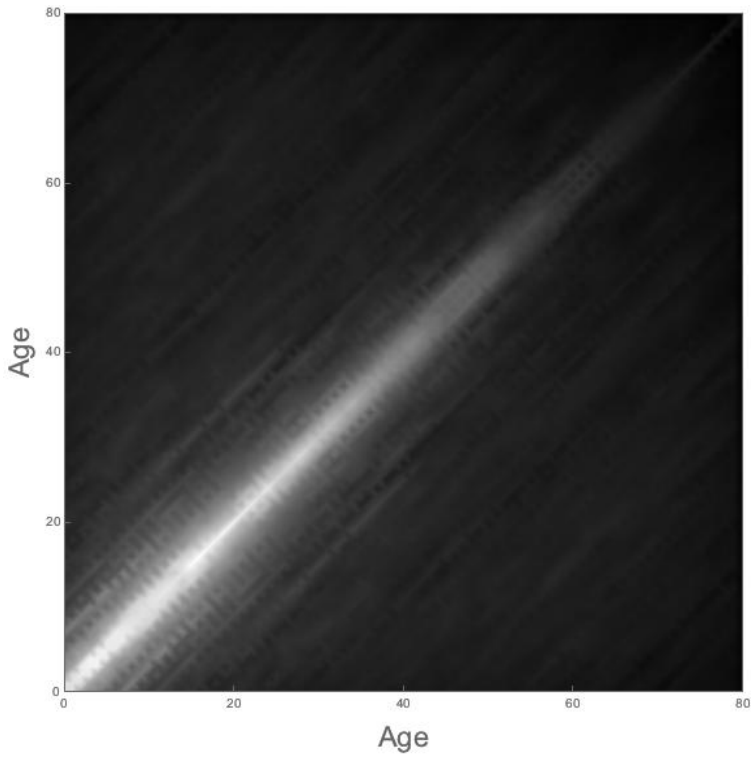


Figure 3: Age Contact Network. Black indicates no contact, and gray scale leading to white indicates an increasing degree of agent association.

We introduce five distinct age groups, child, teenager, young adult, adult, and senior, to which an individual can belong. A strong preference is given to the interaction of people with their own age groups. This was found to be the case in <10> and is clearly reflected in model data connections summarized by figure 3. We can see most of the interactions take place between individuals that belong to the same age group.

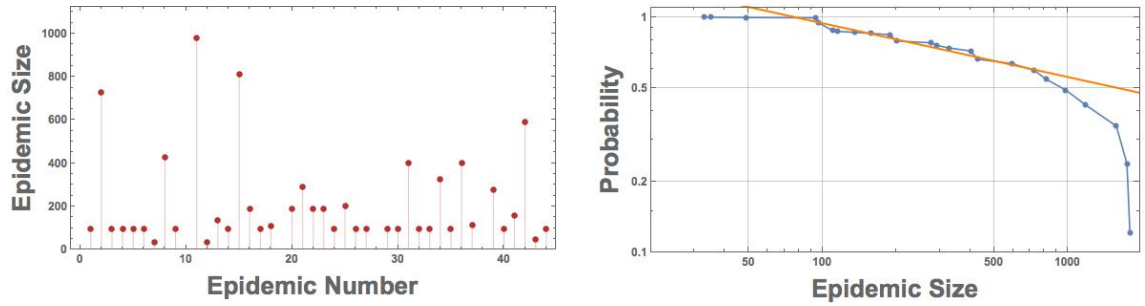
Rules that determine the contact network of a population form the interactions that lead to the formation of sub-populations with characteristic properties. We believe clustering of populations is one of the major factors in contributing to the power law behavior observed in real-world data observed in <9>, as well as our simulations. We discuss this more in the next section.

4.3. Power law

We run our simulation as described above. The simulation takes into account the birth and death of the agents. After the transients in the system have died out, the population settles near 1000 agents. We

consider the epidemics occurring in the span of a 100 years of simulation time after the transients. Using equations (5) and (6), we computed the probability of occurrence of epidemics of a given sizes.

We use C++ to perform the simulation. Depending on the parameters used in the simulation, each run, as described above, takes 1 to 5 hours on a late 2011 model macbook pro with quad core i7 CPU.



(a) Population of infected people over time.

(b) Distribution of epidemic sizes.

Figure 4: Epidemic time-series data, and a log-log probability distribution of epidemic sizes.

