UNIVERSITY OF AMSTERDAM

PROJECT: STOCHASTIC AND SPATIAL MODELS

Introduction to Computational Science

October 28, 2023

Student:
14740583 Vishal Kanteppa Mahesh

Course:
Introduction to Computational Science

Course code:
5284ITCS6Y

Contents

1 Introduction

Stochastic simulations play an indispensable role in the realm of computational epidemiology, offering a nuanced understanding of complex systems where deterministic approaches may fall short. These simulations incorporate elements of randomness, capturing the inherent uncertainty and variability found in real-world processes. Particularly when predicting the dynamics of infectious diseases, stochastic simulations can reveal a range of possible outcomes, providing insights into the likelihood of different scenarios and helping to identify critical factors that might tip the balance in one direction or another.

Networks are conceptual representations of interconnected entities, serve as a foundational tool in studying the spread of diseases. In these networks, nodes typically represent individuals or populations, while edges signify interactions or connections that can facilitate disease transmission. Different network architectures, such as scale-free, small-world, or random networks, exhibit distinct properties that can dramatically influence the dynamics of disease spread. For instance, networks with a high degree of clustering might show rapid initial spread but could be more resilient to widespread epidemics due to their inherent community structures[16].

Analyzing disease spread on networks using stochastic simulations allows researchers to combine the unpredictability of infectious processes with the structured nature of human interactions. By doing so, it becomes feasible to anticipate potential outbreak patterns, evaluate intervention strategies, and ultimately gain a more holistic understanding of the multifaceted interplay between diseases and their host populations.

In the subsequent sections, we will delve deeper into how stochastic simulations function in the context of disease propagation within networks and the diverse factors that contribute to their dynamics. Specifically, we will explore the nuances of Gillespie's Direct Algorithm and its role in these simulations. We will also examine the distinctive characteristics and behaviors of various network models, including the Erdos-Renyi networks, Barabasi-Albert networks, and Watts-Strogatz networks, and their implications for disease spread.

2 Background/Theory

2.1 Stochastic Models

By making use of stochastic simulations we seek methods to understand and predict the dynamics of disease spread within populations. Unlike deterministic models that provide a single outcome based on set initial conditions, stochastic models introduce the element of randomness. This randomness acknowledges the inherent uncertainties tied to disease transmission events[10].

Stochastic simulations consider various factors in disease spread: the random nature of person-to-person contacts, the variable susceptibility of individuals, and other unpredictable events like super-spreader occurrences. This randomness becomes crucial, especially when dealing with small populations or when the probability of infection is low, conditions under which deterministic models might not provide adequate insights. Five key features distinguish stochastic models from their deterministic counterparts[15]. They are:

1. Variability between Simulations
2. Variances and Covariances
3. Increased Transients
4. Stochastic Resonance
5. Extinctions

2.2 Gillespie's Algorithms

Gillespie's algorithms are a group of methods used in stochastic simulations to model the time progression of events in systems influenced by randomness. These algorithms are in particular vital for simulating the kinetics of chemical reactions within systems where the exact sequence and timing of individual reactions matter[11]. Among these methods, Gillespie's Direct Algorithm stands out. It determines the precise time until the next event occurs and simulates that event without resorting to approximations. There's also the First Reaction Method and the Next Reaction Method, both variants developed to enhance computational efficiency. The overarching aim of Gillespie's algorithms is to provide an accurate representation of systems influenced by inherent randomness, often proving more genuine than deterministic models in capturing unpredictable nuances[12].

2.2.1 Gillespie's Direct Algorithm

Gillespie's Direct Algorithm is a method used in stochastic simulations to precisely capture the progression of events in systems driven by randomness. Instead of approximations, it provides an exact trajectory of how a system evolves over time, accounting for the inherent randomness in the process. Its necessity arises from the limitations of deterministic models, which can't accurately represent the unpredictable fluctuations in real-world systems, especially when populations are small or events are infrequent[15].

Algorithm 1 Gillespie's Direct Algorithm as stated by [15]

1: Label all possible events $E_1, \ldots, E_n$.
2: For each event determine the rate at which it occurs, $R_1, \ldots, R_n$.
3: The rate at which any event occurs is $R_{\text{total}} = \sum_{m=1}^{n} R_m$.
4: The time until the next event is $\delta t = -\frac{1}{R_{\text{total}}} \log(\text{RAND1})$.
5: Generate a new random number, RAND2. Set $P = \text{RAND2} \times R_{\text{total}}$.
6: **if** $\sum_{m=1}^{p-1} R_m < P \leq \sum_{m=1}^{p} R_m$ **then**
7: Event p occurs.
8: **end if**
9: The time is now updated, $t \to t + \delta t$, and event p is performed.
10: Return to Step 2.

2.3 Extinction Events and Critical Community Size

Extinction events are events where an infectious disease fades away and becomes absent from a community. This disappearance can arise from several factors, including a depletion in the number of susceptible individuals due to previous infections or vaccinations, successful public health interventions, or even purely random variations[15]. Particularly in smaller communities, diseases can vanish solely because of chance, even if prevailing conditions might seem conducive for their persistence. For diseases that hover near their extinction thresholds, even modest interventions can tip the balance, leading to disease elimination in specific regions.

Critical Community Size is the minimum population size required for an infectious disease to remain endemic over an extended period[3]. When a population's size falls below this threshold, random fluctuations in infection rates due to various factors like births, deaths, and recoveries can lead to such extinction events as mentioned above. In communities exceeding this critical size, while diseases might experience temporary extinctions, the frequency of reintroductions ensures that the disease remains consistently present, becoming endemic[13]. The determination of CCS depends on a large number of factors, including the disease's basic reproductive number (R_0), the infectious period, and birth/death rates within the population. By identifying regions with populations that fall below the CCS, health authorities can strategically channel limited resources where they are most effective.

