A Glimpse of Representing Stochastic Processes

Nathaniel Osgood

CMPT 858

March 22, 2011

Recall: Project Guidelines

- Creating one or more simulation models.
- Placing data into the model to customize it to a particular context (e.g. to a particular region).
- Running a “baseline” scenario with the existing model parameters, and commenting on its plausibility.
- Running one or more “what if” scenarios with the model to explore different possible situations. These situations could reflect the results of implementing different policies, or different possible external conditions.
- Performing one or more “sensitivity analyses”, in which assumptions in the model (in the form of parameter values, or structural elements of the model) are changed.
- A well-structured written report describing the above, and your findings.
- Report Due date: April 24

What I'd like to See in the Report

- Introduction
- Background
- Comment on baseline scenario choice & plausibility of model results
- Sensitivity analyses (Parameters and/or structural)
- Investigation of "what if" scenarios
  - Potential policies
  - External conditions
- Model limitations & ideas for possible extensions
- Process learning
- Implications for our understanding of the world

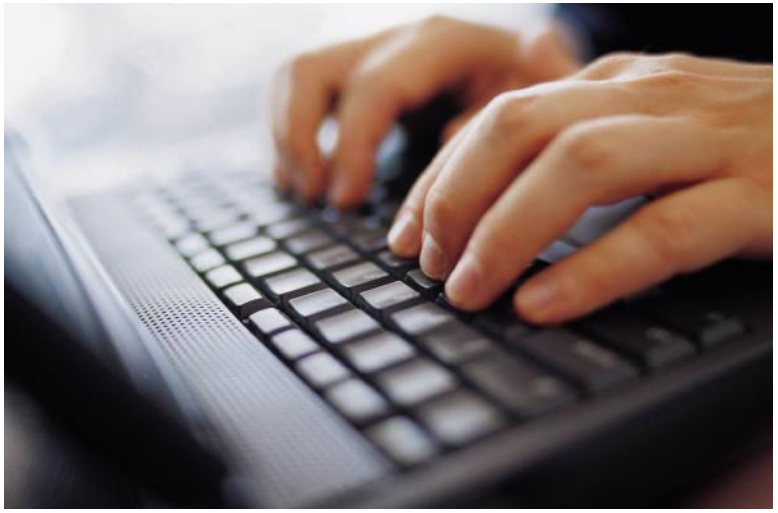
Project Presentations

- Presentations will be 45 minutes in length
- Seeking 2 half days of presentations
- Encourage all students to attend
- Likely scheduling: Late in week of April 24 (just after report due)
  - afternoon of April 28 & 29?

Monte Carlo Analyses in AnyLogic

- When running Monte Carlo analysis, we'd like to summarize the results of multiple runs
- One option would be to display each trajectory over time; downside: quickly gets messy
- AnyLogic's solution
  - Accumulate data regarding how many trajectories fall within given areas of value for a given interval of time using a "Histogram2D Data"
  - Display the Histogram2D Chart

