TEMPORAL AND SPATIAL EPIDEMIC DIFFUSION PATTERNS IN SMALL REGIONAL SETTINGS

James L. Wilson
Dickinson College and
University of North Carolina

ABSTRACT. The patterns of communicable disease epidemics over human space result from the combination of several factors. These factors can be categorized broadly into two sets: the biological characteristics of human host and disease agent, and culture. Several aspects of culture are reified into particular human settlement patterns. These aspects include subsistence modes and social networks. This paper considers the influences of both settlement patterns and the biological characteristics on epidemic patterns. Using monte-carlo techniques, contact probability networks, disease agent and host characteristics, and the Reed-Frost formulation of individual infective and susceptible interaction, communicable disease epidemics can be simulated on a microcomputer. The resulting simulations demonstrate the influence that settlement pattern has on the temporal and spatial spread of epidemics in small regional systems. If the effects of spatial organization can be controlled then it may be possible to gain an understanding of the biological effects on observed epidemic patterns.

The temporal patterns of epidemic morbidity and mortality serve as a barometer of a population's health and biological interaction with a disease agent. These patterns may also reveal information about the nature of a population's spatial organization. Settlement patterns resulting from socioeconomic processes and topography affect interpersonal contact patterns and rates, which then produce measurable effects on observed epidemic time series distributions. Conceivably, the temporal patterns of epidemics occurring in rural places would be different from those occurring in urban areas. Unique combinations of these factors have the potential to yield a wide variety of epidemic patterns.

To illustrate this point, let's look at two examples of observed epidemic patterns, one characterized as "short and severe" and the other "long and mild." Short and severe epidemics observed in a place can be described as time series distributions that possess large numbers of cases clustered within a short time interval. How do the previously introduced epidemic factors combine to produce these effects? A number of explanations are possible: short disease transmission chains, a highly infectious disease organism, the crowding of susceptible individuals, sanitary practices, or the general health and nutritional status (McKeown 1988) of a population. When contacts between susceptible and infectious individuals are sporadic (due to the spatial arrangement of individuals) and infection rates are low, slower and milder epidemics result and tend to produce less clustered time series observations.

The patterns described above could be the result of any combination of factors such as organism infectiousness, population nutritional status and susceptibility, and the social and spatial arrangements of individuals and households (Angulo 1987). The subject of this paper is a method by which the relative effects of these factors on the observed epidemic time series distributions can be controlled and evaluated.

In order to assess the relative effects of spatial and biological factors on epidemic time series distributions, an epidemic simulation model was constructed. The model is a software program (MicroSoft's QuickC v2.51) written for use on a microcomputer. This computer program simulates the movement of infection among susceptible individuals located in various spatial arrangements of villages within a small regional setting. The model generates epidemic time series distributions for the entire region, and these

distributions can then be reduced and compared using statistical parameters that describe the temporal distribution curves. The next section treats the construction of the epidemic simulation model and provides examples of model runs. Following the model section, a synopsis is provided on the method by which simulated epidemic time series can be portrayed and compared.

An Epidemic Simulation Model for Small Regional Settings

Several components are built into this epidemic simulation model. These include functions for writing and reading disease agent, population, and locational data; algorithms calculating contact (infection) probabilities and mobility behavior; as well as a means of keeping track of newly infected cases and when and where they occur. Each of these components is discussed more fully below.

Disease and population characteristics are placed in a small dataset for input prior to the run of the model. This information establishes the initial conditions of the model system. The dataset contains the following information:

1. initial susceptible population size
2. initial contagious population size
3. household size
4. probability of transmission upon contact

Certain aspects of the disease agent's natural history are also included:

5. length of latency period after infection
6. length of the infectious period

After the initial conditions of the model system are entered as data, susceptible individuals and the general population are distributed among the villages within the region. Each village is comprised of a set of average household sizes. The partitioning of villages into these smaller units establishes susceptible population densities at the family or household level. These densities are then used to calculate probabilities of contact within a village and among villages.

