2003

# Monte Carlo Simulation of a Nonlinear Epidemic

Blair Margaret Bannerman

# Monte Carlo Simulation of a Nonlinear Epidemic

by

Blair M. Bannerman

A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of the requirements of the McDonnell-Barksdale Honors College.

Oxford
May 2003

Approved by

Advisor: Professor Wei-Yin Chen

Reader: Professor Peter Sukanek

Reader: Professor Tristan Denley

ACKNOWLEDGEMENTS

I would like to thank Dr. Wei-Yin Chen for the opportunity to perform this research. I would also like to thank Bill Anderson for his help with various computer language crises. I greatly appreciate Dr. Peter Sukanek and Dr. Tristan Denley taking the time to read this thesis and be on my defense committee. Finally, I would like to thank the staff of the Honors College for all of their help throughout this process.

ABSTRACT

BLAIR M. BANNERMAN: Monte Carlo Simulation of a Nonlinear Epidemic Model

An event-driven Monte Carlo method is used to simulate a simple, nonlinear epidemic. This model illustrates how quickly and to what degree an epidemic spreads through a population. Moreover, it yields information concerning the uncertainties of the epidemic. Some basic assumptions involving the probabilistic dependence of the rate of change of each class in the population must be determined. The rate of infection and the rate of removal from infection are based on case studies found in literary sources. The evolution of the populations is estimated on a time scale that is advanced based on the waiting time. The waiting time, in turn, is estimated by the aforementioned rates and a random number generated by a computer program. The simulation is repeated as many times as there are individuals in the population, so that a mean value and the variance can be determined. Two Fortran 77 computer programs are used to obtain the Monte Carlo information. When the simulation results are compared to the data obtained from literature, it is shown that all of the data points fell within the variance outlined by the simulation. The simulation results compare well with those from the master equation.

## Table of Contents



### *List of Tables*



### *List of Figures*

## Introduction

The spread of epidemics has been a concern throughout scientific history. With current deadly diseases such as AIDS and the Ebola virus, techniques for determining how a disease will be transmitted through a population are eagerly sought. One such technique is the Monte Carlo method. This approach uses stochastic probability modeling and random numbers to predict how a population will change with time once an epidemic is introduced. This method is very simple, yet it can be applied to complex, nonlinear equations, such as those involved in the spread of a disease. The Monte Carlo method has additional benefits in that it can also calculate the variance associated with the predicted mean solution. The variance of the solution is also related to the inherent fluctuations within the system.

This research will examine how a Monte Carlo simulation compares to an actual case of an influenza epidemic spreading though a small population.

## Theory and Literature Review

### Stochastic Method

The stochastic approach to solving complex problems is strongly based on probability. This is in contrast to the deterministic method more familiar to most individuals, which involves using equations and mathematical relationships to give an exact solution. These deterministic equations are generally linear and can require very complex equations. Stochastic modeling results in a probability distribution of the solution to a problem. This method has its advantages because it can be used to model very complex, nonlinear situations.

The probability theory associated with stochastic modeling concentrates on determining how a situation changes with time. In this case, the situation is an epidemic progressing though a population. There are three categories into which individuals in the population can be classified; susceptibles (S), infected (I), and removed (R) classes. The susceptibles category includes all members of the population who are not currently infected but have the potential to become infected. The infected class is simply those in the population who are infected with the disease. The removed class consists of those who have had the disease but are no longer infected. This last class is for those who have either recovered or have succumbed to the illness. In general, a member of the population progresses from the susceptibles to the infected to the removed class, hence the SIR model.

$$S \rightarrow I \rightarrow R$$

Probability is used to estimate the size of these classes at later times in the simulation. The probability of finding a number j of infected individuals in the population at a time t is taken to be $P_j(t)$. At a later time $t + \Delta t$, the number infected is equal to another value n. The relationship between these is

$$P_n(t + \Delta t) = \Sigma [ P_j(t) \cdot W_{j,n}(\Delta t) ] \qquad \text{(Equation 1)}$$

where $W_{j,n}(\Delta t)$ is the conditional probability that n infected individuals exist in the population at time $t + \Delta t$, given that there were j infected present at time t. The probability of $P_j(t)$ and $W_{j,n}(\Delta t)$ is summed over all values of j to find the total probability $P_n$ at time $t + \Delta t$.[1] This total probability can be used to obtain an average solution to how the population will change as the epidemic progresses. A benefit of this method it that the variance associated with this average value can also be predicted for the system.

[1] Stienfeld, Jeffery et al. p. 67-68

Therefore, the internal fluctuations that correspond to the population can be observed. If data appears outside of the variance boundaries, this implies that an external force, which is not accounted for in Monte Carlo simulation, is present.

Monte Carlo Method

The Monte Carlo method is a stochastic model which applies random variables to determine the probability of how a population will change with time. A random variable is a value that is not known specifically; however, the range for the value is known. These random variables and the total probability of a population described by equation 1 are used to simulate the progression of an epidemic.

This method relies on the initial concentrations and known characteristics of the system to simulate the solution to the system as time progresses. The Monte Carlo method has several different approaches, which are dependant on the system being considered. The first classification for determining the most appropriate model is to consider the elements in the system. If the elements of the population being considered are equivalent and change independently of one another, the system is a model A. Model A elements are not dependent on the size of the group being considered. In contrast, model B looks at the change of each element within the system simultaneously as the system changes with time. This model is best for populations in which the probability that an element in the system will change is based on a continuously changing population.

The next consideration when modeling a system is to determine the manner in which time will be updated during the simulation. This is important because most Monte

Carlo simulations concentrate on the change of a population with time. The first method is called time-driven because it regulates the sampling of simulation at certain predetermined time intervals. The time intervals are very small so that it is unlikely that more than one change if any will occur during this time period. An event-driven method can also be used to update the time during a simulation. Using this method, a random time interval is applied to the system during which an event or change is assumed to occur. The nature of this change is determined by the probability of change at the end of the random time interval.

Once the appropriate model has been chosen, one may begin the simulation of the system. The simulation is dependent on certain probability factors or transition intensities. These factors strongly influence whether or not a change will occur and are dependent on the system.

For this example, a model B Monte Carlo simulation is used. The changes in the population are dependant on the size of the three different epidemic categories. The rate of change of the individual classes with time is

$$dS / dt = - r \cdot S \cdot I \qquad \text{(Equation 2)}$$

$$dI / dt = r \cdot S \cdot I - a \cdot I \qquad \text{(Equation 3)}$$

$$dR / dt = a \cdot I \qquad \text{(Equation 4)}$$

where r is the rate of infection and a is the rate of removal[2]. These rates are the transition intensities associated with the population. An event-driven model is chosen to simulate the epidemic.

Another method of simulating the epidemic is using the master equation. This is a deterministic approach which is based on the same basic assumption as Monte Carlo.

[2] Murray, J.D. p. 612

The common point between these methods is the Markov assumption. This assumption essentially states that the current change in a simulation is dependent upon current conditions, not the changes that occurred prior to the current variation.