Hands on Model Use Ahead

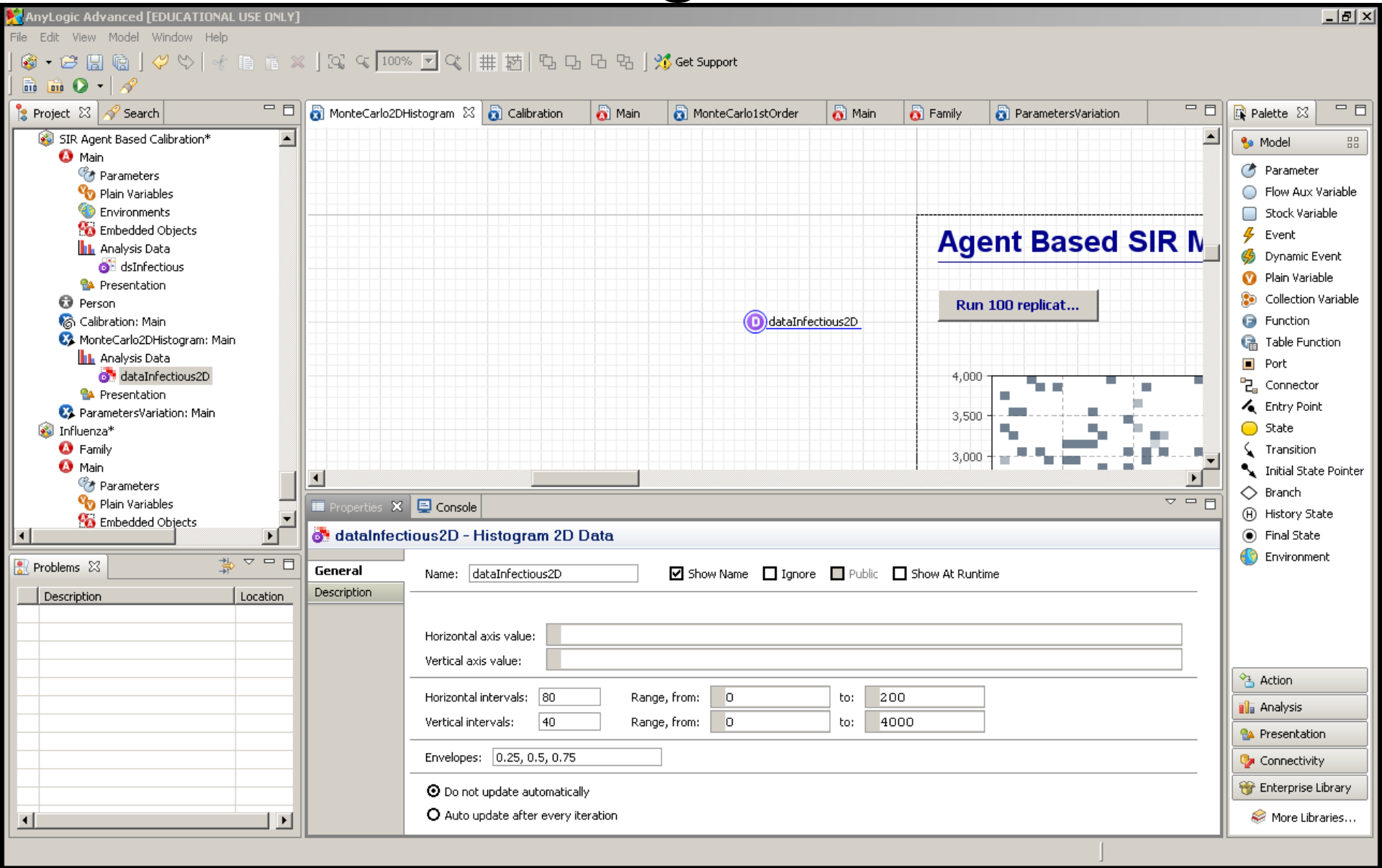


Load Sample Model:

SIR Agent Based Calibration

(Via “Sample Models” under “Help” Menu)

2D Histogram Data



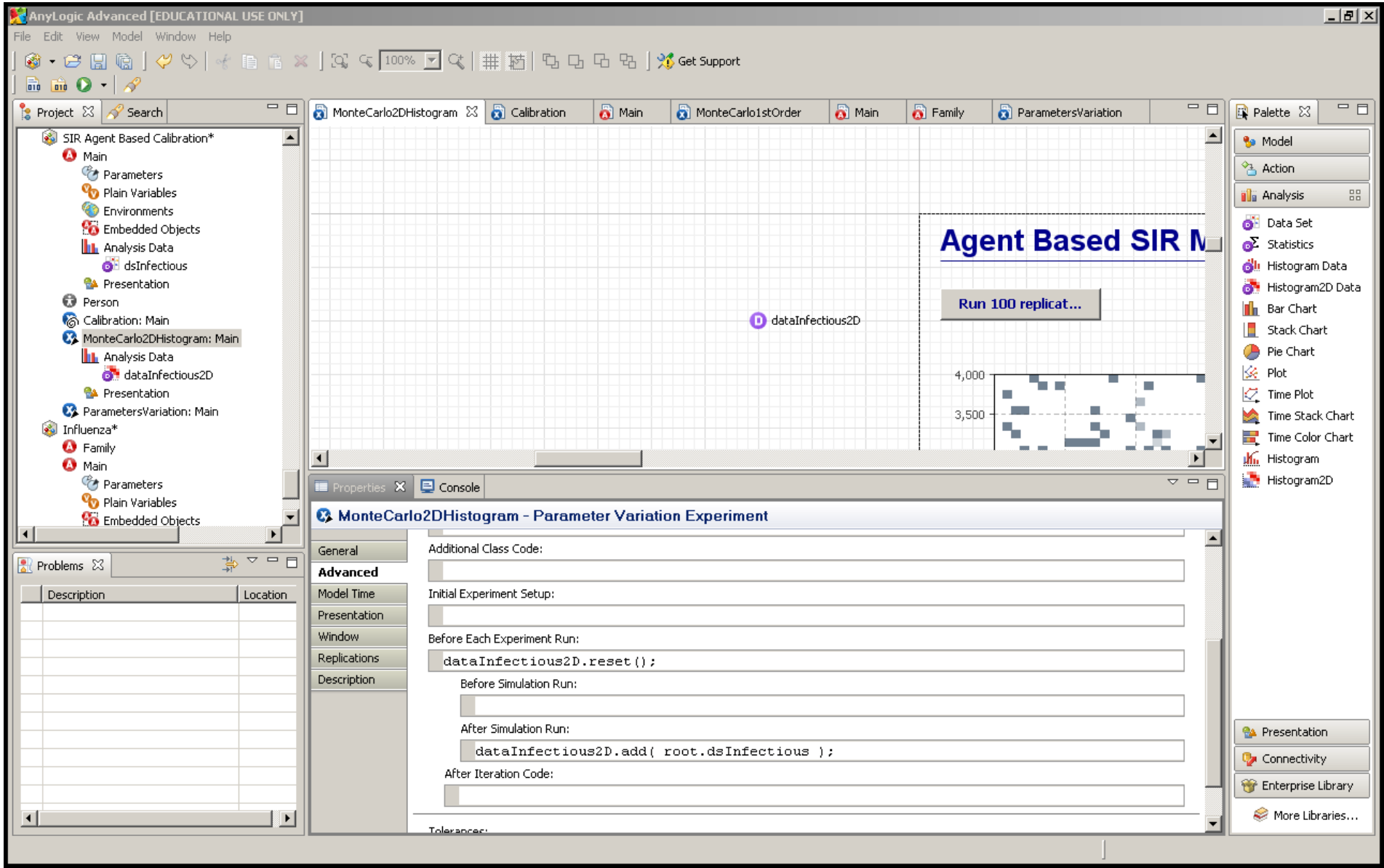
Important Distinction (Declining Order of Aggregation)

- Experiment
  - Collection of simulation
- Simulation
  - Collection of replications that can yield findings across set of replications (e.g. mean value)
- Replication
  - One run of the model