Our simulated data results in a power law, figure 4b, just as found in <9>. This indicates that the proposed model is capable of capturing a wide range of behaviors, from the simple SIR when it is fully simplified, to the results of a study using real-world data. As noted in <9>, the traditional SEIR model proved incapable of reproducing these results. It is then natural to ask the question: what is the minimal level of complexity added to our model that will result in a power law? In order to approach this question, we perform a parameter sweep and consider the correlation coefficient obtained by computing linear regression of the probability data. However, picking a meaningful parameter to perform this parameter sweep is not simple. There is no single parameter that we can change to reflect the added complexity to the system. At one extreme, we start with the simulation that resulted in reproducing the SIR model. From this point on, we add more complexity to the model through six stages described in table 1. Figure 5 displays the correlation coefficient obtained from a linear fit of the log-log probability distribution. At each stage, we perform four independent simulations and consider the average result.

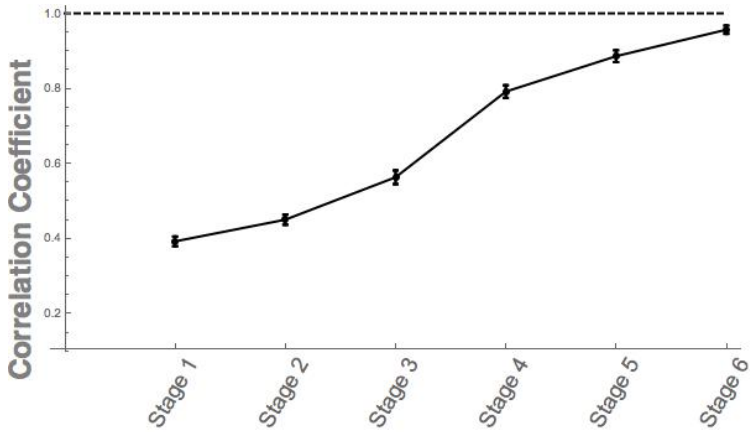


Figure 5: Correlation coefficients obtained from performing a linear regression on the log probability distribution at each stage. Each stage shown on the horizontal axes represents a cumulative addition of complexity to our simulation. At each step we perform four simulations. The error bars represent standard deviation of these results.

The correlation coefficient approaches 1 as we advance through the stages, showing how additional model complexity modifies the system to respond in a manner best described by a power law. We believe the mechanism that drives the system towards a power law also builds additional rules for hierarchical

diversity that add degrees of freedom to the population. Agent interactions diversify to create distinct subpopulations. This in turn causes the resulting epidemics to be contained mostly within subpopulations. Rarely do we experience an epidemic that crosses the subpopulation borders to infect large numbers of people. For example when the geographic domain is partitioned, we have effectively limited communication among the agents and created subpopulations with their own properties. Similarly, when we introduce preferential motion of agents based on their age groups, we are creating subpopulations that may not always interact with other subpopulations during an epidemic event. In this way, diversity in agent properties constrains the size of infectious outbreaks, which adds stability to the population in terms of population size.

x-axis on fig. 5	Description
Stage 1	Single location; Identical properties for individuals; No variation in rate of birth or motion; Large contact neighborhood around individuals.
Stage 2	Single location; Identical properties; Variation in motion and rate of birth; Large contact neighborhood around individuals.
Stage 3	Single location; Varied properties; Variation in motion and rate of birth; Large contact neighborhood around individuals.
Stage 4	Multiple locations; Varied properties; Variation in motion and rate of birth; Large contact neighborhood around individuals.
Stage 5	Multiple locations; Varied properties; Variation in motion and rate of birth; Small contact neighborhood around individuals.
Stage 6	Multiple locations; Varied properties; Variation in motion and rate of birth; Small contact neighborhood around individuals; Age group clustering; Restricted motion for infected people.

Table 1: The progression of stages that describe agent properties and rules of interaction that move the modeled system from a homogeneous, SIR model probability distribution to a complex power law probability distribution.

It is important to note that the specific order in which we add complexity to our simulation is irrelevant. Figure 5 demonstrates one path of increasing complexity to the system. However, by re-ordering the steps outlined in table 1, we get a result similar to figure 5.

Figure 6 displays details of simulation results for the six stages. From top to bottom, we move from stage 1 to stage 6, by adding property diversity for the population. The first column on the left represents the size of each epidemic that occurs in order of occurrence through time. These epidemic sizes were computed using equation (5). The second column shows the distribution of these epidemic sizes for a given stage. Finally, the last column shows the probability distribution computed from equation (6) that is plotted on log-log axes, along with best linear fit of the data. The correlation coefficient of this line is the value plotted in figure 5.

The homogeneous population in stage 1 means that epidemics as large as 500 agents have equal probability. At stage 1, epidemic size is primarily controlled by the size of the susceptible sub-population at the time of the next outbreak, which is why epidemic sizes less than 500 agents are equally likely.