2.4 Spatial Models - Networks

Networks are fundamental in epidemiology, representing the interactions and relationships within populations that influence disease transmission. Each individual can be viewed as a node, and their interactions as edges forming the network. Different network structures, such as random, scale-free, or small-world networks, provide insights into how diseases might spread.

2.4.1 Erdos Renyi Network

The Erdos-Renyi model involves creating a network with n nodes and connecting each pair of nodes with a fixed probability p. As a result, the degree of each node follows a binomial distribution[1]. One hallmark of ER networks is that they don't exhibit the hub like structure seen in many real world networks. Instead most nodes have roughly the same degree. A critical feature of ER networks is the emergence of a giant connected component once the connection probability surpasses a certain threshold[15]. This threshold behavior is important for understanding phase transitions in complex networks.

2.4.2 Barabasi Albert Network

The Barabasi–Albert network is a model that generates scale-free networks, capturing the inherent growth and preferential attachment observed in many real world networks. This model starts with a small number of nodes and grows by adding new nodes one by one. As each new node is introduced it forms connections to existing nodes with a probability proportional to their degree. This preferential attachment mechanism ensures that nodes that have gained more connections early on (often called hubs) continue to gain more connections as the network grows[2]. The resulting degree distribution of nodes in a BA network follows a power law, indicating that while most nodes have only a few links there are a few highly connected nodes that play a crucial role in the network's structure and dynamics.

2.4.3 Watts-Strogatz Network

The Watts–Strogatz model is known for producing networks that exhibit both small-world properties and a high clustering coefficient. Starting with a regular lattice where each node is connected to its nearest neighbors, the WS model introduces randomness by rewiring edges with a probability p. When $p = 0$ no edges are rewired and the network remains a regular lattice. As p increases, more edges are randomly rewired, reducing the average path length between nodes[14]. When $p = 1$ the network becomes completely random. The uniqueness of the WS model lies in the intermediate values of p, where the network maintains a high clustering coefficient like a regular lattice yet has small average path lengths when compared to a random network.

2.5 Network Statistics

2.5.1 Degree Distribution

provides insight into the network's structure and the connectivity patterns of its nodes. In essence, degree distribution, often denoted as *P(k)*, gives the probability that a randomly selected node from the network has degree k. The degree of a node in a network signifies the number of connections or contacts it has with other nodes. In the context of disease spread, a node's degree represents the number of potential direct disease transmission pathways it possesses. Equations 1, 2 and 3 show the degree distributions for the ER, BA and WS networks respectively [4][5].

$$P_{\mathrm{ER}}(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k} \tag{1}$$

$$P_{\mathrm{BA}}(k) \sim k^{-3} \tag{2}$$

$$P_{\mathrm{WS}}(k) = \delta(k - k_0) \tag{3}$$

- k is the degree of a node

- k_0 is a characteristic degree for the Watts–Strogatz model.
- δ is the Dirac delta function, indicating the network has a fixed degree k_0 for the Watts-Strogatz model.

2.5.2 Betweenness

Betweenness centrality is a measure used to quantify the significance of a node in terms of its position within the overall network topology. It represents the proportion of shortest paths between pairs of nodes that pass through a given node. In other words it measures the extent to which a particular node acts as a bridge between other nodes in the network. Nodes with high betweenness centrality are crucial in disease transmission dynamics, as they serve as potential bridges for disease spread between different clusters or communities. Equation 4 shows how to numerically calculate betweenness in a network [6].

$$B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{4}$$

Where:

- σ_{st} is the total number of shortest paths from node s to node t.
- $\sigma_{st}(v)$ is the number of shortest paths from node s to node t that pass through node v.

2.5.3 Clustering Coefficient

Clustering coefficient is a measure that captures the degree to which nodes in a network tend to cluster or form tightly knit groups[7]. It quantifies the likelihood that two neighbors of a given node will themselves be neighbors. The clustering coefficient offers insights into how diseases might spread within communities. In networks with a high clustering coefficient infections can spread rapidly within closely knit groups. However these local clusters can sometimes act as barriers, limiting the disease's spread to broader parts of the network. Equation 5 shows how to calculate the clustering coefficient for a given network [15].

$$C(i) = \frac{2e_i}{k_i(k_i - 1)} \tag{5}$$

- k_i is the degree of node i
- e_i is the number of edges between the neighbors of node i.

2.6 Vaccination Strategies

By making use of knowledge on network structures, our vaccination strategie prioritize vaccinating nodes or individuals with numerous connections making them potential hubs for disease transmission. This targeted method portrays a huge difference to random vaccinations which provide vaccines without considering a node's connectivity or risk profile. Given the limited resources such as tests and vaccines, targeted vaccination ensures that each dose is used to its utmost potential, concentrating on those who, if unvaccinated, might significantly contribute to disease spread. The strategic importance of testing with varying degrees of accuracy emphasizes the need for careful resource deployment. Even with potential test inaccuracies focusing on high risk nodes can offer valuable insights into transmission dynamics. In essence, while both targeted and random vaccination methods provide protection the former does with high precision ensuring a more robust defense against disease spread.

3 Programming Methods

3.1 Numerical Integration

As in Assignment 1, the *odeint()* function was used to obtain the values of S,I and R over some time period to obtain the Deterministic dynamics when comparing to the equivalent Stochastic one.