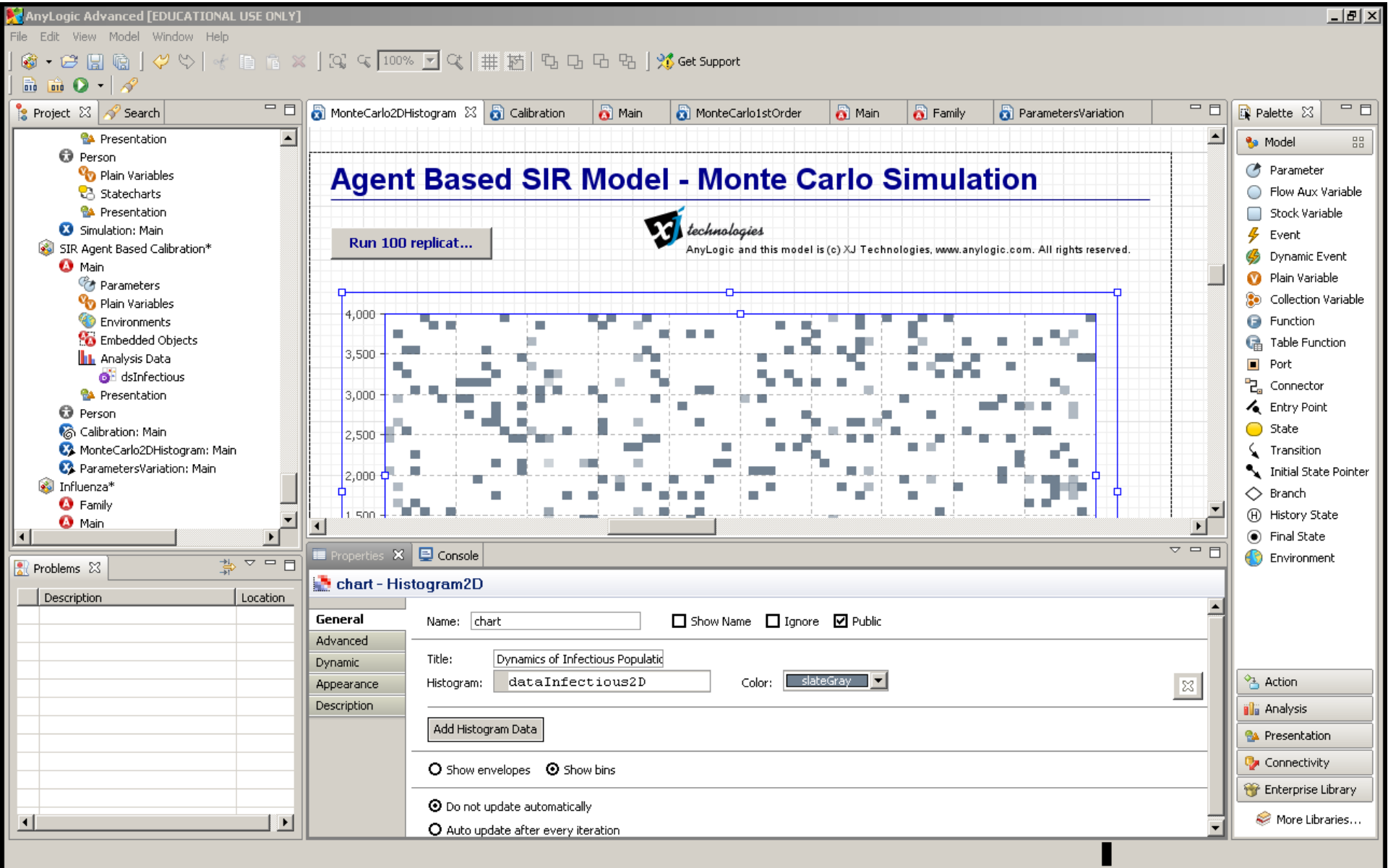
Flexibility Typically Ignored

- In most AnyLogic models, an Experiment is composed of a single Simulation, which is composed of a single Replication
- In most AnyLogic models which run “ensembles” of realizations, a simulation is composed of only a single realization