Villages are located in a two-dimensional graph space referenced by x and y coordinates. Distances are linear. The spatial system of villages is considered as a node network over which a contagion travels. The network is also represented by a two-dimensional adjacency matrix. Joins between two villages (node and neighbor) are indicated as the linear distances between them. When no joins are present, a zero is placed in the appropriate matrix element. The matrix, as input data for the simulation model, is actually a "look-up table" for determining where the epidemic proceeds. This method of spatial representation has the advantage of being both conceptually simple and computationally fast. For a more extensive treatment of network and matrix construction and manipulation, see Unwin (1981) and Haggett et al. (1977).

The driving mechanism of the simulation model is the probability of an infection occurring when a susceptible individual makes contact with a contagious individual. The probability of infection is based on a simple version of the Reed-Frost formulation (Ackerman et al. 1984). In this formulation, individuals are infected according to the probability density of contact and subsequent infection. The greater the number of susceptibles available to contagious individuals, the greater is the likelihood of infection. As the local (village) population of contagious and latent individuals increases and the local pool of susceptibles is exhausted, the likelihood of infection will decrease. The algorithm determining infection is based on the Reed-Frost formulation and can be represented as:

$$p(infect) = 1 - q^{C_t}, \qquad (1)$$

where $p(infect)$ is the probability of contact and infection before the number of contagious individuals at the village node is considered at time t, $q = 1 - p$ is the probability of escaping infection, and C_t is the the number of contagious individuals found at the village at time t.

It can be seen from this equation that as the number of contagious individuals increases, then $p(infect)$ increases. The value of p is recalculated for each subsequent iteration. It is based on the current local susceptible population density. The results of this process are two proportions totaling 1.00: $p(infect)$ and q^{C_t}.

After the probabilities of infection and escape from infection are calculated, a random number, R, is selected and scaled between 0.00 and 1.00. If the probability of an individual being infected is .73, and R is less than .73, then an infection of a susceptible takes place. (This is basically a monte-carlo method.) As a result of this infection, one individual is subtracted from the current pool of susceptibles. For the next iteration, the density of susceptibles decreases and hence the likelihood of contact between contagious and susceptible individuals diminishes (this affects the term p). The algorithm is thus recursive in the sense that output from the prior iteration serves as input to the process at the next iteration.

The above discussion describes how infection occurs at the intravillage level in the simulation model. Changing scales, we now turn to a consideration of intraregional mobility and how infectious contacts are made among villages.

In this study, villages are connected to each other by roads, bridges, and footpaths. This spatial structure implies that the spread of infection will proceed in a steplike fashion from one village to the next (the neighborhood diffusion effect). Since the distances are known between a village experiencing infection and its neighbors, the probabilities of making intra- and intervillage contacts can be calculated. The inverse distances between a village and its neighbors are first summed. Next, these inversed distances are calculated as individual proportions relative to the sum of inversed distances. The total of these proportions equals 1.00. Nearby villages will have larger proportions than more distant villages. Another random number is chosen and scaled to a value between 0.00 and 1.00. This value is then compared to the relative interval sizes of the village distance proportions. For example, if the closest village is proportioned from 0.00 to 0.30 and the scaled random number is 0.33, then the village that occupies the interval between 0.31 and 0.38 (say, a relatively distant neighboring village) is selected and contacted by a contagious individual from the currently infected village.

Intravillage contact between susceptible and contagious individuals is possible and is usually the most likely. This behavior is based on the predetermined propensity of an individual to be mobile within the village system. The distance between a village and itself is technically 0.00. To incorporate this into the proportional distance scheme would be problematic, since division by zero would occur. Instead, another "distance" is chosen that is typical of intervillage distances within the regional system and can be incorporated into the proportional distance scheme. The smaller the distance, the larger the relative proportion or interval, and the greater the likelihood of a contact between contagious and susceptible individuals within the village.

As the epidemic percolates across the village network, a tally is made of newly infected individuals for each village within each time interval. The result is a two-dimensional array with columns representing village nodes and rows representing time intervals. Row totals are the frequencies of new cases for each time period. These frequencies can be arranged into epidemic time-series distributions and graphed.

The model can control individual biological, population, and spatial factors involved with disease spread. By holding one or more of these variables constant, the relative effects of these variables can be determined.