3.2 Gillespie's Algorithms

Gillespie's Direct algorithm is implemented in a function called *gssa()*. The implementation is relatively straightforward and all individual steps mentioned in Algorithm 1 have been manually put into code.

3.3 NetworkX and NDlib

The *networkx* library[8] is used for constructing various networks in this assignment. Methods such as *erdos_renyi_graph()*, *barabasi_albert_graph()*, and *watts_strogatz_graph()* are employed to create the Erdos–Renyi, Barabasi-Albert, and Watts-Strogatz networks respectively. Upon establishing the desired network the SIR dynamics are introduced using the *SIRmodel()* function from *ndlib.models.epidemics*[9]. This function facilitates the application of the SIR model to the network allowing for the simulation of disease spread. After the model's parameterization, the simulation is executed. Metrics like susceptible, infected, and recovered counts over time are extracted utilizing the *iteration_bunch()* and *build_trends()* methods. The functions dedicated to each specific network have intuitive names reflecting the network they represent (eg- Erdos–Renyi network is present within a function titled *erdos_renyi()*). The *networkx* library was also used to create the graph from the sociopatterns dataset. As the degree of a node can be obtained with this library this assisted in the smart vaccination strategy of vaccinating based on the degree of the node.

4 Experiment Results

4.1 Simulation Variability

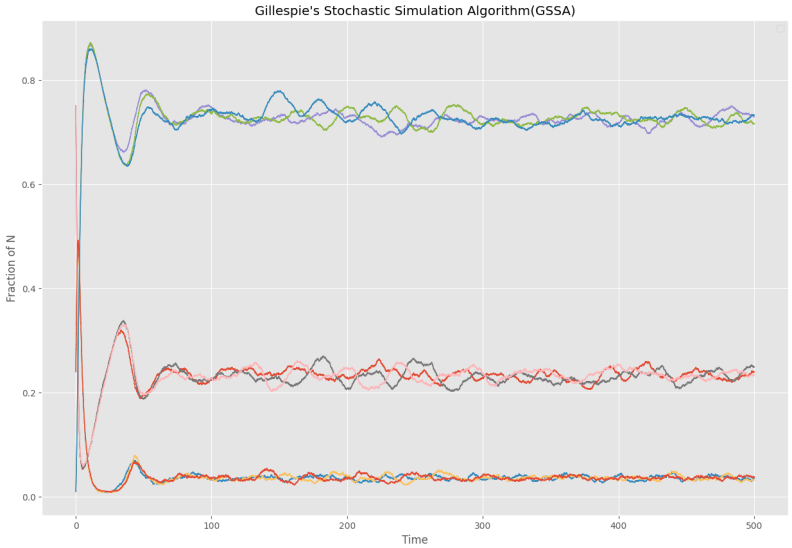


Figure 1: Variation between multiple GSSA simulations

Figure 1 displays the results of three simulations(**with the same set of parameters**) using Gillespie's Stochastic Simulation Algorithm (GSSA) for the dynamics of susceptible, infectious, and recovered individuals. The top most, middle and bottom curves represent the fraction of S, R and I respectively. The variations in the trajectories of S, I, and R observed between the simulations are primarily attributed to **stochasticity**.

Stochasticity, or randomness inherent in the GSSA method, introduces natural fluctuations into the simulations. These variations are a reflection of the unpredictable nature of disease spread dynamics, where **individual-level interactions**, infection events, and recoveries are subject to chance. Therefore, while the underlying model remains consistent, the specific outcomes can differ due to the random interplay of factors. Table 1 shows the parameters for this simulation.

Parameter	Parameter Name	Values
$\boldsymbol{\beta}$	Transmission Rate	1.8
$\boldsymbol{\gamma}$	Recovery Rate	0.4
$\boldsymbol{\mu}$	Birth/Death Rate	$\frac{1}{50}$
$\boldsymbol{N}$	Population Size	1×10^4
$\boldsymbol{S}$	Initial Susceptible Fraction	0.75
$\boldsymbol{I}$	Initial Infected Fraction	0.24
$\boldsymbol{R}$	Initial Recovered Fraction	0.01

Table 1: Parameter values for GSSA simulations

4.2 Negative Co-Variances and Increased Transients

Figure 2 presents a comparative visualization of stochastic simulation versus the deterministic counterpart. The stochastic dynamics exhibit a tendency to oscillate around the natural frequency inherent to the deterministic model. A closer examination reveals clear **stochastic perturbations**, marked by the presence of red arrows in the graph. These perturbations momentarily push the system away from its **natural frequency**. The response of the system to these stochastic perturbations is quite interesting. In a seemingly orchestrated fashion, deterministic forces within the model counteract these perturbations, compelling the system to display transient-like behavior as it returns to its natural frequency. These **increased transients** are highlighted by the presence of orange arrows in the graph.

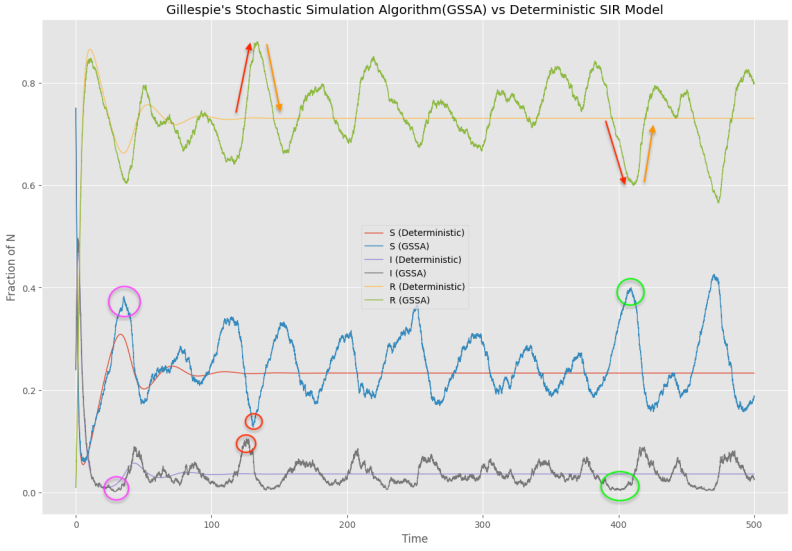


Figure 2: Graph Showcasing Negative Co-Variance between S and I and Transient Nature of Stochastic Simulations