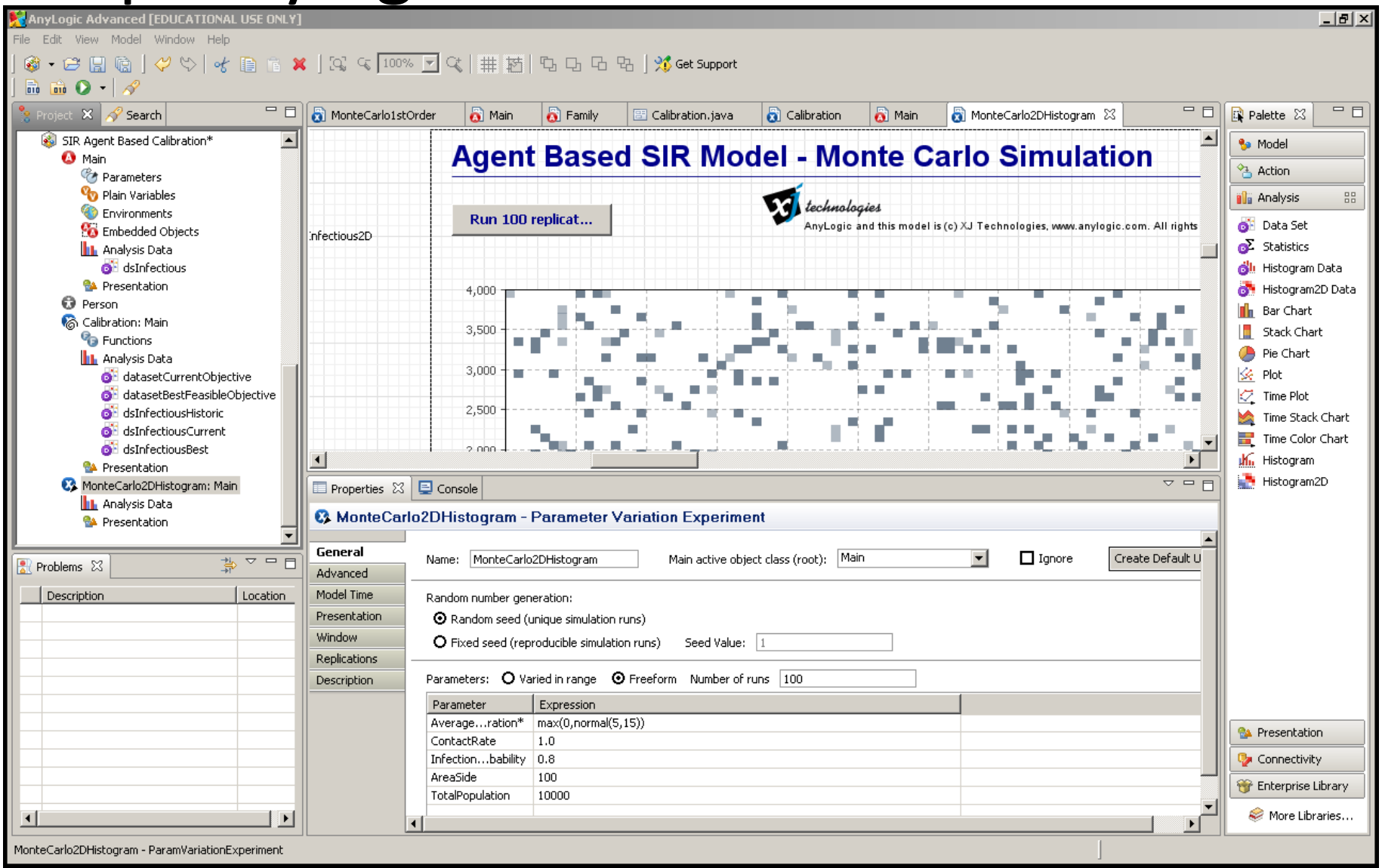
Accumulating the Histogram2D dataset from other datasets



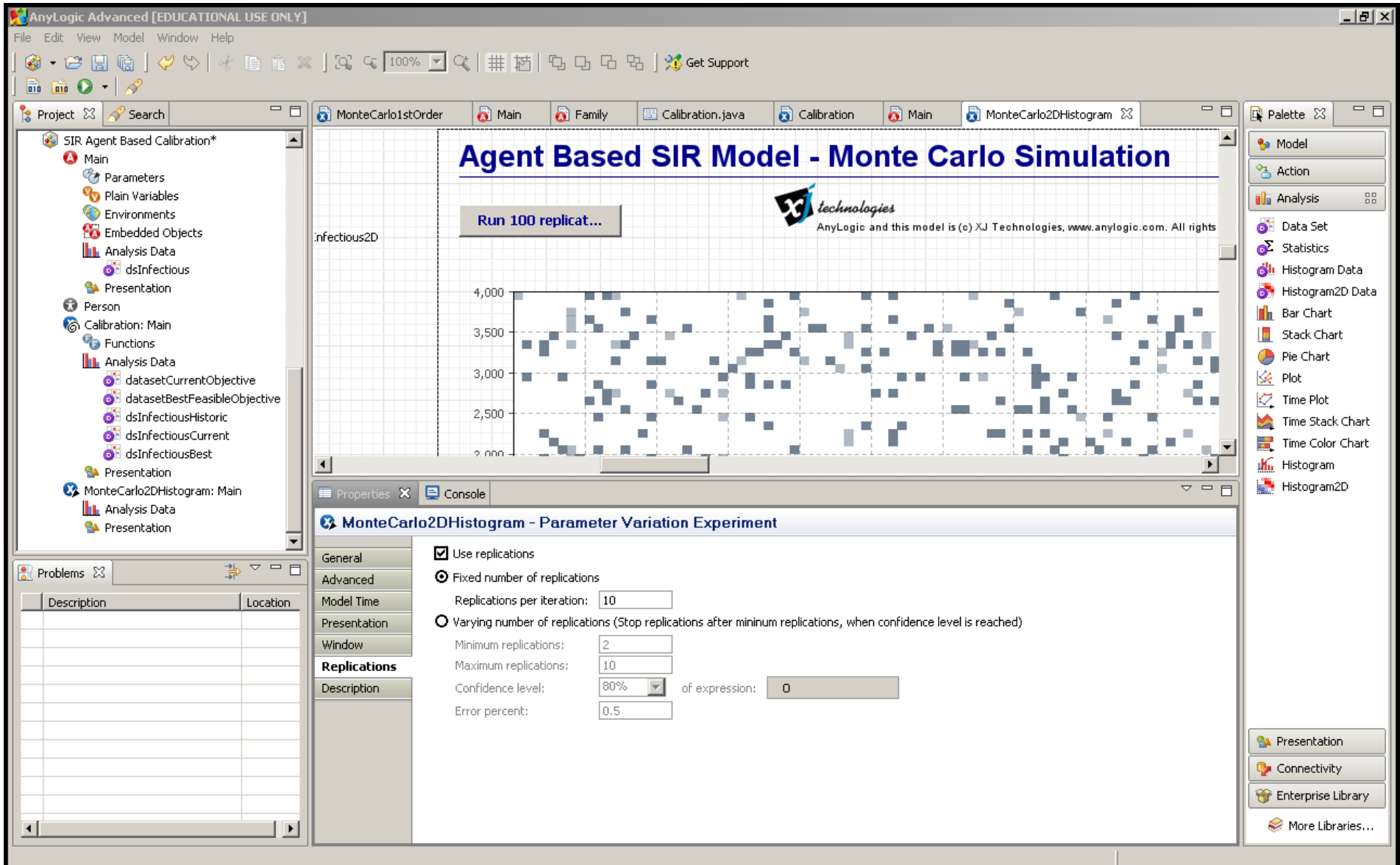
Monte Carlo Sensitivity Analyses in AnyLogic