Simulated Epidemic Time Series

Three regional networks and three types of disease characteristics are used to illustrate the results of model simulations. Each network is held constant while the characteristics of the disease agent are allowed to vary in three different ways. The resulting nine simulations are then graphed as time series distributions.

One network representing an actual village system found within a small region, and two model networks are used in these simulation examples (see Figure 1). Each of these networks possesses twenty-four village nodes: 1) *Parish* is based on an actual riverine system found in western Finland, 2) *Ring* is a contrived symmetrical village system which can be used to examine the effects of epidemic spread from two different directions, 3) *Gateway* is an ideal representation of a village system found within a "partially commercialized" or colonial socioeconomic environment (see Hodges 1988).

Each regional system possesses an overall population of 3500 individuals. Ten percent of each of the regional populations is susceptible to infectious disease. Historically, this proportion of susceptible individuals corresponds to slightly less than the proportion of children under five (12% to 16%) found in a study by the author of population registers for an 18th- and 19th-century Finnish parish. The general and susceptible populations are then distributed evenly throughout the network.

In this example, each of the 24 villages has a population of 146 individuals. There are approximately 14 susceptible individuals per village. The susceptible individuals are assigned to households that have an average of 8 individuals. Therefore, each household averages less than 1 susceptible individual. Contacts are made, however, at the village level, and the incorporation of household partitions is meant to reduce the likelihood of unrealistically direct contacts between contagious individuals from one village and susceptible individuals in a neighboring village.

The simulation model, in addition to its ability to use different network representations of village systems, can also vary the transmission characteristics of human diseases. For the purpose of illustration, three contrived disease types are used in the simulation runs: 1) Transmission "efficiency" is .10 (one out of ten contacts results in infection) and is incorporated into the rate of contact between susceptible and contagious individuals. Once infected, the individual has a latency period of four weeks and remains contagious for three weeks; 2) Transmission efficiency is .50. The latency period is one week and the infectious period is two weeks; 3) Transmission efficiency is .10. The latency period is one week and the infectious period is two weeks. For each week, an infectious person can contact any combination of two susceptible, immune, latent, and infectious individuals in either the individual's village of residence or in a neighboring village. It is also possible to change the number of contacts according to regional mobility patterns.

Figures 2-4 depict the effects of the three types of disease characteristics on the three human landscape arrangements (see Figure 1). The effects are represented as weekly times series data aggregated to the regional level.

The biological effects on the time-series distributions are evident in the graphs. Longer transmission chains result in longer epidemic durations. Higher transmission efficiencies result in greater numbers of new cases per week. This also has the effect of

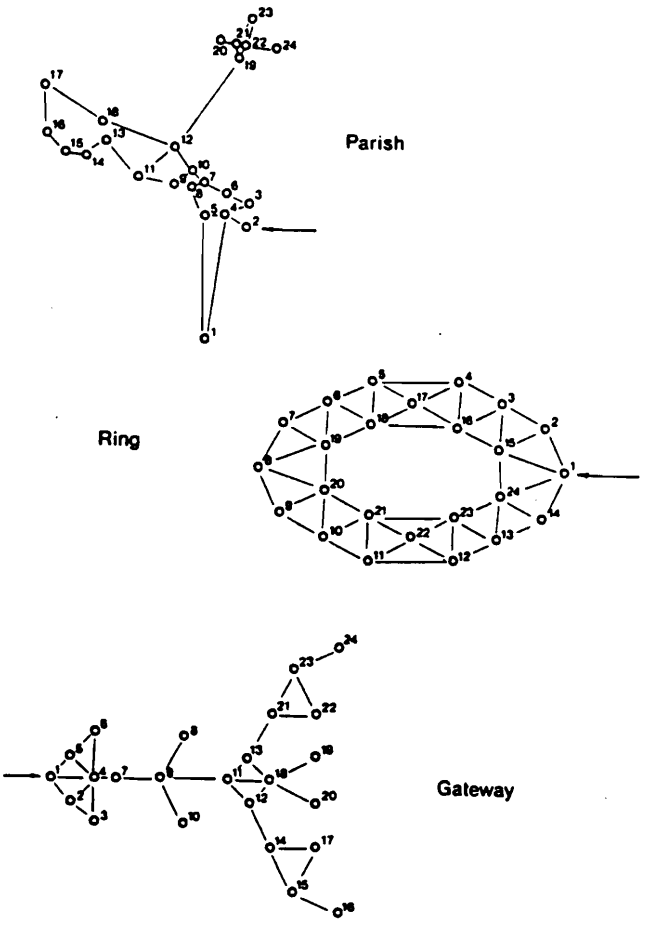

Figure 1: Village Patterns

exhausting village pools of susceptibles at a higher rate, thereby shortening the duration of the epidemic.