Figure 2 also prominently illustrates the **negative covariance** between the Susceptible and Infectious fractions. In this context, the blue line corresponds to the S fraction, while the black line represents the I fraction. A distinctive pattern emerges as highlighted by the circular markers on the graph: **when one variable experiences a notable spike, the other undergoes a significant dip**.

To provide clarity in visualizing this relationship, we've employed a color-coding scheme for the circular markers. Each pair of markers sharing the same color serves as a visual representation of instances where an increase in one variable corresponds to a decrease in the other. This complementary behavior, signified by the colors, highlighting the negative covariance between S and I fractions.

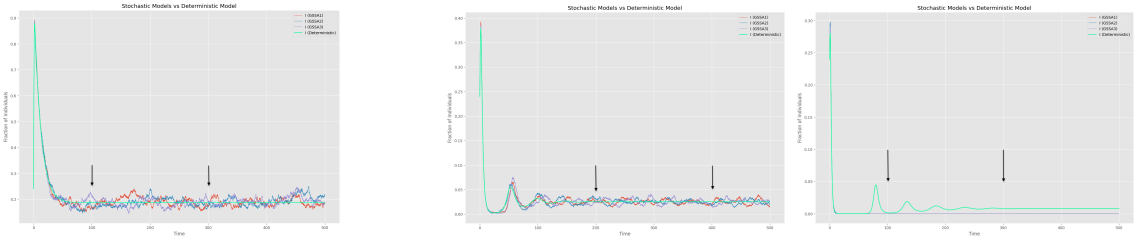
Parameter	Parameter Name	Values
$\boldsymbol{\beta}$	Transmission Rate	1.8
$\boldsymbol{\gamma}$	Recovery Rate	0.4
$\boldsymbol{\mu}$	Birth/Death Rate	$\frac{1}{50}$
$\boldsymbol{N}$	Population Size	1×10^3
$\boldsymbol{S}$	Initial Susceptible Fraction	0.75
$\boldsymbol{I}$	Initial Infected Fraction	0.24
$\boldsymbol{R}$	Initial Recovered Fraction	0.01

Table 2: Parameter values to display Negative Co-Variance and Increased Transients

Table 2 shows the parameters used for simulation which ended up creating large transients and peturbations. In this simulation, we kept all the parameters consistent with those detailed in Table 1, with the sole alteration being a reduction in the population size. In smaller populations, the effect of stochasticity becomes more readily apparent. This arises from the diminished scale of interactions among individuals. Each individual interaction assumes increased significance within the context of a smaller population. Furthermore, random events, such as the timing of infections and the nature of interpersonal encounters, take on heightened importance. In these reduced populations, the stochastic impact of these events can result in more noticeable oscillations in disease propagation.

4.3 Comparing Mean b/w Stochastic and Deterministic Simulations

Figures 3a, 3b and 3c show the the Infected fractions of 3 different simulations run with 3 different set of parameters. The arrows present in these figures show the time points around which the mean was calculated. Since the exact time points cannot be compared in stochastic models, the mean was calculated over a very small window around the arrow. The values of the parameters used are shown in Table 4.



(a) Simulation 1 (b) Simulation 2 (c) Simulation 3

Figure 3: Comparing Means of Multiple Simulations

The means of the **infected fractions** for each simulation is shown in Table 3. For our comparison, we employed three distinct sets of parameters(shown in Table 4) and conducted simulations to explore the impact of different infection strengths. Each set consisted of one deterministic simulation and three stochastic simulations, providing a perspective on the dynamics of infectious diseases under varying conditions.

The first parameter set was designed to represent a scenario with a relatively strong infection. To achieve this, we assigned a high **Basic Reproductive Ratio**, which signifies the average number of secondary infections generated by an infected individual.

In contrast, the second parameter set was characterized by a comparatively weaker infection. Here, the Ro value was intentionally reduced in comparison to the first set, indicating a lower potential for widespread transmission.

The third parameter set was crafted to represent a scenario with a notably weak infection. We anticipated that the infection in this case would struggle to propagate extensively. Therefore, the Ro value for the third set was set at a level that reflected this limitation.