Monte Carlo Analyses in AnyLogic: Specifying Distributions for Parameters



Automatic Throttling of Monte Carlo Analyses



Reminder: Statistical Scaling

- Consider Taking the sample mean of *n* samples that vary independently around a mean
- If two samples x and y are independent samples of random variables X and Y, then Var[x+y]=Var[X]+Var[Y]
  - So if we have n indep. samples x_i from distribution X

$$Var\left(\sum_{i=1}^{n} x_i\right) = nVar(X)$$

- If we scale a random variable by a factor α, the standard deviation scales by the same factor of α => the variance scales by α^2
  - i.e. StdDev[αX]= α StdDev[X], Var[αX]= α^2 Var[X]

Statistics of Sample Mean

- Recall: Sample Mean: $m = \frac{\sum_{i=1}^{n} x_i}{n}$
- From the preceding, we have

$$Var(m) = Var\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) = \frac{Var\left(\sum_{i=1}^{n} x_i\right)}{n^2} = \frac{nVar(X)}{n^2} = \frac{Var(X)}{n}$$

- This means that standard deviation for the sample mean of n samples varies as

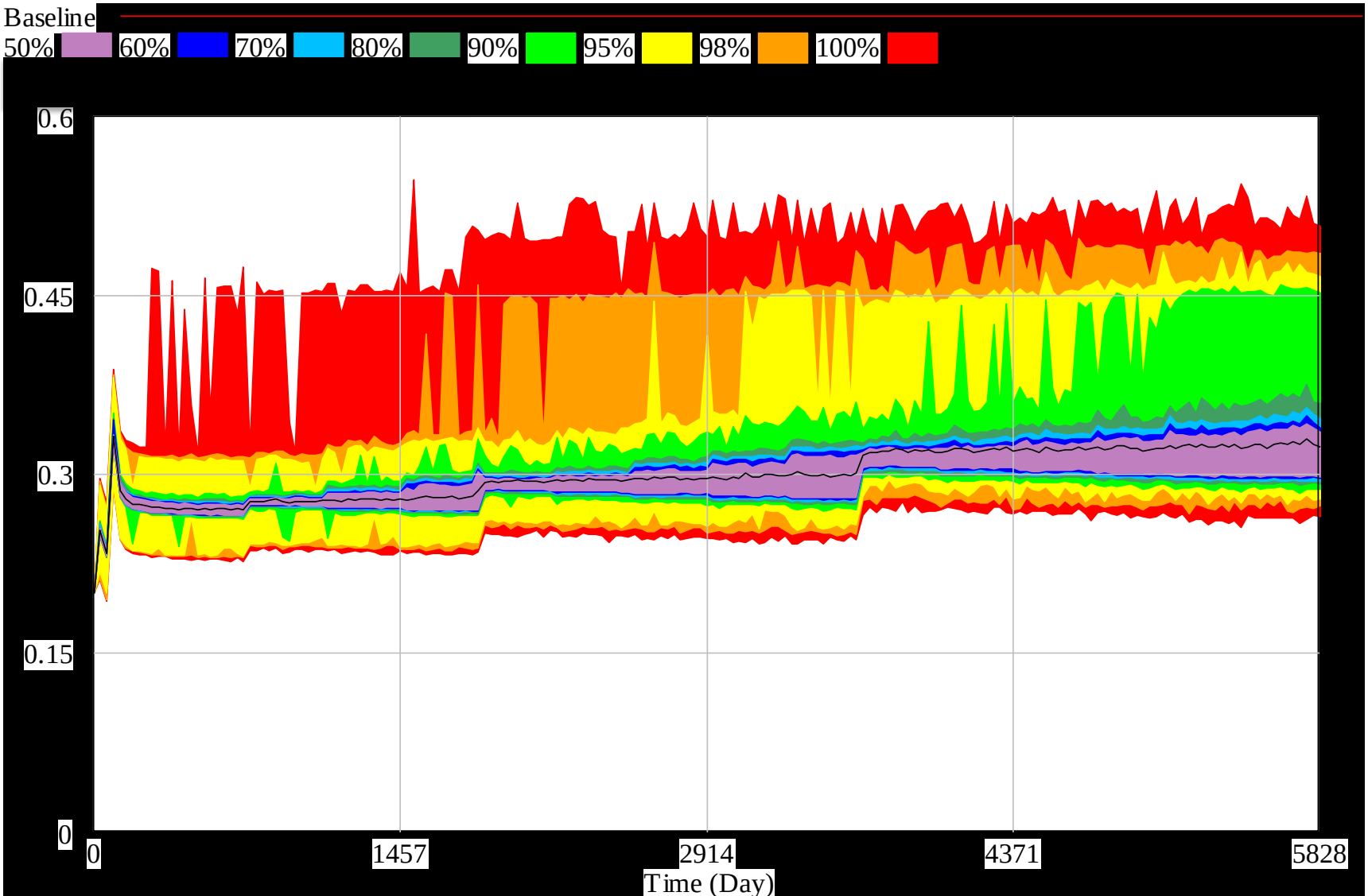
$$StdDev(m) = \sqrt{Var(m)} = \sqrt{\frac{Var(X)}{n}} = \sqrt{\frac{\left(StdDev(X)\right)^2}{n}} = \frac{StdDev(X)}{\sqrt{n}}$$