Gillespie Algorithm

The Gillespie algorithm is a numerical method that uses the Monte Carlo stochastic model to simulate a changing population with time. There are two possible events which can happen to an individual in the population: (1) the change from a susceptible to an infected or (2) the change from the infected to the removed class. This simulation is based on the event probability density function, $P(\tau,i)$, which determines the probability of one specific event i occurring during a specific time interval $\tau$. When this probability density function described by Gillespie is applied to the first event of this epidemic, equation 5 is the resulting probability formula.

$$P(\tau,i) = (r \cdot S \cdot I) \cdot [\exp(-(r \cdot S \cdot I + a \cdot I) \cdot \tau)] \qquad \text{(Equation 5)}$$

This probability can be broken down into two parts; the probability of i occurring, P(i) and the probability of an event occurring at $\tau$, $P(\tau)$.

$$P(\tau,i) = P(\tau) \cdot P(i) \qquad \text{(Equation 6)}$$

From the addition theorem of probability, one can determine $P(\tau)$ as the sum of $P(i, \tau)$ for both events.

$$P(\tau) = \Sigma P(\tau,i) \qquad \text{(Equation 7)}$$

Therefore, P(i) can be expressed by substituting equation 7 into equation 6.

$$P(i) = P(\tau,i) / \Sigma P(\tau,i) \qquad \text{(Equation 8)}$$

One can also substitute equation 5 into equations 7 and 8 to relate the probabilities to the epidemic class sizes and rates of infection and removal.

$$P(\tau) = \exp(-(r \cdot S \cdot I + a \cdot I) \cdot \tau) \qquad \text{(Equation 9)}$$

$$P(i) = (r \cdot S \cdot I) / (r \cdot S \cdot I + a \cdot I) \qquad \text{(Equation 10)}$$

These last two equations are used to update the time by the time interval $\tau$ and to determine which of the two events, a susceptible becoming an infected or an infected becoming a removed, will happen during this time interval. By solving equation 9 for $\tau$ and replacing $P(\tau)$ with a random number, $Rn_1$, one obtains equation 11.

$$\tau = [1 / (r \cdot S \cdot I + a \cdot I)] \cdot \ln (1/Rn_1) \qquad \text{(Equation 11)}$$

In order to determine which event will occur during this time interval, one substitutes P(i) with a second random number, $Rn_2$, in equation 10 and compares this value to the right side of the equation. If the random number is greater than this value, the second event will take place. Otherwise, the first event will transpire.[3]

$$i = 1 \text{ if } Rn_2 < [(r \cdot S \cdot I) / (r \cdot S \cdot I + a \cdot I)]$$

$$i = 2 \text{ if } Rn_2 > [(r \cdot S \cdot I) / (r \cdot S \cdot I + a \cdot I)] \qquad \text{(Equation 12)}$$

Therefore, the basic outline of the Gillespie algorithm is as follows. First, one must obtain the initial sizes of categories, rates of change, and time over which the simulation is to take place. Using this information about the population sizes and rates of change, the probability density function is calculated for the first event. At this point, the two random numbers, each between zero and one, are necessary to determine $\tau$ and which of the two events will transpire. The time must now be updated by $\tau$ and the class sizes adjusted according to which event took place. At this point, the time of the simulation must be compared to the stopping time of the simulation. If the time of the

[3] Stienfeld, Jeffery et al. 98

simulation is less than this stopping time, the probability density function is recalculated with the updated category sizes and the process continues. Otherwise, the simulation ends.[4] This algorithm is presented in a flowchart later in the report. The entire process described above is carried out from a starting time of zero to the stopping time for as many individuals as there are in the population so that an accurate probability may be obtained for the solution.

## Design Plan

### Software

In order to model this simulation quickly, a computer program was designed to follow the Gillespie algorithm described above. The Fortran 77 computer language was used to write the program. This language was appropriate because it contained a random number generation function and it could handle a large number of files. The program is written so that it will perform a simulation of the epidemic from time zero to the specified stopping time and store the generated time, number of susceptibles, and number of infected individuals into three separate matrices. Once these results have been entered into the three matrices, the same program begins to sort the class sizes by the time at which they occurred in the individual simulations. The total time is divided up into 190 increments. The program goes through the time matrix sorting the time and the corresponding classes into the appropriate time files. This process is repeated within the program for as many times as there are individuals in the total population. A second program reads through these time files individually and determines the average and the variance associated with the values in each file. These averages and variances are stored

[4] Stienfeld, Jeffery et al. p. 99

in separate files which can be opened once all of the files have been averaged and the program is complete. The average and variance values are transferred to MS Excel® where they can be plotted.

Since such a large amount of information is involved in this process, it was necessary to use the Mississippi Center for Supercomputing Research (MCSR) so that the program could be run in a short amount of time. Fortunately, MCSR has a Fortran 77 compiler in one of its machines known as sweetgum.

## Base Case

This simulation is based on an influenza epidemic that took place at an English boarding school in 1978. Detailed statistics of how the epidemic spread are outlined in an study in *British Medical Journal*.[5] The influenza epidemic follows the forward progression of S to I to R, which is necessary for this simulation. Also, the population is a constant number and the size of each population is documented as the disease progressed with time. The important information from this article is listed in Table 1 below, and Figure 1 on the following page shows how the three epidemic classes change with time.

Table 1. Influenza information necessary for proper Monte Carlo simulation.

| | |
|---|---|
| Total Population | 763 |
| Initial Number of Susceptibles, So | 762 |
| Initial Number of Infected, Io | 1 |
| Initial Number of Removed, Ro | 0 |
| Rate of Infection, r | 0.00218 / day |
| Rate of removal, a | 0.44036 / day |
| Total Time | 15 days |