	Timestamp	**Infected Fraction**
Deterministic (Sim 1)	100-105	0.205
GSSA 1	100-105	0.186
Deterministic (Sim 1)	300-305	0.205
GSSA 1	300-305	0.185
Deterministic (Sim 2)	200-205	0.027
GSSA 2	200-205	0.021
Deterministic (Sim 2)	400-405	0.027
GSSA 2	400-405	0.016
Deterministic (Sim 3)	100-105	0.001
GSSA 3	100-105	0.0
Deterministic (Sim 3)	300-305	0.008
GSSA 3	300-305	0

Table 3: Means of Different Stochastic Simulations Compared to Equivalent Deterministic Simulation

The data in Table 3 reveals subtle variations in mean values between deterministic and stochastic simulations for the same timestamps. Notably, in the case of 'Sim3,' the mean values for the deterministic simulation are exceedingly close to 0, while the corresponding values for the stochastic simulation are exactly 0. This phenomenon can be attributed to the specific parameters chosen for 'Sim3,' reflecting a scenario of a weak infection. The deterministic simulation maintains values near 0 due to the reduced infectiousness associated with this scenario. In contrast, the stochastic simulation results in precisely 0 values because of the inherent randomness in stochastic modeling. This randomness, particularly in cases of small populations, can lead to the population's extinction

Parameter	**Parameter Name**	**Simulation 1(Fig:3a)**	**Simulation 2(Fig:3b)**	**Simulation 3(Fig:3c)**
$\boldsymbol{\beta}$	Transmission Rate	3	1.5	2
$\boldsymbol{\gamma}$	Recovery Rate	0.07	0.5	1
$\boldsymbol{\mu}$	Birth/Death Rate	$\frac{1}{50}$	$\frac{1}{60}$	$\frac{1}{60}$
$\boldsymbol{N}$	Population Size	1×10^3	1×10^4	1×10^4

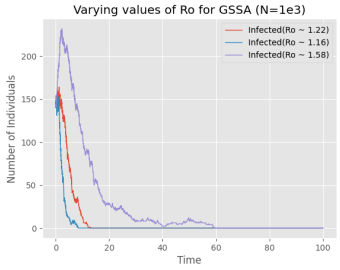
Table 4: Parameter values for Simulations

4.4 Extinction Events and Critical Community Size

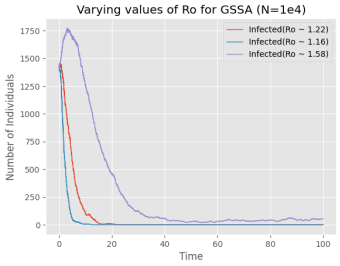
In Figure 4, we examine the variation in the fraction of infected individuals across three distinct R_0 values for different population sizes.

In Figure 6a, with a population size N of 1×10^3, despite all R_0 values leading to eventual disease extinction, the duration of persistence differed among them. The infected fraction corresponding to $R_0 = 1.58$ (represented by the purple line) persisted noticeably longer than the other two. However, it did eventually diminish to zero.

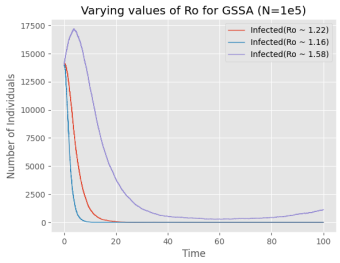
For Figure 6b, where the population size increased tenfold to 1×10^4, while the trajectories for the lower two R_0 values still underwent extinction, the curve for $R_0 = 1.58$ sustained, albeit at minimal levels of infected individuals.



(a) $N = 1 \times 10^3$



(b) $N = 1 \times 10^4$



(c) $N = 1 \times 10^5$

Figure 4: Impact of Basic Reproductive Ratio and Population Size on Disease Extinction

Lastly, Figure 4c, with the population size set to 1×10^5, exhibited similar trends as observed in Figure 4b. Nevertheless, the progression of disease dynamics for $R_0 = 1.58$ appeared smoother compared to the slightly more erratic fluctuations observed in the previous scenario. Table 5 shows the individual values of β, γ and μ used for these simulations.

Table 5: Parameter Values for Figure 4

Simulation	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\mu}$
$R_0 = \mathbf{1.22}$	1	0.8	$\frac{1}{60}$
$R_0 = \mathbf{1.16}$	2	1.7	$\frac{1}{60}$
$R_0 = \mathbf{1.58}$	0.5	0.3	$\frac{1}{60}$

5 Spatial Models - Networks

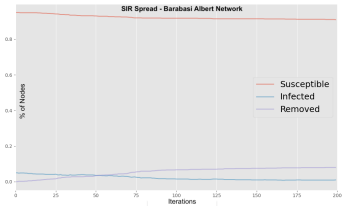
Figure 5 depicts the behavior of the SIR spread across three distinct networks using three different sets of parameters. Each row in the figure demonstrates the SIR spread within a specific network under varied parameters, while each column represents the consistent application of a particular parameter set across different networks. The first row illustrates the Barabási-Albert network, the second row displays the Erdős-Rényi network, and the third row, the Watts-Strogatz network.

The parameters, as detailed in Table 6, were created to simulate the progression of disease strength. Set 1 denotes a milder disease with a transmission rate (β) of 0.001. Set 2 indicates a moderately strong disease with β=0.01, while Set 3 exemplifies a considerably stronger disease with β=0.1. While there were variations in other parameters, the general trend and underlying rationale for distinguishing these sets revolved around the disease strength.

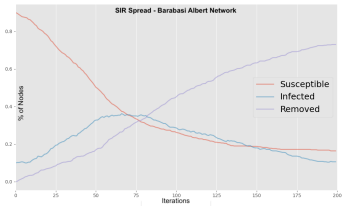
In the Watts-Strogatz network, the SIR spread is notably subdued for the first two parameter sets. It becomes evident only with parameter set 3, which characterizes a more potent disease. This observation might be attributed to the unique structure of the Watts-Strogatz network, known for its small-world properties, leading to a limited spread until a highly contagious disease, represented by Set 3, breaks these structural barriers.

The Barabasi-Albert network, characterized by its scale-free properties and hubs, reveals noticeable disease spread dynamics starting from parameter set 2. By the time parameter set 3 is applied, the number of susceptibles plummets rapidly. This could be due to the hubs or highly connected nodes in the Barabasi-Albert network. When a strong disease (like that in Set 3) gets introduced, these hubs accelerate the spread, quickly depleting the pool of susceptible individuals.