Adding a small amount of population variability at stage 2, we see the equal probability range is reduced significantly to about 200, and as a result the correlation coefficient shown in figure 5 increases moving from stage 1 to stage 2. As additional variability in agent-population properties is added during stages 3-6, there is a continuing trend in the probability plot toward a power law.

Note that in the later stages, epidemics of very small sizes still maintain equal probability close to 1.

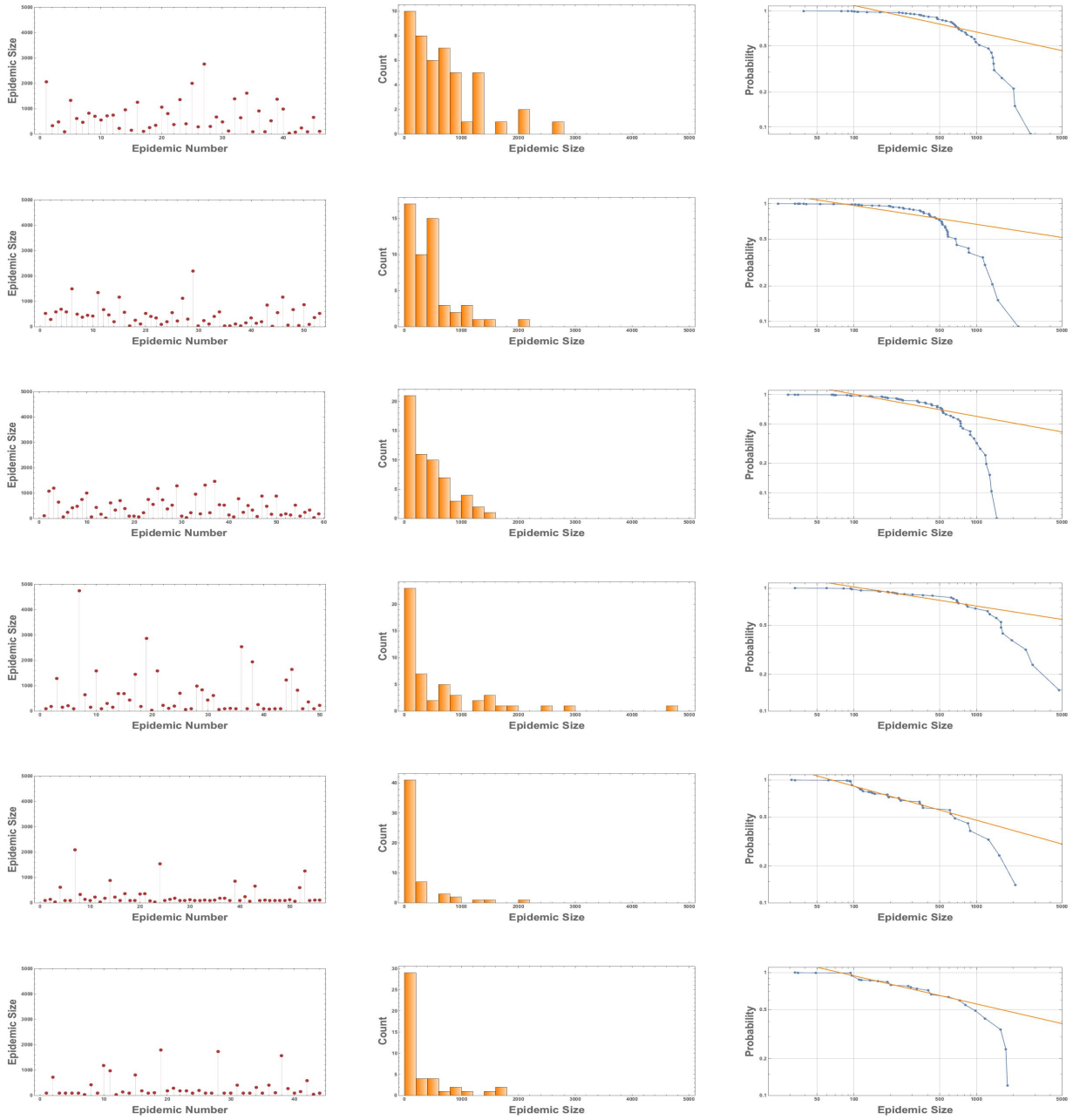


Figure 6: Illustrating the effects that increasing property variability has on model complexity as measured by the probability of observing specific epidemic size. The top row are examples of stage 1 results and the bottom row are from stage 6. The left column shows epidemic sizes in order of occurrence. The middle column is the distribution of the sizes of these epidemics. The right most column represents the probability of occurrence of epidemics on a log scale. From top to bottom we see that as we increase variability, the probability approaches a power law.