[5] British Medical Journal p. 587

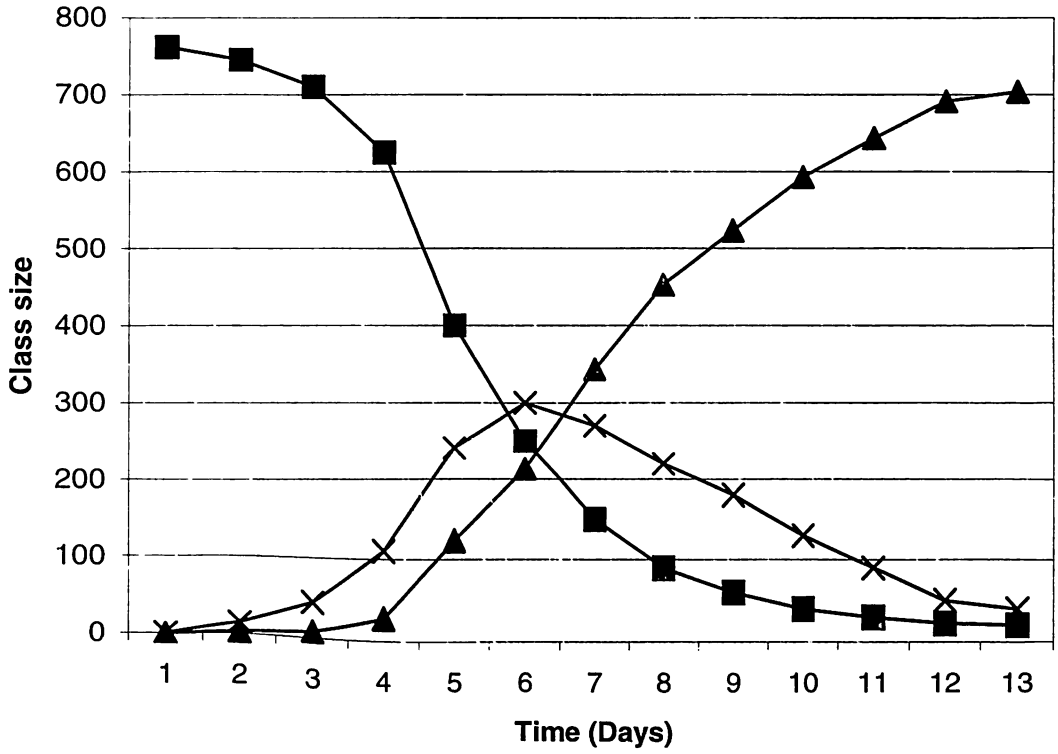
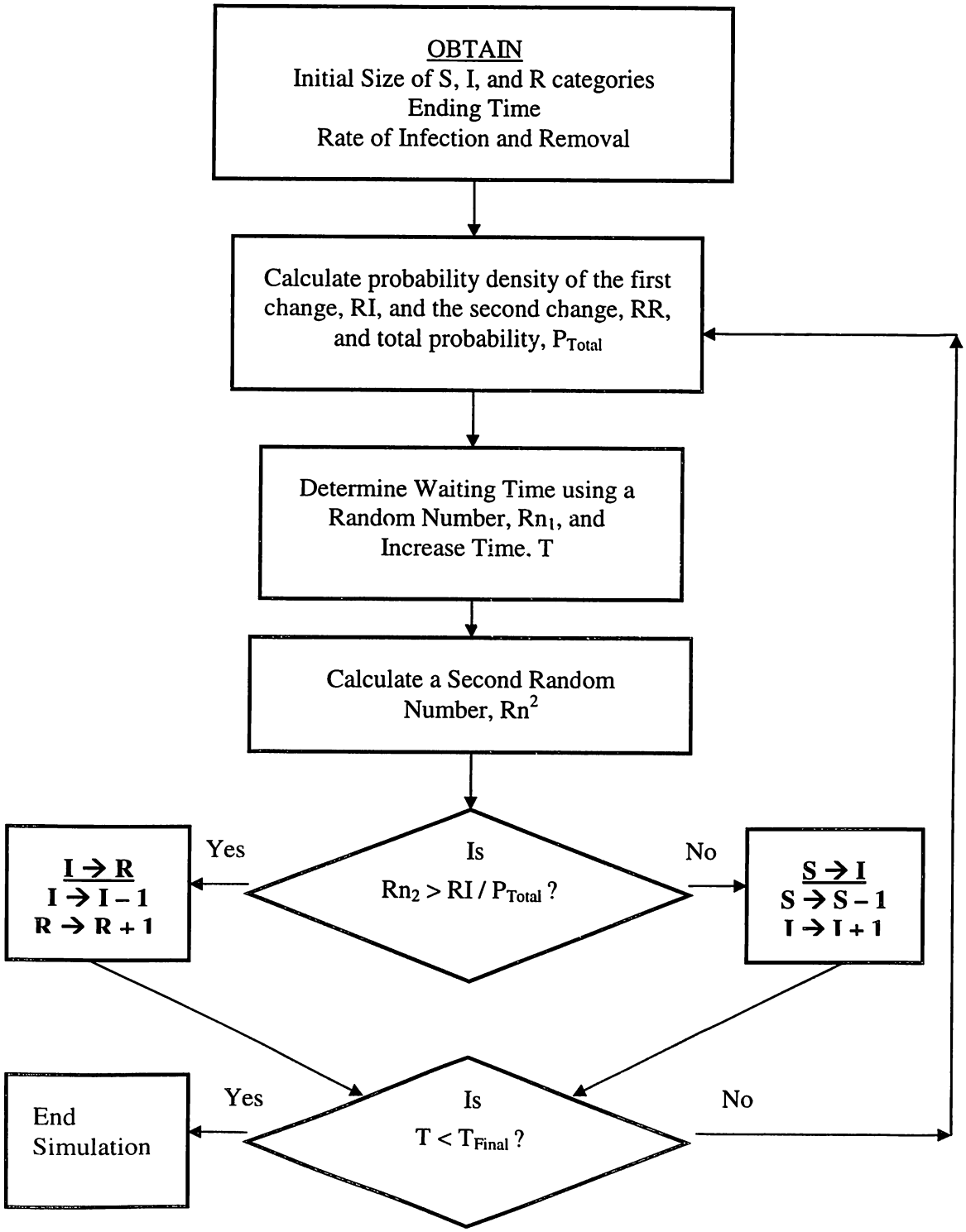



Figure 1. Changes in class sizes vs. time as recorded by the British Medical Journal for a 1978 influenza outbreak in a boys boarding school. (■) Susceptibles (x) Infected (▲) Removed

Flowchart

The following flowchart shows the Gillespie algorithm as applied to the SIR epidemic model.



**Results and Discussion**

Figures 2, 3, and 4 on pages 13, 14, and 15 demonstrate graphically the changes in the susceptibles, infected, and removed classes respectively as determined by the Monte Carlo computer simulation. In order to illustrate the accuracy of the simulation, the actual population sizes as they varied with time are plotted on the same graph as the Monte Carlo results. These graphs show that the actual data fall within the upper and lower variance determined by the simulation in all cases. This implies that only internal variations are present in this epidemic.

The overall trend of the epidemic is for there to be a small period of stagnation where all the classes remain about the same. After this point, the susceptibles class decreases rapidly and the infected class increases in size. The infected class reaches a peak, after which more individuals are entering the removed class than the infected class. Very soon after this peak point, the susceptibles class nears its final value. The infected and removed classes turn toward their final values as well shortly after the susceptibles class.

There are several interesting features in the variance curves, the first of which is the point in the upper variance of the susceptibles category where the initial stagnation time is extended by approximately one day. This extension is present in the infected and removed classes lower variances as well. The susceptibles class lower variance begins to turn toward its final value at least a half day before the mean value. The infected upper variance has two peaks. While the second peak is smaller than the first, it is still significant. This second peak corresponds to the time at which the susceptibles lower variance curve begins its early stagnation toward its final value. The susceptibles curve

decreases rapidly over a period of five days, while the removed class curve increases rapidly over a period of eight days. This is due to the second event of an individual entering the removed class from the infected class becoming the dominant change later in the epidemic time frame.

When the Monte Carlo approach is compared to the results of the master equation, the results are very similar. A graduate student, Sankar Bokka, who was doing similar research at the time of this work, compiled the master equation results. Both methods illustrate the double peak in the infected population and the period of stagnation in the susceptibles class as shown in figures 5 and 6 on pages 16 and 17. The master equation has a much larger variance associated with the information collected for the simulation than the Monte Carlo method. There is a small deviation between the times at which the first upper variance peak occurs. Also the second peak is more apparent in the master equation data. Overall, the Monte Carlo and master equation results compare quite favorably.