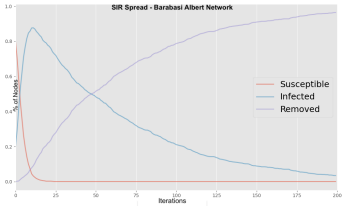
Finally, the Erdos-Renyi network starts to exhibit changing dynamics right from parameter set 1. For parameter sets 2 and 3, the decline in the number of susceptibles is precipitous, indicating the strong impact of the disease on this network. The Erdos-Renyi network's randomness and equal likelihood of connections could make it more vulnerable, even to milder diseases represented by Set 1. As the disease strength increases with Sets 2 and 3, the network's structure



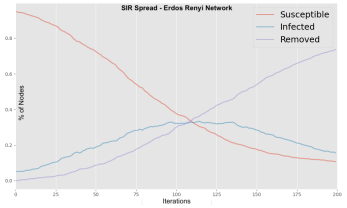
(a) BA - Paramset1



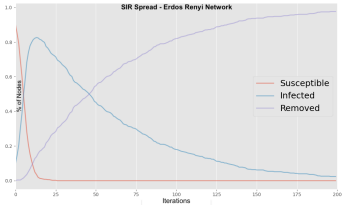
(b) BA - Paramset2



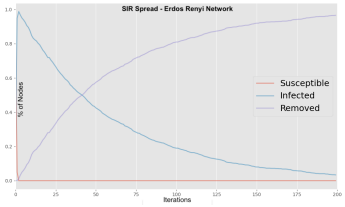
(c) BA - Paramset3



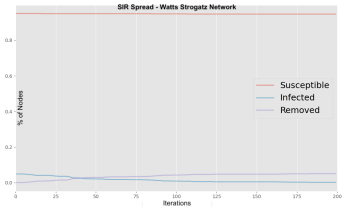
(d) ER - Paramset1



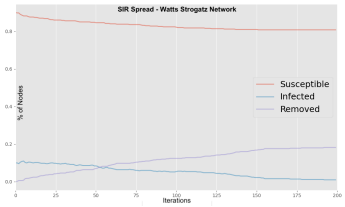
(e) ER - Paramset2



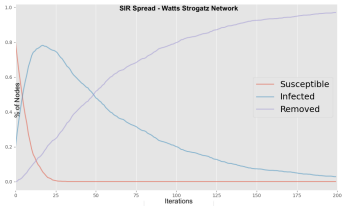
(f) ER - Paramset3



(g) WS - Paramset1



(h) WS - Paramset2



(i) WS - Paramset3

Figure 5: Comparison of SIR spread across different networks and parameter sets. The first row shows the spread on the Barabasi Albert network with varying parameter sets, while the subsequent rows display the spread on Erdos Renyi and Watts Strogatz Networks. Each column corresponds to a distinct parameter set.

offers little resistance, leading to rapid disease proliferation.

Parameter	Set 1	Set 2	Set 3
$\boldsymbol{\beta}$	0.001	0.01	0.1
$\boldsymbol{\gamma}$	1/60	1/60	1/60
Initial I Fraction	0.05	0.1	0.2
Number of Nodes	1000	500	1000
Connecting Probability (Erdos Renyi Network)	0.05	0.1	0.15
m (Barabasi Albert Network)	5	3	3
k (Watts Strogatz Network)	5	3	5
Rewiring Probability (Watts Strogatz Network)	0.05	0.1	0.15

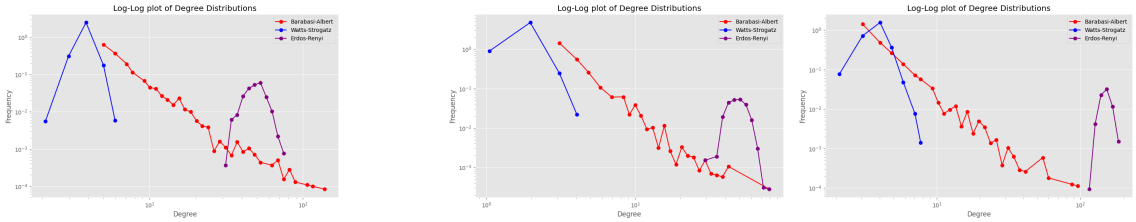
Table 6: Parameter sets for different network models

5.1 Comparison of Network Statistics

5.1.1 Degree Distributions

Figure 6 presents the degree distributions of three distinct networks across different parameter sets as detailed in Table 6. These networks are: Watts-Strogatz represented in blue, Barabasi-Albert in red, and Erdos-Renyi in purple.

The WS network, particularly in Parameter Sets 1 and 2, displays a clear tapering off in frequency as the degree rises. This hints at the presence of hubs or clusters of nodes with higher connectivity, especially under these parameter configurations.



(a) Parameter Set 1 (b) Parameter Set 2 (c) Parameter Set 3

Figure 6: Degree Distributions of BA, ER and WS Networks across 3 parameter sets

The BA network showcases a distinctive peak in its degree distribution, most noticeable in Parameter Set 1. This behavior aligns with scale-free networks which often adhere to a **Power-Law distribution**. Such a pattern emerges from preferential attachment phenomena, where newer nodes tend to connect with those that are already well-linked.

The ER network is recognized for its randomness. This is evident from the somewhat bell-shaped distribution in Parameter Set 1, becoming flatter in subsequent sets. Typically, the degree distribution in ER networks leans towards a **Poisson Distribution**, indicative of its inherent unpredictability in connections.