- So if we wish to divide the standard deviation of the sample mean by a factor of 2, we need to take 4x the number of Monte Carlo samples

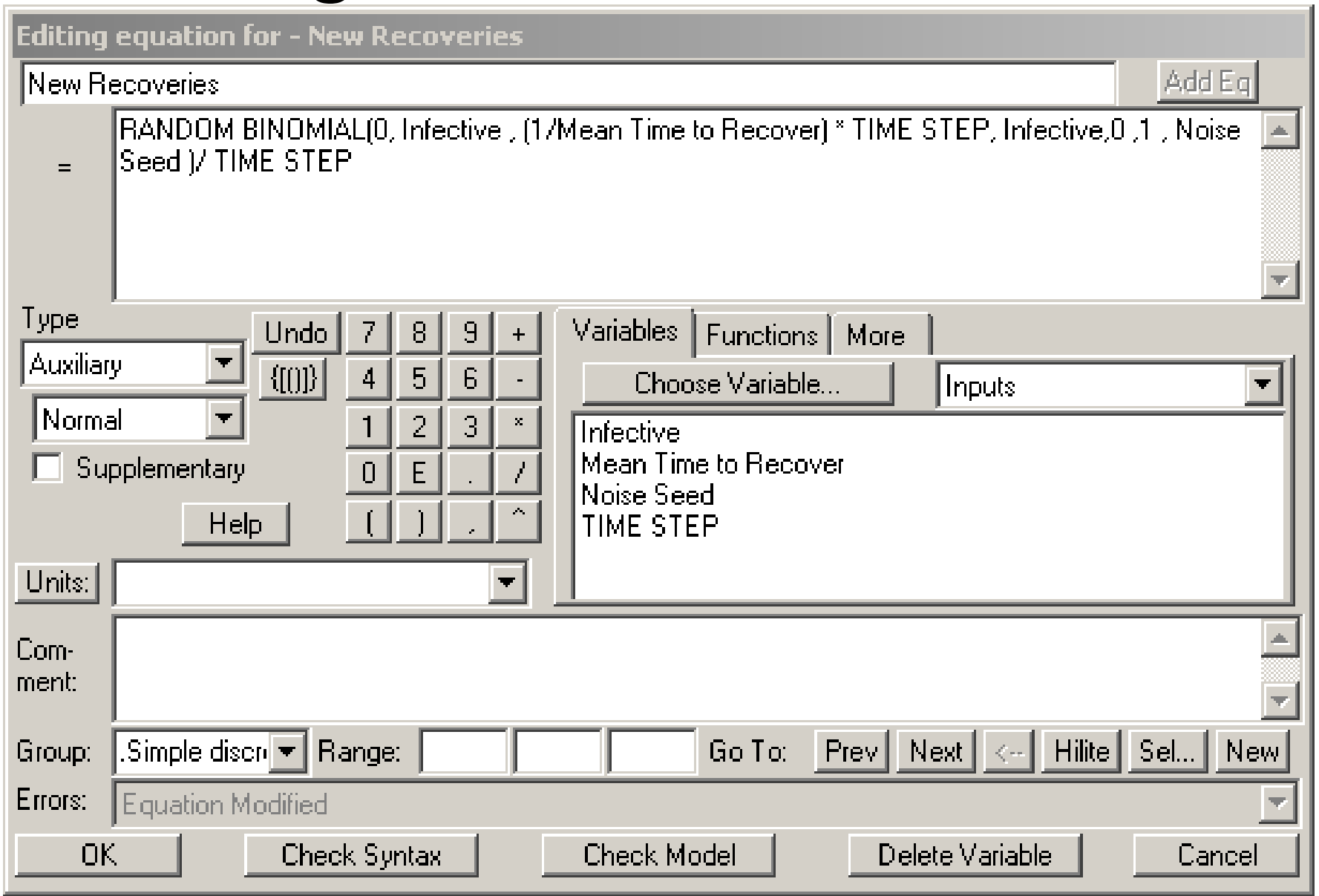
Dynamic Uncertainty: Stochastic Processes

- Examples of things commonly stochastically approximated
  - Stock market
  - Rainfall
  - Oil prices
  - Economic growth
- What considered “stochastic” will depend on the scope of the model
  - Detailed model: Individual behaviour, transmission, etc.
  - A meteorological model may not consider rainfall stochastic