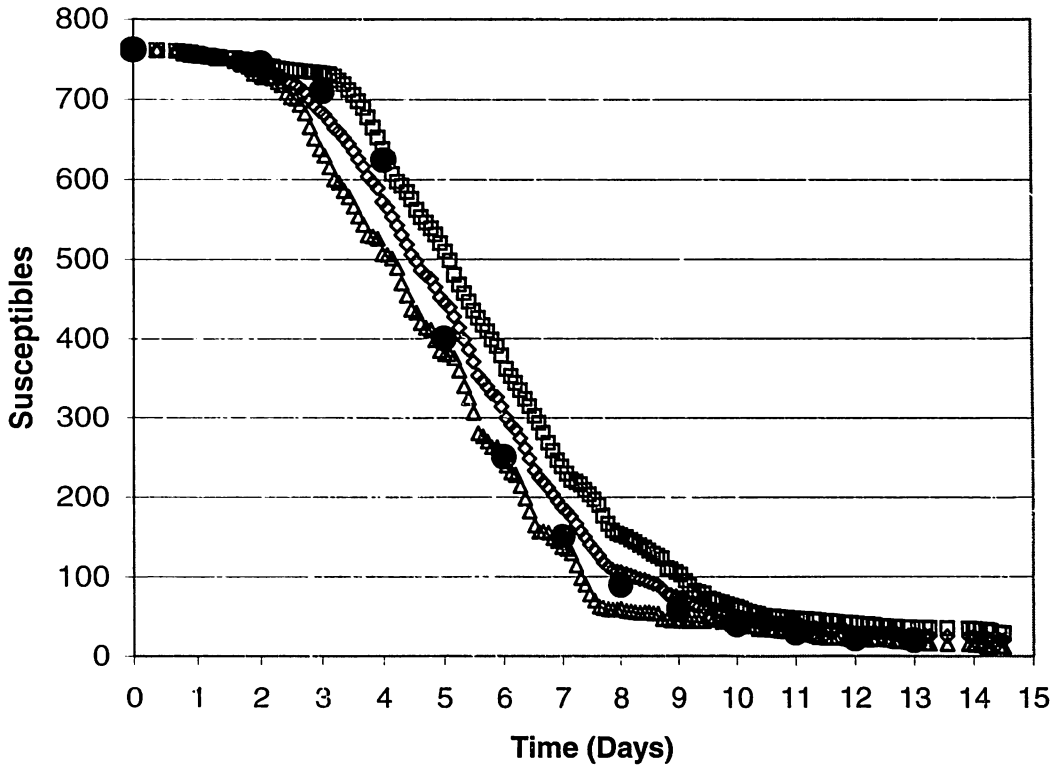



Figure 2. Susceptible class size mean and upper and lower variance as determined by Monte Carlo vs. time, as well as the actual recorded susceptibles class vs. time.

(●) Actual (◊) Mean (□) Upper Variance (Δ) Lower Variance

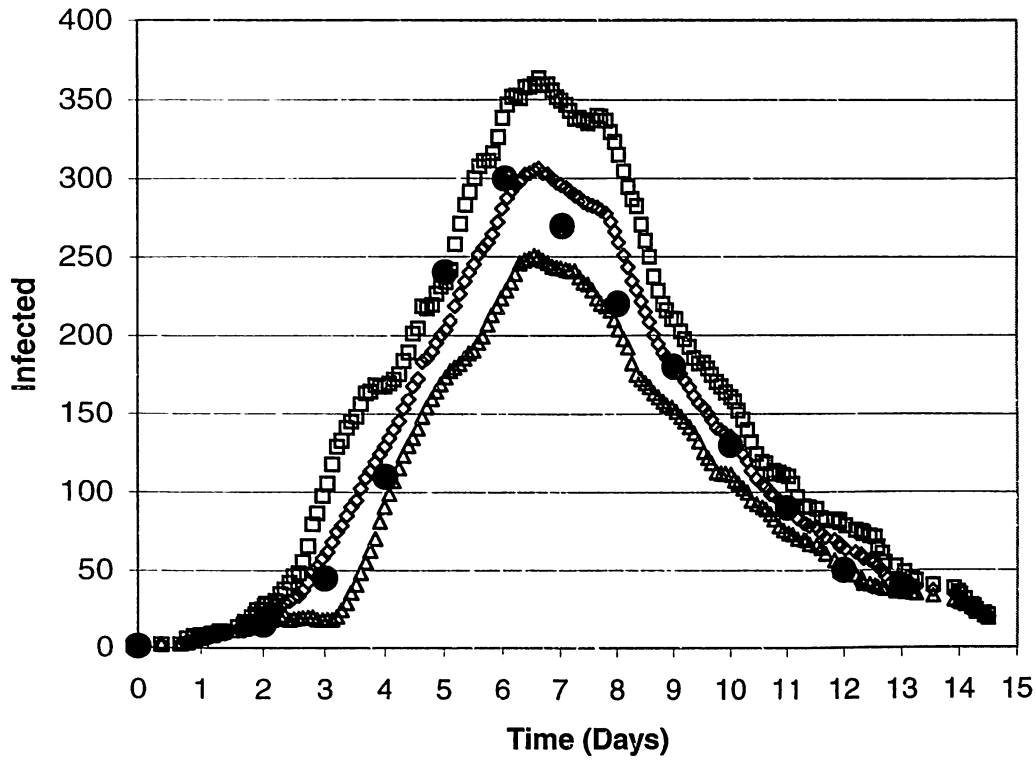



Figure 3. Infected class size mean and upper and lower variance as determined by Monte Carlo vs. time, as well as the actual recorded infected class vs. time.

(●) Actual (◊) Mean (□) Upper Variance (Δ) Lower Variance

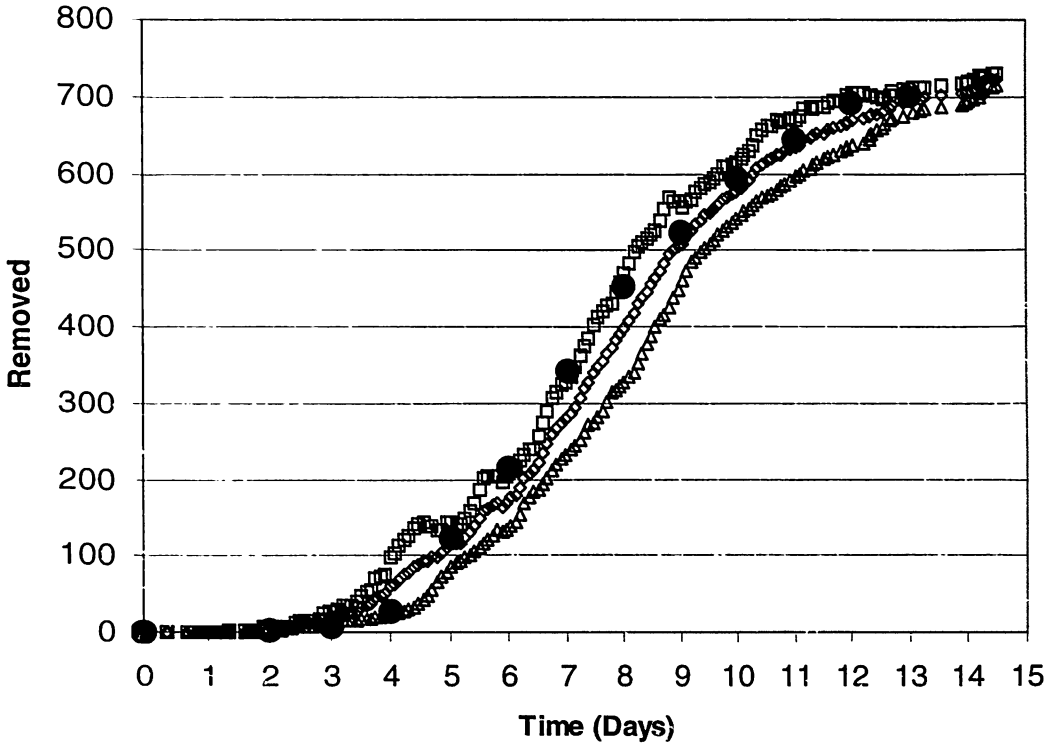



Figure 4. Removed class size mean and upper and lower variance as determined by Monte Carlo vs. time, as well as the actual recorded removed class vs. time.