5.1.2 Betweenness

Figure 7 illustrates a detailed comparison of betweenness centrality across the BA, ER, and WS networks for three different parameter sets. The BA network, depicted by the three violins on the left, exhibits a stark variation in the distribution of betweenness values. This is particularly evident for parameter set 1 (blue), which showcases a notably thin profile, implying a high disparity in node centralities within this configuration. Similarly, the WS network, represented by the violins on the right, shows an even sharper contrast in its distribution, especially for parameter set 1 and 2, suggesting that certain nodes significantly dominate the information flow in these setups.

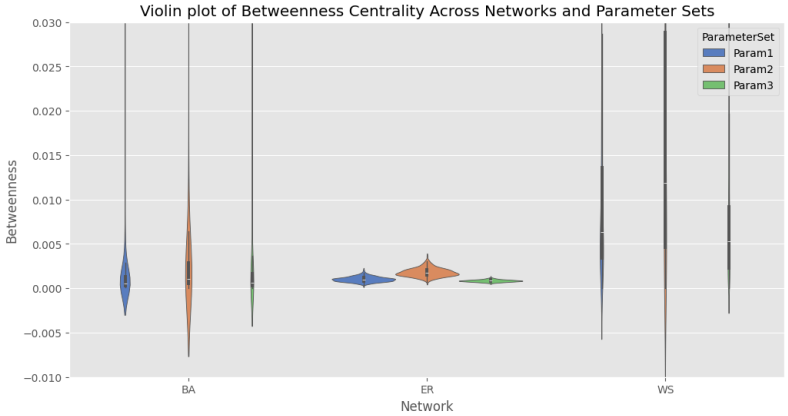


Figure 7: Betweenness Comparison across Networks and Parameter Sets

The ER network, visualized by the central violins, offers a different narrative. Across all parameter sets, its violins maintain a thicker, more uniform profile, indicating that betweenness centrality is more evenly distributed among nodes. In essence, this suggests that no singular node or cluster dominantly dictates the flow of information in the ER network, leading to a more homogenous spread.

Such trends can be rationalized by the inherent structure of these networks. The BA and WS networks are characterized by **Preferential Attachment** and **Small-World Properties**, respectively. These features can lead to the emergence of hubs or specific nodes that command

a significant portion of the network's connections. In contrast, the ER network, being random, naturally steers toward a balanced distribution, ensuring no node is disproportionately central.

5.1.3 Clustering Coeffiecient

In Figure 8, we compare the clustering coefficients of three different network types – Erdos Renyi, Barabasi Albert, and Watts Strogatz – across three sets of parameters. These sets were created to show increasing strengths of infectious diseases.

The Watts Strogatz Network consistently shows the highest average clustering coefficient values. It's followed by the Erdos Renyi Network and then the Barabasi Albert Network. One main reason for this order is how each network model works. The Barabasi Albert model works on a **preferential attachment** system. This means new nodes tend to connect to the big hubs in the network that already have many connections. Because of this, the network ends up with a few big hubs and many nodes connected to them, which leads to a lower clustering coefficient. On the other hand, both Erdos Renyi and Watts Strogatz networks don't use this system, so they show higher clustering coefficients.

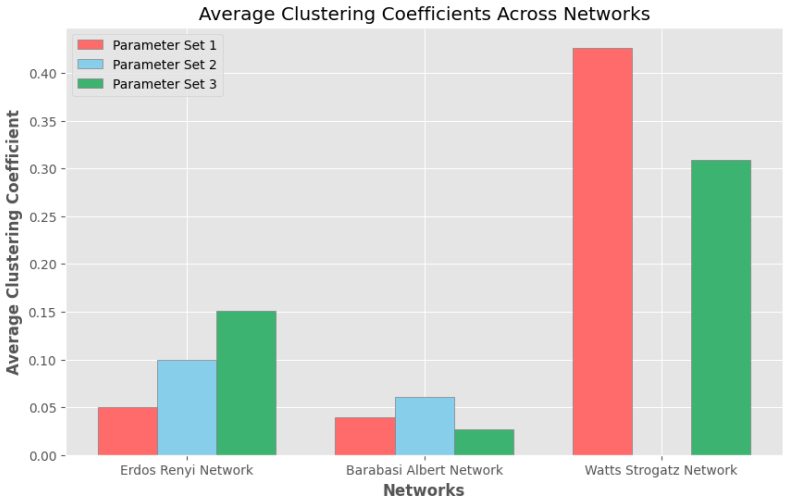


Figure 8: Clustering Coefficient Comparison across Networks and Parameter Sets

Looking closer, the Erdos Renyi Network has a clear rise in the clustering coefficient as the disease strength goes up. This shows that a stronger disease makes nodes in the network connect more closely to each other. Despite the Watts Strogatz Network typically exhibiting pronounced clustering coefficients, a marked deviation is observed for Parameter Set 2, with the coefficient plummeting to zero. In efforts to modulate the network, the rewiring probability was adjusted. Simultaneously, the **k value** was reduced from 5 to 3. These modifications evidently had a significant impact on the network's characteristics, highlighting the pivotal role of individual parameters in determining a network's behavior.

5.2 Dynamic Vaccination Campaign

Figures 9 and 10 provide a comparative visualization of the performance between a smart vaccination strategy, which prioritizes nodes with the **highest degrees**, and a random vaccination approach. It's clear that the smart strategy consistently outperforms the random method, achieving a significant reduction in the number of infections, especially at higher vaccination budgets. On the other hand, the random vaccination, while showing some decline in infections with increased budget, does not achieve as pronounced a decrease, emphasizing the importance of strategic allocation over mere quantity in vaccination drives.