The spatial effects of the different human landscapes on the case distributions can also be seen in these figures. Networks that have greater distance variations among villages (e.g., parish and gateway) tend to produce simulated time series distributions that are asymmetrical and are often bimodal. This is the result of the contagion circulating in local village clusters until the susceptible population is exhausted. These distributions are actually a composite of two or more smaller epidemics.

The ring network, which has little intervillage distance variation, produces epidemic patterns that show a unimodal trend. Because contagion spread can proceed more quickly in a more integrated node network such as the ring arrangement, susceptible individuals are contacted at a higher rate, which in turn has the effect of shortening the duration of the epidemic. This fact illustrates the potential for confusing the effects between the biological characteristics of a disease and the spatial properties of the human landscape.

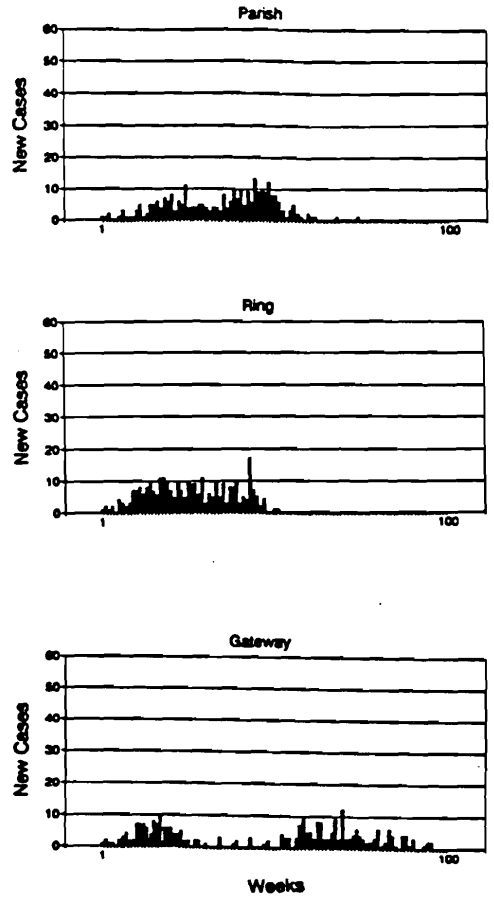

Figure 2: Type 1 Disease

Visual inspection of simulated epidemic time series distributions provides a means for some general interpretations of the effects of both geography and biology on epidemic patterns. It is possible to examine these effects in a more objective manner using descriptive statistics. The next section is an example of this method.

Describing Simulated Epidemic Time Series Distributions

Cliff and Haggett (1982) have proposed a technique for the description and comparison of epidemic patterns based on descriptive statistical parameters. The parameters are derived from the moments of the curve that describes the temporal distribution of new cases and are the familiar mean, standard deviation, skewness, and kurtosis measures. Other useful measures, such as the coefficient of variation, can be derived from these parameters as well. How these parameters and measures can be applied to the analysis of epidemic time series description is summarized below.

The mean of the temporal distribution of new cases is taken to be the average time of epidemic onset. If observations are made in weekly intervals, this measure indicates the average week of infection based on the distribution of epidemic cases. Units of time are still contained in this measure, and so it is not dimensionless. Likewise, the standard