(●) Actual (◊) Mean (□) Upper Variance (Δ) Lower Variance

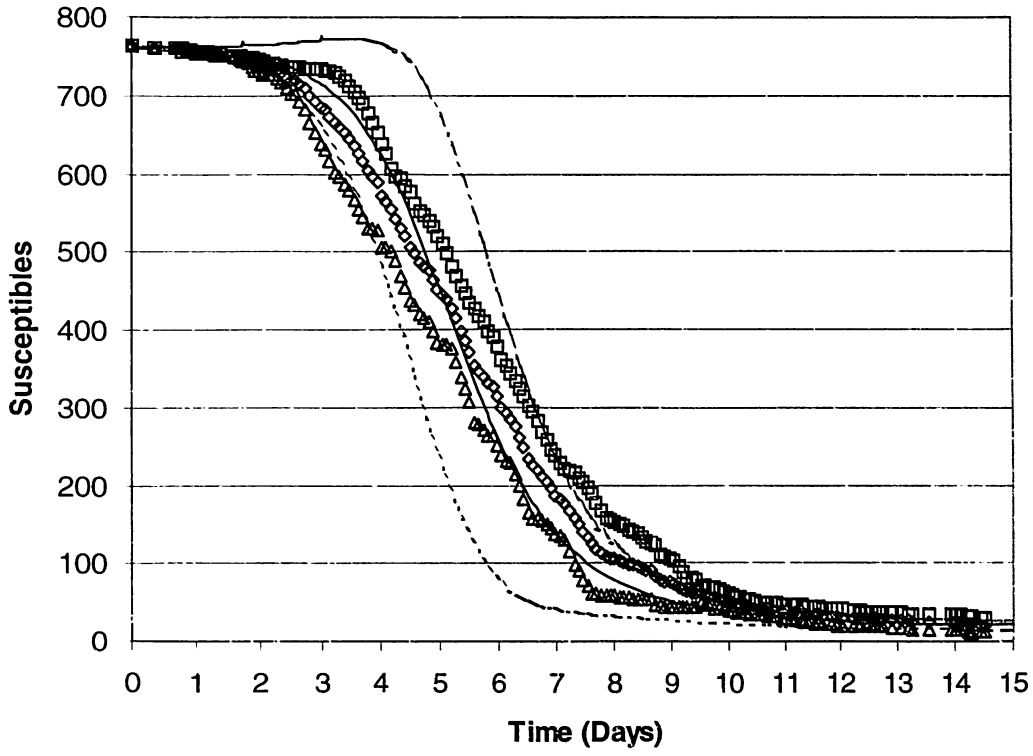



Figure 5. Susceptibles class size mean and upper and lower variance as determined by Monte Carlo (MC) vs. time, as well as the mean and upper and lower variance as determined by the master equation (ME) vs. time.

(◊) Mean - MC (□) Upper Variance - MC (Δ) Lower Variance - MC

(——) Mean - ME (----) Lower Variance - ME (— —) Upper Variance - ME

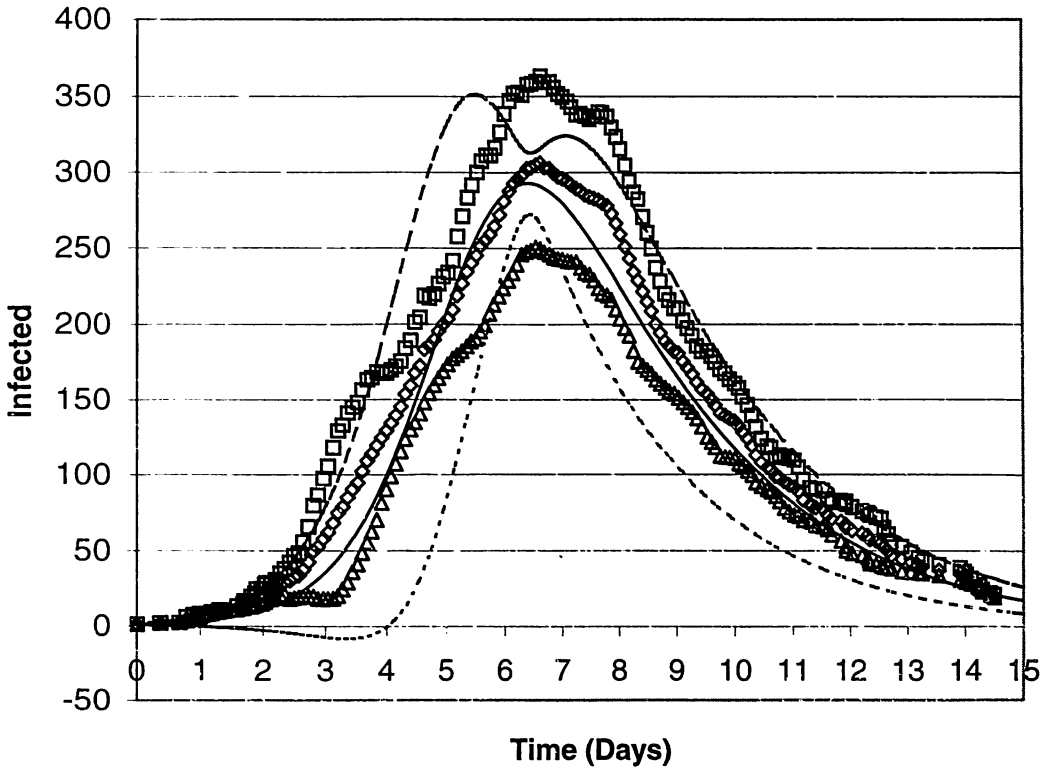



Figure 6. Infected class size mean and upper and lower variance as determined by Monte Carlo (MC) vs. time, as well as the mean and upper and lower variance as determined by the master equation (ME) vs. time.

(◊) Mean - MC (□) Upper Variance - MC (Δ) Lower Variance - MC

(——) Mean - ME (·----) Lower Variance - ME (— —) Upper Variance - ME

## Conclusion

The Monte Carlo modeling proves to be quite accurate when compared to the actual data. The fluctuations of the data are all contained within the upper and lower variance of the various class sizes. This simulation demonstrates the effectiveness of modeling epidemics using the Monte Carlo method. Given the results of this simple case, further research may investigate the application of the Monte Carlo method to more complex diseases. This information could aid in the study of diseases and may also be useful in determining the severity of an epidemic.

When the Monte Carlo results are side by side on the same graph as the master equation results, many of the same characteristics are present in both methods. This is to be expected since the models stem from the same assumptions. The main difference between these methods is their application. The master equation involves solving differential equations simultaneously, while the Monte Carlo method uses random numbers and computer programs. The master equation can be solved quickly, but has a greater variance than the Monte Carlo method.