These correspond to very few agents (even one or two) becoming infected, which is always a possibility in the simulation, hence the probability is always 1. One can consider theses events not as epidemics, but as random cases of infection in the population. However, after that when actual epidemics start forming, we can clearly see the power law behavior when there is a high degree of variability. for these cases, the epidemic

size range for which the power law fits well increases in range to 100-1000 in stage 6 as the correlation coefficient steadily increases toward one. This behavior is an indication that additional property variability is generating subpopulations of different sizes that contribute additional degrees of freedom. We referred to this effect in the introduction as hierarchical variability that is responsible for the emergent properties characteristic of system that is growing in complexity.

5. Concluding Remarks

We have introduced a new agent-based approach to modeling infectious disease in populations. The model allows consideration of various details down to the individual level in the population. Many characteristics of the system such as variability in behavior among individuals, birth and death, age group clustering, spatial variations, changes in behavior of individuals in response to environmental effects and others are explored. We use this model to reproduce qualitative behavior of existing models such as SIR. Additionally, we show that our simulation can qualitatively reproduce results of a study of the problem, <9>, wherein the model was analogous to the forest fire model. Infection growth within individuals is modeled on a separate time scale than the overall simulation time scale. The embedding of this model in the overall simulation, as well as the modifications necessary to include environmental and population effects on each individual are discussed.

We compute the probability of occurrence of epidemics of a certain size as described in <9>, and show that this results in a power law as desired. Starting with a simple model, we show the system approaches self-organized criticality as complexity increases by adding diversity to agent properties and their rules of interaction. At each stage added variability increases spatial clustering in the population. This in turn decreases the likelihood of large epidemics sweeping through the entire population, giving rise to power law behavior. We surmise that the emergence of a power law depends on the presence of variability in the system and the mechanism by which this happens is through introduction of hierarchical subpopulations with distinct properties giving rise to multiple degrees of freedom in the system.

[1] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality. *Physical review A*, 38(1):364, 1988.
[2] Fred Brauer. The Kermack–McKendrick epidemic model revisited. *Mathematical Biosciences*, 198(2):119–131, 2005.
[3] James H Brown, Vijay K Gupta, Bai-Lian Li, Bruce T Milne, Carla Restrepo, and Geoffrey B West. The fractal nature of nature: power laws, ecological complexity and biodiversity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 357(1421):619–626, 2002.
[4] Charly Favier, Delphine Schmit, Christine DM Müller-Graf, Bernard Cazelles, Nicolas Degallier, Bernard Mondet, and Marc A Dubois. Influence of spatial heterogeneity on an emerging infectious disease: the case of dengue epidemics. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1568):1171–1177, 2005.
[5] Zhilan Feng, Jorge Velasco-Hernandez, Brenda Tapia-Santos, and Maria Conceição A Leite. A model for coupling within-host and between-host dynamics in an infectious disease. *Nonlinear Dynamics*, 68(3):401–411, 2012.
[6] Henryk Fuks and Anna T Lawniczak. Individual-based lattice model for spatial spread of epidemics. *Discrete Dynamics in Nature and Society*, 6(3):191–200, 2001.
[7] Matt J Keeling and Ken TD Eames. Networks and epidemic models. *Journal of the Royal Society Interface*, 2(4):295–307, 2005.
[8] Matthew J Keeling. The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1421):859–867, 1999.
[9] CJ Rhodes, HJ Jensen, and RM Anderson. On the critical behaviour of simple epidemics. *Proceedings of the Royal Society of London B: Biological Sciences*, 264(1388):1639–1646, 1997.
[10] Pejman Rohani, Xue Zhong, and Aaron A King. Contact network structure explains the changing epidemiology of pertussis. *Science*, 330(6006):982–985, 2010.
[11] Hal L Smith and Patrick De Leenheer. Virus dynamics: a global analysis. *SIAM Journal on Applied Mathematics*, 63(4):1313–1327, 2003.
[12] Shelby H Steinmeyer, Claus O Wilke, and Kim M Pepin. Methods of modelling viral disease dynamics across the within- and between-host scales: the impact of virus dose on host population immunity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1548):1931–1941, 2010.
[13] S Hoya White, A Martín del Rey, and G Rodríguez Sánchez. Modeling epidemics using cellular automata. *Applied Mathematics and Computation*, 186(1):193–202, 2007.
[14] Yingcun Xia, Ottar N Bjørnstad, and Bryan T Grenfell. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *The American Naturalist*, 164(2):267–281, 2004.
[15] Mina Youssef and Caterina Scoglio. An individual-based approach to SIR epidemics in contact networks. *Journal of theoretical biology*, 283(1):136–144, 2011.