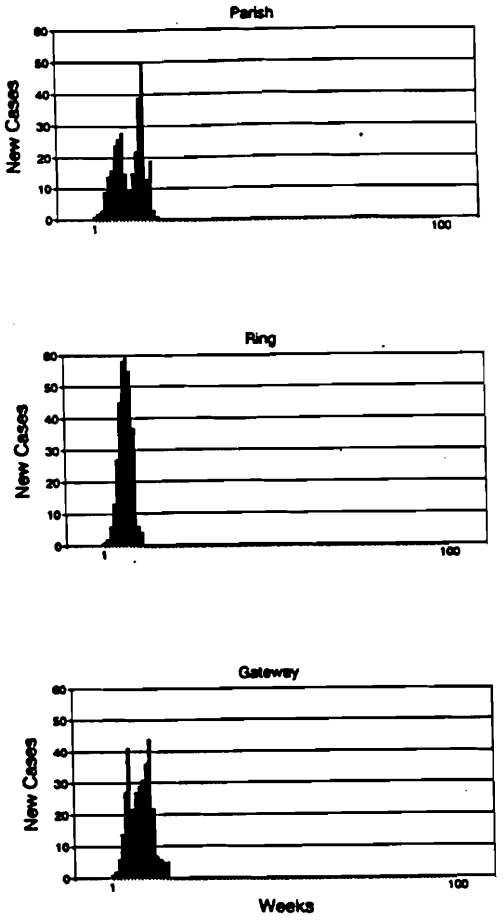

Figure 3: Type 2 Disease

deviation, which measures the dispersal of cases about the mean week of infection, contains time units. Because these two parameters still contain these measurement units in their denominators, and because they are also a function of the overall length of the epidemic, they are not of much value for comparing epidemic time series distributions of different lengths. To accommodate this problem, they can be combined to create a new measure, the coefficient of variation (Cliff and Haggett 1982). This new measure is derived by simply dividing the distribution's standard deviation by the distribution's mean and multiplying the result by 100. The coefficient of variation measures the degree of case dispersal relative to the length of the epidemic. Small values for this measure indicate more temporal clustering of cases, while larger values indicate that cases are dispersed throughout the time series.

Epidemic timing is indicated by the skewness parameter. Using this measure, it is possible to determine if an epidemic is either "fast" or "slow" with respect to where the greatest accumulations of cases are found during the course of the epidemic. Epidemics that appear suddenly in an area and are characterized by rapid accumulations of new cases will have positive skewness values (the distribution mode tends to be to the left of the distribution's mean). Slow epidemics, indicated by negative skewness values, have larger accumulations near the end of their durations. This parameter, used in conjunc-

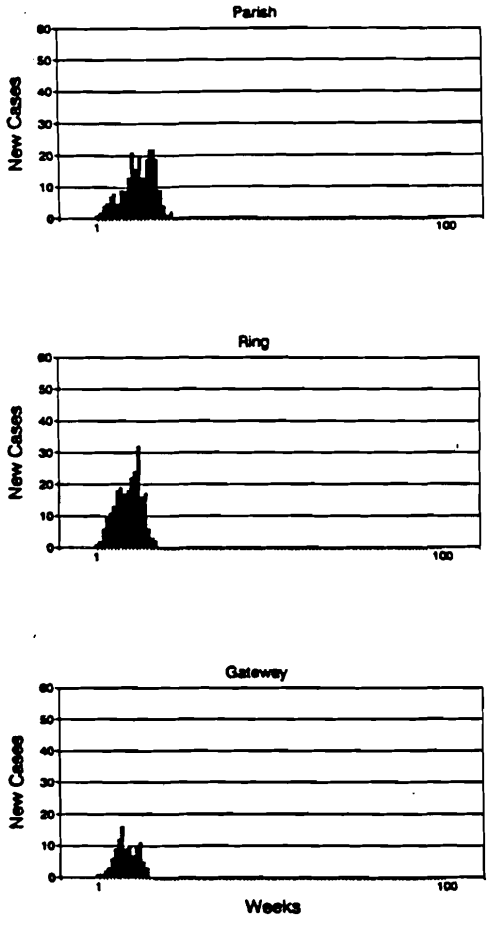

Figure 4: Type 3 Disease

tion with other information, offers the possibility for interpreting which factors, spatial or biological, are affecting epidemic timing.

Kurtosis measures the relative peakedness of a distribution. For epidemic time series, kurtosis indicates the rate of case accumulation. When case accumulation is rapid, the distribution curve becomes more peaked (leptokurtic). If the case accumulation is relatively slow, then the curve will take on a flattened appearance (platykurtic). This measure is useful for comparing the relative velocities or speed of contagion spread within an area.