## References

Murray, J.D. *Mathematical Biology.* Berlin: Springer-Verlag, 1989

Sobol, I.M. *The Monte Carlo Method.* Chicago: University of Chicago Press, 1974.

Stienfeld, Jeffrey, et al. *Chemical Kinetics and Dynamics.* Englewood Cliffs: Prentice Hall, 1989.

*British Medical Journal.* March 4, 1978.

## Appendix A. Computer programs

The following 11 pages contain the two Fortran 77 computer programs necessary for the Monte Carlo simulation of an epidemic model.

```
      PROGRAM SIR
*****************************************************************
*
*     THIS PROGRAM PERFORMS A MONTE CARLO SIMULATION BASED
*     ON USER ENTERED INFORMATION AND AN EVENT DRIVEN MODEL.
*
*****************************************************************
*     DEFINE VARIABLES
*
*     NST          - CONTROL VARIABLE FOR RUNNING PROGRAM
*     COUNT        - DETERMINES WHICH EVENT IN THE SYSTEM OCCURS
*     NSTOP        - DETERMINES WHEN TO STOP THE PROGRAM
*     X            - NUMBER OF TIMES THE SIMULATION IS REPEATED
*                    BASED ON THE NUMBER OF INDIVIUALS IN THE
*                    POPULATION
*     M            - MAXIMUM DATA POINTS RECORDABLE DURING ONE
*                    SIMULATION
*     MNEW         - CURRENT SET OF DATA WITHIN A SIMULTAION BEING
*                    SORTED INTO TIME FILES
*     NT           - TIME FILE UNIT NUMBER TO WHICH DATA IS BEING
*                    ENTERED
*     K            - INDEX VARIABLE USED TO CLOSE SORTED FILE
*     a            - RATE CONSTANTOF REMOVAL
*     r            - RATE CONSTANT OF INFECTION
*     Density      - RATIO OF REMOVAL RATE CONSTANT TO INFECTED
*                    RATE CONSTANT
*     N            - POPULATION SIZE
*     T            - INITIAL TIME OF PROGRAM
*     T2           - STOPPING TIME OF PROGRAM
*     S            - NUMBER OF CURRENT SUSCEPTIBLES IN POPULATION
*     S0           - INITIAL NUMBER OF SUSCEPTIBLES IN THE
*                    POPULATION
*     I0           - INITIAL NUMBER OF INFECTED CLASS IN POPULATION
*     I            - SIZE OF CURRENT INFECTED CLASS IN POPULATION
*     RE           - NUMBER OF CURRENT REMOVED IN THE POPULATION
*     EPS          - CONTROL VARIABLE
*     R1           - RANDOM NUMBER USED TO UPDATE TIME INDEX
*     R2           - RANDOM NUMBER USED TO DECIDE WHICH CHANGE
*                    OCCURS
*     A1           - PROBABILITY OF A SUSCEPTIBLE BECOMING AN
*                    INFECTED
*     A2           - PROBABILITY OF AN INFECTED BECOMING A
*                    REMOVED
*     A0           - TOTAL PROBABILITY OF AN EVENT OCCURING
*     SEED         - NUMBER USED TO CALCULATE RANDOM NUMBERS
*     SEED2        - NUMBER USED BY PROGRAM TO CALCULATE SECOND
```

```
*                        RANDOM NUMBER
*   TDIV               - TIME INCREMENT BETWEEN FILES
*   TNEW               - CURRENT TIME DURING THE SORTING OF GATHERED
*                        DATA INTO THE FILES
*   TIME(1,1450)       - MATRIX TO HOLD TIME DATA GATHERED
*                        DURING ONE SIMULATION
*   SUS(1,1450)        - MATRIX TO HOLD SUSCEPTIBLES DATA
*                        GATHERED DURING ONE SIMULATION
*   INF(1,1450)        - MATRIX TO HOLD INFECTED DATA GATHERED
*                        DURING ONE SIMULATION
*
*************************************************************
*
      INTEGER  NST, COUNT, NSTOP, X, M, MNEW, NT, K
      REAL     a, r, Density, N, T, T2, S, S0, I0, I, RE, EPS, R1,
      REAL     R2, A1, A2, AO, SEED, SEED2, TDIV, TNEW
      REAL     TIME(1,1450), SUS(1,1450), INF(1,1450)
*
*     OPEN A SET OF TIME FILES TO HOLD DATA GENERATED BY THE
*     SIMULATION
*
      OPEN (UNIT=12,FILE='/tmp/sirnew/TIME01',STATUS='NEW')
      OPEN (UNIT=13,FILE='/tmp/sirnew/TIME02',STATUS='NEW')
      OPEN (UNIT=14,FILE='/tmp/sirnew/TIME03',STATUS='NEW')
      OPEN (UNIT=15,FILE='/tmp/sirnew/TIME04',STATUS='NEW')
*
*     THE MAXIMUM NUMBER OF FILES WHICH COULD BE OPENED IS 190.
*     THESE FILES ARE OPENED IN THE TEMPERARY SECTION OF
*     SWEETGUM TO ALLOW FOR THEIR LARGE SIZE
*
      OPEN (UNIT=199,FILE='/tmp/sirnew/TIM188',STATUS='NEW')
      OPEN (UNIT=200,FILE='/tmp/sirnew/TIM189',STATUS='NEW')
      OPEN (UNIT=201,FILE='/tmp/sirnew/TIM190',STATUS='NEW')
*
*     GET NECESSARY INFORMATION ABOUT THE EPIDEMIC FROM THE USER
*
5     PRINT*, "r= ? Density= ? N= ? I=?"
      READ*, r, Density, N, I
*
*     FOR THIS EPIDEMIC THE CONDITIONS ARE AS FOLLOWS
*
      r = .00218
      Density = 202
      N = 763
      I = 1
*
```

```
      a = r * Density
*
*     INITIALIZE CATEGORY SIZES
*
      I0 = I
      S0 = N - I
      S = N – I
      RE = 0
*
*     SET CONTROL VARIABLE WITHIN PROGRAM TO STOP PROGRAM IF
*     ERROR OCCURS
*
      EPS = (1/(10*10*10*10*10))
*
*     PRINT VALUES ENTERED BY THE USER TO CONFIRM
*
      PRINT*, "EPS = ", EPS
      PRINT 101, r,a,S,I
101     FORMAT (F9.6, 1X, F9.7, 1X, F10.4, 1X, F9.4)
*
*     GET NECESSARY TIME INFORMATION FROM THE USER
*
      PRINT*, "INITAL TIME, To? STOPPING TIME, T2?"
      READ*, T,T2
*
*     PRINT DATA TO CONFIRM
*
      PRINT*, "T= ",T,"   T2= ",T2
*
*     FOR THIS EPIDEMIC THE TIME DATA IS AS FOLLOWS
*
      T = 0.0
      T2 = 15
*