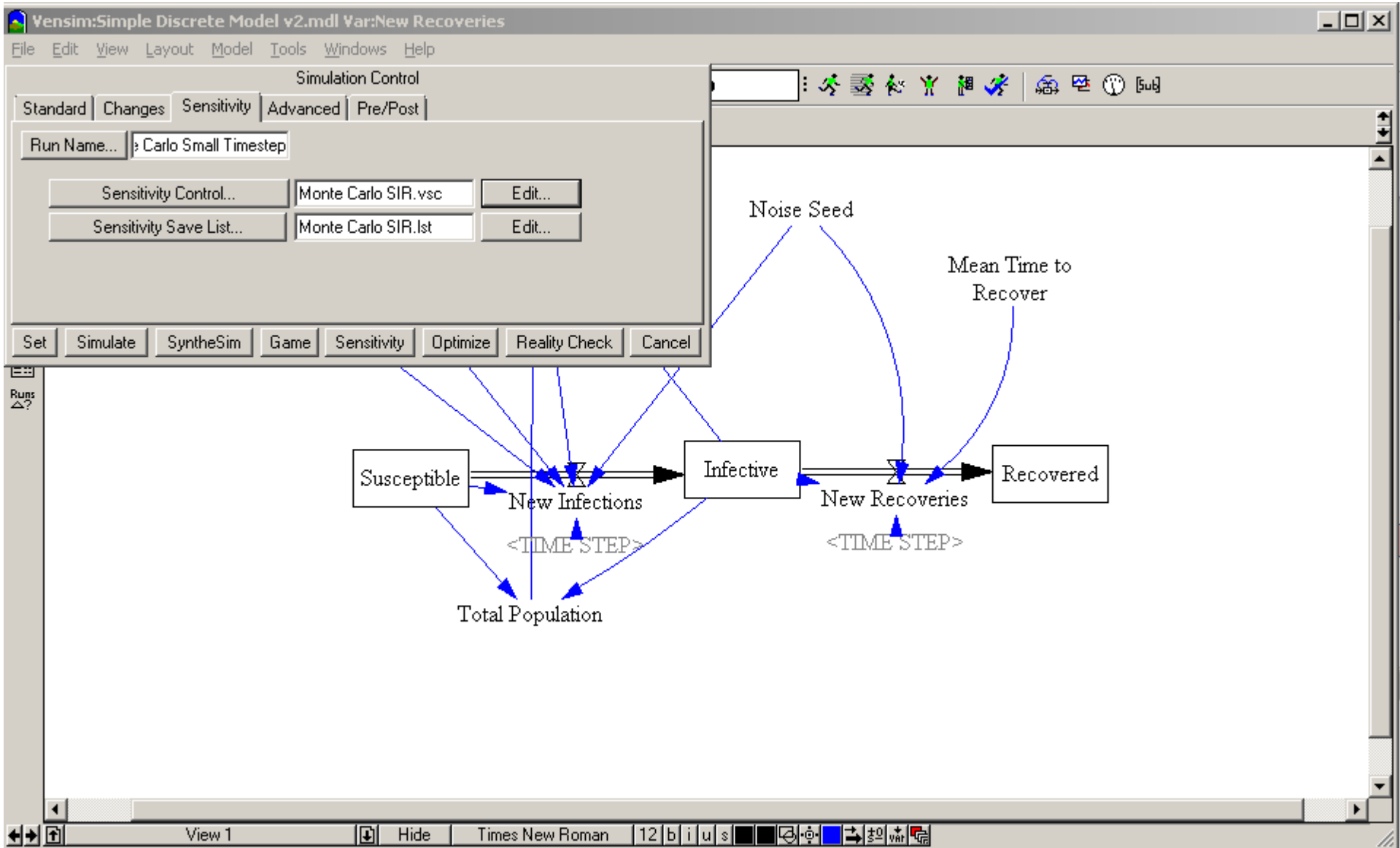
Stochastic Processes in Vensim



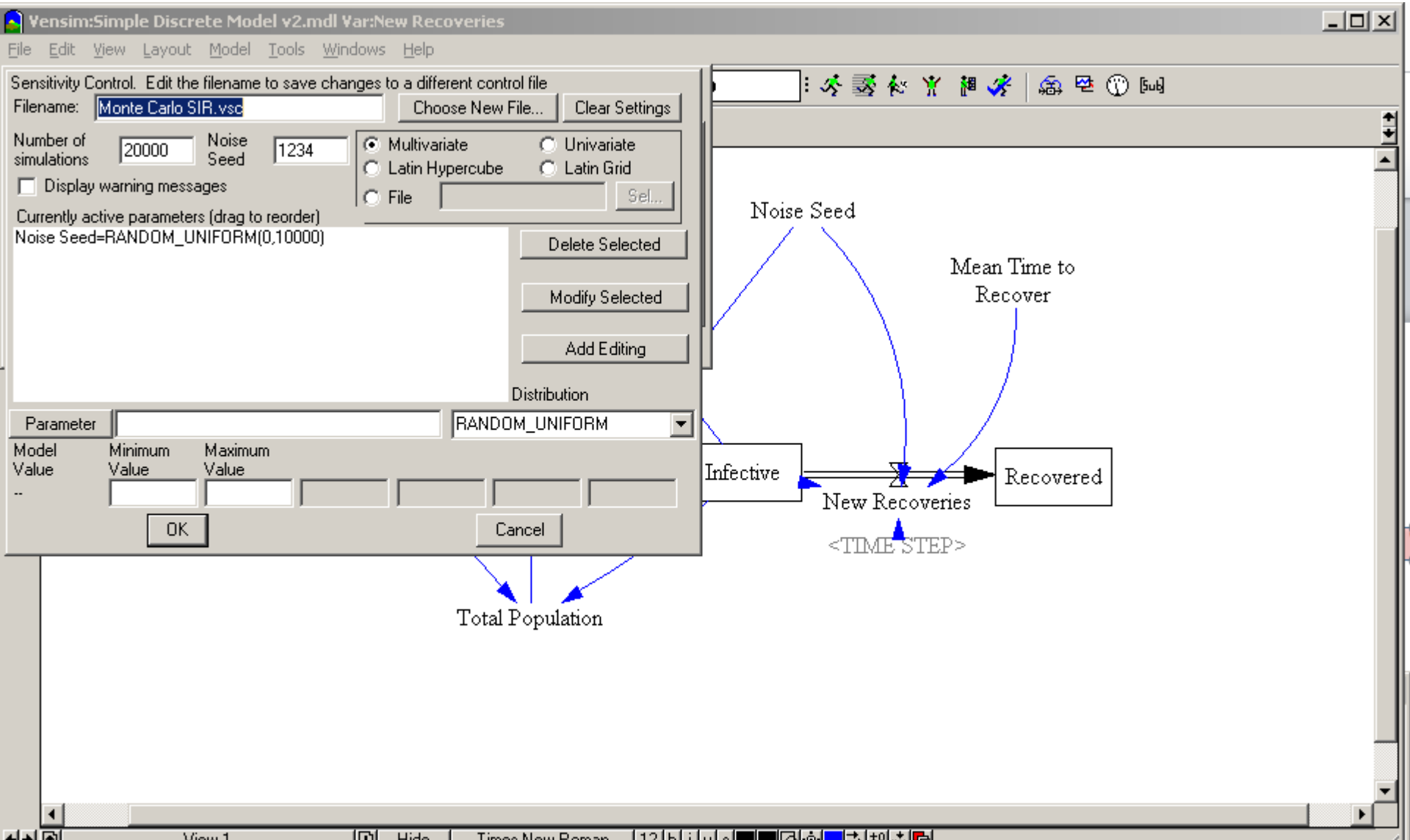