The heat maps highlight a crucial observation: both the smart and random vaccination strategies achieve their highest performance when the vaccinations per timestep are at their maximum (10) and the accuracy of the vaccinations is also optimal (1). This could be due to

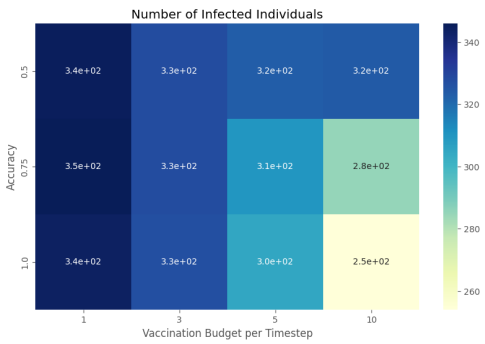


Figure 9: Smart Vaccination Strategy

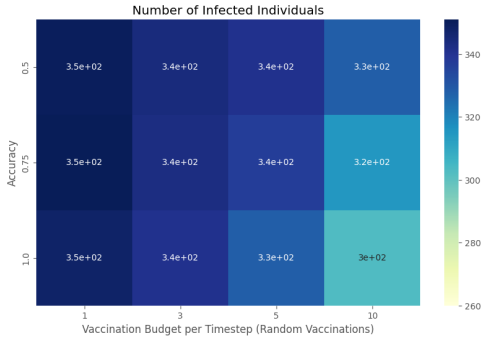


Figure 10: Random Vaccination Strategy

vaccinating larger numbers simultaneously substantially diminishes the pool of susceptibles, transitioning them to a recovered state where they are immune to further infections. When the pace of vaccination is slower, a more significant portion of the population remains susceptible for longer, increasing the chances of them contracting the infection.

6 Discussion and Summary of Findings

In our exploration of stochastic simulations in disease spread, we encountered distinctive differences that set them apart from deterministic models. One of the more striking observations was the pronounced **noise** and **variability** present in stochastic simulations. Despite operating under a **basic reproductive ratio** of approximately 2, we found that the infected fraction rapidly declined to zero shortly after simulation initiation. This is interesting given our expectations aligned with a prolonged infection scenario based on the R_0 value. Our findings further suggest that **smaller populations** are more prone to sudden infection extinctions while larger populations yielded more consistent less volatile results.

Our analysis of network models reinforced some well established concepts. The degree distributions for both the Barabasi-Albert and Erdos–Renyi models behaved predictably, with the BA model adhering to the **power-law distribution** and the ER model aligning with a **Poisson distribution**. In contrast, the degree distribution of the Watts-Strogatz model, characterized as a **small-world network**, did not adhere to a straightforward trend. This might be attributed to the WS model's inherent balance between regularity and randomness, an attribute that also endows it with a significant **clustering coefficient** across a range of parameters.

Our evaluation of vaccination strategies yielded a clear correlation: increasing the vaccination budget enhances the strategy's efficacy. This can be understood by seeing that a more generous budget allows for broader vaccinations, especially targeting high risk nodes or individuals. Such **targeted vaccination** reduces the susceptible pool quickly cutting down on potential transmission pathways and hence showcasing the advantages of a targeted approach over a randomized one.

References

1. July 2022. URL: https://www.geeksforgeeks.org/erdos-renyl-model-generating-random-graphs/.
2. May 2022. URL: https://www.geeksforgeeks.org/barabasi-albert-graph-scale-free-models/.
3. Sept. 2023. URL: https://en.wikipedia.org/wiki/Critical_community_size.
4. Oct. 2023. URL: https://en.wikipedia.org/wiki/Erd%C5%91s%E2%80%93R%C3%A9nyi_model.

5. Aug. 2023. URL: https://en.wikipedia.org/wiki/Barab%C3%A1si%E2%80%93Albert_model.
6. July 2023. URL: https://en.wikipedia.org/wiki/Betweenness_centrality.
7. Feb. 2023. URL: https://en.wikipedia.org/wiki/Clustering_coefficient.
8. URL: https://networkx.org/documentation/stable/tutorial.html.
9. URL: https://ndlib.readthedocs.io/en/latest/tutorial.html.
10. Linda J.S. Allen. *A Primer on stochastic epidemic models: Formulation, numerical simulation, and analysis*. Mar. 2017. URL: https://www.sciencedirect.com/science/article/pii/S2468042716300495.
11. Jun Chu. URL: http://drp.math.umd.edu/Project-Slides/ChuFall2012.pdf.
12. Lewis Cole. *Gillespie algorithm*. Apr. 2020. URL: https://lewiscoleblog.com/gillespie-algorithm.
13. Daniel T Haydon et al. *Identifying reservoirs of infection: A conceptual and Practical Challenge*. Dec. 2002. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2738515/.
14. Chih-Ling Hsu. *Watts-Strogatz model of small-worlds*. URL: https://chih-ling-hsu.github.io/2020/05/15/watts-strogatz.
15. Rohani Keeling. "Modeling Infectious Diseases in Humans and Animals by KEELING, M. J. and ROHANI, P." In: *Biometrics* 64.3 (Aug. 2008), p. 993. DOI: 10.1111/j.1541-0420.2008.01082\{_}7.x. URL: https://doi.org/10.1111/j.1541-0420.2008.01082_7.x.
16. Danon L;Ford AP;House T;Jewell CP;Keeling MJ;Roberts GO;Ross JV;Vernon MC; *Networks and the epidemiology of infectious disease*. URL: https://pubmed.ncbi.nlm.nih.gov/21437001/.