*     OBTAIN A NUMBER TO CALCUALTED THE RANDOM NUMBERS NEEDED
*     FOR THIS PROGRAM
*
      PRINT*, "ENTER A 6 DIGIT, ODD INTEGER"
      READ*, SEED
*
*     ALLOW THE USER TO RUN THE PROGRAM IF ENTERED DATA IS AS
*     THEY WISH
*
      PRINT*, "EXECUTE PROGRAM? (Y=1,N=0)"
      READ*, NST
*
```

```
*   IF THE USER DOES NOT WISH TO EXECUTE THE PROGRAM, ALLOW
*   THEM TO END THE PROGRAM
*
      IF (NST .EQ. 0) GO TO 50
*
*   BEGIN SIMULATION OF THE EPIDEMIC FOR THE NUMBER OF
*   INDIVIDUALS IN THE POPULATION
*
      X = 1
60    IF (X .LE. 763) THEN
        M = 1
        SEED = SEED + 1372
*
*   ENTER INITIAL DATA INTO THE FIRST ROW OF EACH MATRIX
*
        I = 1.0
        INF(X,M) = I
        S = 762.0
        SUS(X,M) = S
        T = 0.0
        TIME(X,M) = T
*
*   CALCULATE THE RATE ASSOCIATED WITH EACH EVENT OCCURING
*
10      A1 = r * S * I
        A2 = a * I
        A0 = (r * S * I) + (a * I)
*
*   IF AN ERROR HAS OCCURED ALLOW THE PROGRMA TO QUIT
*
        IF (A0 .LT. EPS) GO TO 50
*
*   DETERMINE THE RANDOM NUMBERS FOR THE PROGRAM
*
        R1 = RAN(SEED)
        SEED2 = SEED + 202
        R2 = RAN(SEED2)
*
*   UPDATE THE TIME BY AN AMOUNT BASED ON THE FIRST RANDOM
*   NUMBER AND THE TOTAL PROBABILITY
*
21      T = T + (LOG(1./R1))/A0
*
*   CONTROL VARIABLE TO END PROGRAM IF TIME HAS INCREASED
*   BEYOND STOPPING TIME
*
```

```
      IF (T .LT. T2) THEN
        GO TO 25
      END IF
*
*   USE THE SECOND RANDOM NUMBER AND THE TOTAL PROBABILITY
*   TO DETERMINE WHICH EVENT OCCURS
*
25    PRINT*, "R2 =", R2, "T= ", T
      IF (R2 .GT. (A1/A0)) THEN
        COUNT = 2
      ELSE
        COUNT = 1
      END IF
*
*   IF THE FIRST EVENT OCCURS, DECREASE THE SUSCEPTIBLE SIZE AND
*   INCREASE THE INFECTED CLASS.  IF THE SECOND OCCURS, DECREASE
*   THE INFECTED CATEGORY AND INCREASE THE REMOVED CLASS
*
      IF (COUNT .EQ. 1) THEN
        S = S - 1
        I = I + 1
        PRINT*, "COUNT1S = ", S, "COUNT1I = ", I
        GO TO 40
      END IF
      IF (COUNT .EQ. 2) THEN
        I = I - 1
        RE = RE + 1
        PRINT*, "COUNT2S = ", S, "COUNT2I = ", I
        GO TO 40
      END IF
*
*   SEND THE DATA OBTAINED FOR THIS TIME VALUE TO THE
*   APPROPRIATE MATRIX.  THE PROCESS CONTINUES UNTIL THE
*   STOPPING TIME IS REACHED
*
40    IF (T .LE. T2)  THEN
        M = M + 1
        TIME(1,M) = T
        SUS(1,M) = S
        INF(1,M) = I
        PRINT*, "T=", T, "S=", S, "I=", I, "X = ",X,"M= ",M
        GO TO 10
      ELSE
        X = X + 1
        GO TO 65
      END IF
```

```
      END IF
*
*     THE USER IS ALLOWED TO STOP THE PROGRAM IF THEY ARE
*     SATISFIED WITH THE DATA OBTAINED
*
50    PRINT*, "STOP? (YES=1, NO=0)"
      READ*, NSTOP
*
*     IF THE USER WISHES TO RERUN THE DATA, THEY MAY DO SO
*     THROUGH THE IF – STATEMENT BELOW
*
      IF (NSTOP .EQ. 0) THEN
         GO TO 5
      ELSE
         GO TO 90
      END IF
*
*     NOW THE RESULTS FROM ALL OF THE TOTAL NUMBER OF
SIMULATIONS
*     IS SORTED INTO FILES BASED ON THE TIME TO WHICH THE
*     SUSCEPTIBLE AND INFECTED CLASS DATA CORRESPOND
*
65    TDIV = T2/190
*
*     THIS STARTS A LOOP WHICH WILL READ THROUGH EACH OF THE
*     COLUMNS OF THE MATRIX
*
         NT = 12
         TNEW = TDIV
         MNEW = 1
*
*     THIS BEGINS A LOOP WHICH WILL READ THROUGH EACH ROW IN ONE
*     COLUMN OF THE MATRIX
*
85       IF (MNEW .LE. 1450) THEN
            IF (TIME(1,MNEW) .GE. T2) THEN
               GO TO 60
            END IF
*
*     ENTER TIME, SUSCEPTIBLE AND INFECT VALUE INTO CURRENT FILE OR
*     MOVE TO NEXT FILE IF TIME VALUE IS GREATER THAN THE FILE LIMIT
*
            IF (TIME(1,MNEW) .LE. TNEW) THEN
              WRITE (NT,*) TIME(1,MNEW), SUS(1,MNEW), INF(1,MNEW)
               MNEW = MNEW +1
               GO TO 85
```

```
            END IF
*
            IF (TIME(1,MNEW) .GT. TNEW) THEN
               TNEW = TNEW + TDIV
                NT = NT + 1
               WRITE (NT,*) TIME(1,MNEW), SUS(1,MNEW), INF(1,MNEW)
               MNEW = MNEW + 1
               GO TO 85
            END IF
         ELSE
            GO TO 60
         END IF
*
      ELSE
         GO TO 90
      END IF
*
*   CLOSE FILES CONTAINED SORTED DATA FROM SIMUTLATION
*
90    PRINT*, "GOOD"
      DO 95 K = 12,201
         CLOSE(UNIT=K,STATUS='KEEP')
95    CONTINUE
*
      STOP
      END
```

```
      PROGRAM AVERAGE
*************************************************************************
*     THIS PROGRAM WILL READ THE DATA IN THE SORTED FILES INTO AN
*     ARRAY AND THEN FIND THE AVERAGE AND VARIANCE OF THAT
*     ARRAY.  THE AVERAGE AND VARIANCE WILL THEN BE READ INTO
*     ANOTHER FILE, SO THAT IT MAY BE TRANSFERRED TO A GRAPHING
*     PROGRAM LIKE MS EXCEL
*
*************************************************************************
*
*     DEFINE VARIABLES
*
*     AT          - INDEX VARIABLE FOR OPENING SORTED DATA FILES
*     COUNT       - THE NUMBER OF SET OF VALUES IN THE FILE
*     K           - INDEX VARIABLE FOR NAMES OF SORTED DATA FILES
*     SUM         - SUM OF TIME VALUES WITHIN A DATA FILE, USED
*                   TO DETERMINE AVERAGE TIME VALUE
*     SUMX        - SUM OF SUSCEPTIBLE VALUES WITHIN A FILE, USED
*                   TO DETERMINE AVERAGE SUSCEPTIBLE VALUE