Table 1 lists the values of the various measures for the sample runs of the epidemic simulation model. The percentages of susceptible individuals infected are also included. Recall that two general factors are affecting these measures: spatial organization (see Figure 1) and biology. The spatial arrangement of susceptibles and their mobility behavior may affect the spatial and temporal patterns of contagion circulation within an area. The length of latency and contagious periods can affect epidemic durations and the number of potential infections. Although the sample size is very small, one can gain a tentative picture of the spatial and biological effects.

Time Series Type	Average Week of Onset	Standard Deviation	Skewness	Kurtosis	Coefficient of Variation	Percent Infected
Parish 1	29.6	13.26	.005	.113	44.8	84
Ring 1	24.3	9.49	.048	.312	39.1	78
Gateway 1	34.7	14.94	-0.276	.081	43.0	79
Parish 2	9.0	3.04	-0.290	.190	33.6	91
Ring 2	7.6	2.27	.010	.700	29.9	90
Gateway 2	12.8	5.61	-0.264	.044	43.7	93
Parish 3	16.9	7.55	-0.213	.094	44.6	65
Ring 3	9.6	3.39	-0.045	.310	35.3	72
Gateway 3	15.3	6.66	-0.179	.119	43.5	64

Table 1: Descriptive Curve Parameters for Simulated Epidemic Time Series

Considering spatial effects, epidemics occurring in the gateway region possess longer durations (average week of onset), regardless of the length of the disease agent transmission chain. In this series of simulation runs, regional networks that possess variable distances among villages and different paths for contagion spread (parish and gateway) have larger coefficient of variation measures. This can be interpreted to mean that local pools of susceptible individuals remained isolated until later in the epidemic. Once infected, the conversion of these individuals has the effect of "spreading out" the distribution of new cases in opposition to temporal clustering. A symmetrical network, such as the ring arrangement, produces epidemic distributions that are symmetrical as indicated by skewness values close to zero. In each case, the ring network also possesses the highest kurtosis values, which suggests that symmetrical networks with low intervillage distance variation are spatial settings conducive to rapid case buildups.

Biological effects can be determined from the percentage of susceptible individuals infected in each of these networks. High transmission efficiencies result in higher percentages of infection. Lower transmission efficiencies coupled with longer transmission chains produce epidemics where smaller percentages of individuals are infected. Longer transmission chains appear to have the effect of maintaining infections longer within a region which also tends to increase the percentage of infected individuals.

Using simple curve descriptive measures such as those described above enables the researcher to compare epidemic patterns from different times and places. It is also possible to sort out and interpret the effects of geography and biology in both simulated and observed epidemics.

Conclusions

The results of this simulation study show that, in theory, the effects of human landscape and biological conditions can be measured and controlled. The simulated epidemics can be viewed as archetypes for comparing historical epidemic series in order to determine which factors may be responsible for the observed pattems. If the spatial organization and mobility patterns of a population are relatively constant over time but the parameters describing the observed time series distribution vary, then it must signify changes in population characteristics or changes in the disease organism, or both. In

addition to recreating possible agent-host coevolution scenarios, this model can be used to simulate epidemics in reconstructed human landscapes (e.g., hearth distributions in archaeological records). Spatial processes of disease can then be simulated for extinct populations.

References

Ackerman, E., L. Elveback, and J. Fox (1984). Simulation of Infectious Disease Epidemics (Springfield, IL: Charles C Thomas).

Angulo, J. (1987). "Interdisciplinary Approaches in Epidemic Studies--II: Four Geographic Models of the Flow of Contagious Disease" Social Science and Medicine 1:57-69.

Cliff, A., and P. Haggett (1982). "Methods for the Measurement of Epidemic Velocity from Time-Series Data" International Journal of Epidemiology 1:82-89.

Haggett, P., A. Cliff, and A. Frey (1977). Locational Models (New York: Halsted Press).

Hodges, R. (1988). Primitive and Peasant Markets (Oxford: Basil Blackwell).

McKeown, T. (1988). The Origins of Human Disease (Oxford: Basil Blackwell).

Unwin, D. 1981. Introductory Spatial Analysis (New York: Methuen).