*     SUMY        - SUM OF INFECTED VALUES WITHIN A FILE, USED
*                   TO DETERMINE AVERAGE INFECTED VALUE
*     AVET        - AVERAGE OF TIME ELEMENTS IN A SINGLE FILE
*     AVEX        - AVERAGE OF SUSCEPTIBLE POPULATION IN A SINGLE
*                   FILE
*     AVEY        - AVERAGE OF INFECTED POPULATION IN A SINGLE FILE
*     TIME        - TIME OF A SINGLE SAMPLING DURING A SIMULATION
*     X           - SUSCEPTIBLES POPULATION AT TIME
*     Y           - INFECTED POPULATION AT TIME
*     SQX         - USED TO CALCULATE SQUARE OF DEVIATION OF
*                   SUSCEPTIBLE VALUE FROM AVERAGE
*     SQY         - USED TO CALCULATE SQUARE OF DEVIATION OF
*                   INFECTED VALUE FROM AVERAGE
*     SUMX2       - SUM OF SQX WITHIN A FILE, USED TO DETERMINE
*                   VARIANCE SUSCEPTIBLE VALUE
*     SUMY2       - SUM OF SQY WITHIN A FILE, USED TO DETERMINE
*                   VARIANCE INFECTED VALUE
*     AVEUX       - AVERAGE VALUE OF SUSCEPTIBLES POPULATION PLUS
*                   THE VARIANCE OF THIS POPULATION
*     AVELX       - AVERAGE VALUE OF SUSCEPTIBLES POPULATION MINUS
*                   THE VARIANCE OF THIS POPULATION
*     AVEUY       - AVERAGE VALUE OF INFECTED POPULATION PLUS THE
*                   VARIANCE OF THIS POPULATION
*     AVELY       - AVERAGE VALUE OF INFECTED POPULATION MINUS THE
*                   VARIANCE OF THIS POPULATION
*     J(190)      - ARRAY CONTAINING NAMES OF SORTED FILES
*
```

```
*************************************************************************
*
      INTEGER          AT, COUNT, K
      REAL             SUM, SUMX, SUMY, AVET, AVEX, AVEY, TIME, X, Y,
                      SQX, SQY, SUMX2, SUMY2, AVEUX, AVELX, AVEUY,
                      AVELY
      CHARACTER*6  J(190)
*
*************************************************************************
*
*     OPEN THE 190 SORTED FILES FOR ANALYSIS IN THIS PROGRAM
*
      OPEN (UNIT=12,FILE='tmp/sirnew/TIME01',STATUS='OLD')
      OPEN (UNIT=13,FILE='tmp/sirnew/TIME02',STATUS='OLD')
      OPEN (UNIT=14,FILE='tmp/sirnew/TIME03',STATUS='OLD')
*
      OPEN (UNIT=199,FILE='tmp/sirnew/TIM188',STATUS='OLD')
      OPEN (UNIT=200,FILE='tmp/sirnew/TIM189',STATUS='OLD')
      OPEN (UNIT=201,FILE='tmp/sirnew/TIM190',STATUS='OLD')
*
*     OPEN NEW FILES TO HOLD THE AVERAGE AND VARIANCE VALUES OF
*     THE POPULATIONS IN EACH OF THE 190 TIME DIVISTIONS
*
      OPEN (UNIT=312,FILE='AVET',STATUS='NEW')
      OPEN (UNIT=313,FILE='AVETU',STATUS='NEW')
      OPEN (UNIT=314,FILE='AVETL',STATUS='NEW')
      OPEN (UNIT=315,FILE='AVEX',STATUS='NEW')
      OPEN (UNIT=316,FILE='AVEXU',STATUS='NEW')
      OPEN (UNIT=317,FILE='AVEXL',STATUS='NEW')
      OPEN (UNIT=318,FILE='AVEY',STATUS='NEW')
      OPEN (UNIT=319,FILE='AVEYU',STATUS='NEW')
      OPEN (UNIT=320,FILE='AVEYL',STATUS='NEW')
*
*************************************************************************
*
*     DEFINE THE ARRAY WHICH WILL HOLD THE NAMES OF THE SORTED
*     FILES
*
      AT = 12
      J(1) = 'tmp/sirnew/TIME01'
      J(2) = 'tmp/sirnew/TIME02'
      J(3) = 'tmp/sirnew/TIME03'
*
      J(188) = 'tmp/sirnew/TIM188'
      J(189) = 'tmp/sirnew/TIM189'
      J(190) = 'tmp/sirnew/TIM190'
```

```
*
*    THIS STARTS A LOOP WHICH WILL DETERMINE THE AVERAGE AND
*    VARIANCE WITHIN EACH OF THE FILES
*
5     SUM = 0
      COUNT = 0
      SUMX = 0
      SUMY = 0
*
*    AVERAGE THE TIME, SUSCEPTIBLES, AND INFECTED VALUES WITHIN
*    ONE TIME INTERVAL FILE
*
      IF (AT .LE. 201) THEN
10       READ (AT,*,END=15) TIME, X, Y
         SUM = SUM + TIME
         COUNT = COUNT + 1
         SUMX = SUMX + X
         SUMY = SUMY + Y
         GO TO 10
15       AVET = SUM / REAL(COUNT)
         AVEX = SUMX / REAL(COUNT)
         AVEY = SUMY / REAL(COUNT)
*
         CLOSE (UNIT=AT,STATUS='KEEP')
         K = AT - 11
*
         OPEN (UNIT=AT,FILE=J(K),STATUS='OLD')
         GO TO 30
*
*    FIND THE VARIANCE OF THE SUSCEPTIBLES AND INFECTED
*    POPULATION IN ONE TIME INTERVAL FILE
*
30       SUMX2 = 0
         SUMY2 = 0
40       READ (AT,*,END=45) TIME, X, Y
         SQX = (X - AVEX) * (X - AVEX)
         SUMX2 = SQRT(SQX) + SUMX2
         SQY = (Y - AVEY) * (Y - AVEY)
         SUMY2 = SQRT(SQY) + SUMY2
         GO TO 40
45       AVEUX = (SUMX2 / REAL(COUNT)) + AVEX
         AVELX = AVEX - (SUMX2 / REAL(COUNT))
         AVEUY = (SUMY2 / REAL(COUNT)) + AVEY
         AVELY = AVEY - (SUMY2 / REAL(COUNT))
*
         CLOSE (UNIT=AT,STATUS='KEEP')
```

```
*
*     WRITE THESE AVERAGES AND VARIANCES TO THEIR RESPECTIVE
*     FILES
*
          WRITE (312,*) AVET
          WRITE (313,*) AVEX
          WRITE (314,*) AVEUX
          WRITE (315,*) AVELX
          WRITE (316,*) AVEY
          WRITE (317,*) AVELY
          WRITE (318,*) AVEUY
          AT = AT + 1
        GO TO 5
      END IF
*
*     CLOSE THESE FILES SO THAT THEY MAY BE OPENED LATER AND
*     GRAPHED IN MS EXCEL®
*
      CLOSE(UNIT=312,STATUS='KEEP')
      CLOSE(UNIT=313,STATUS='KEEP')
      CLOSE(UNIT=314,STATUS='KEEP')
      CLOSE(UNIT=315,STATUS='KEEP')
      CLOSE(UNIT=316,STATUS='KEEP')
      CLOSE(UNIT=317,STATUS='KEEP')
      CLOSE(UNIT=318,STATUS='KEEP')
*
      STOP
      END
```