Making a Vensim Flow Stochastic



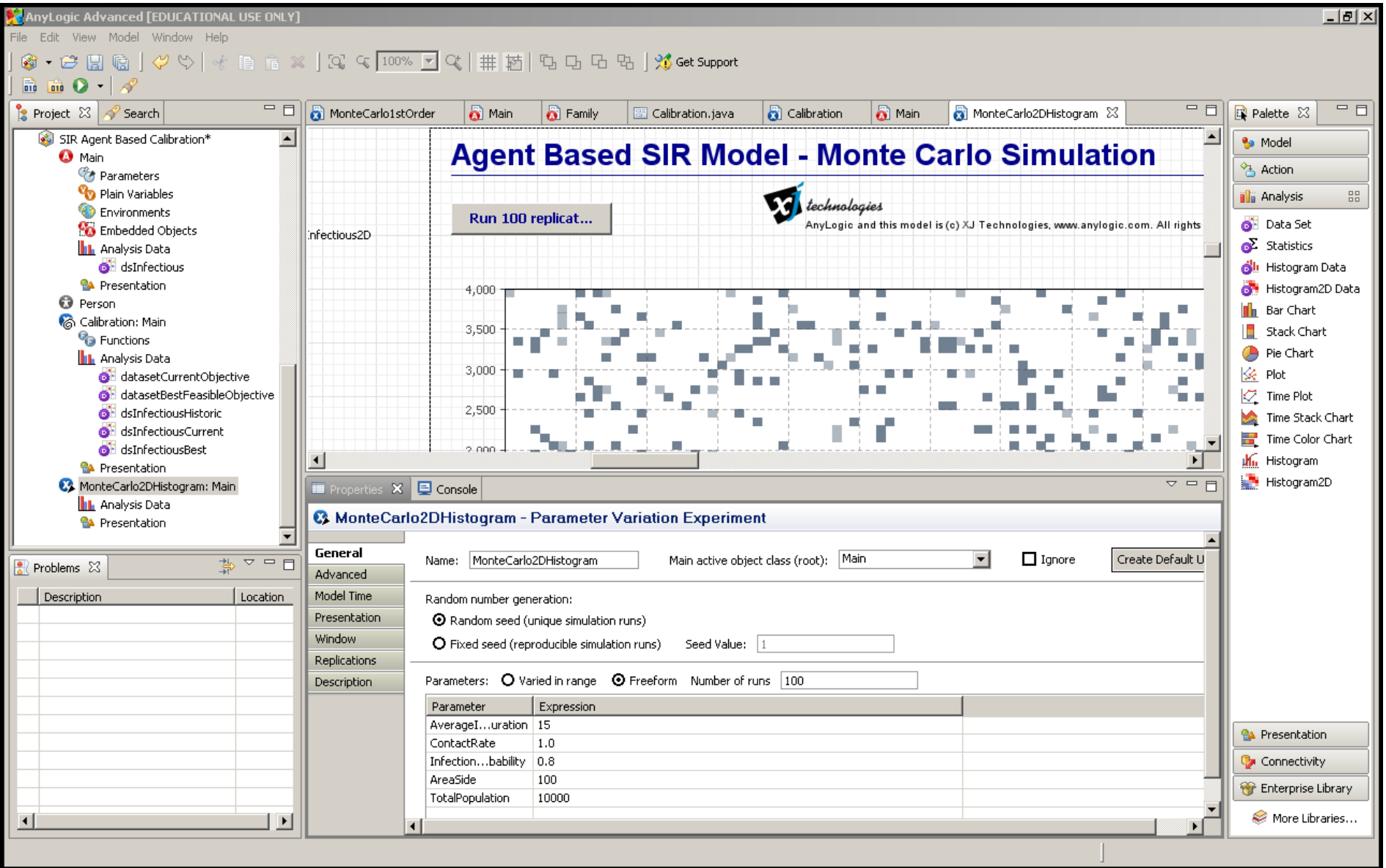
Treat as a Sensitivity Analysis



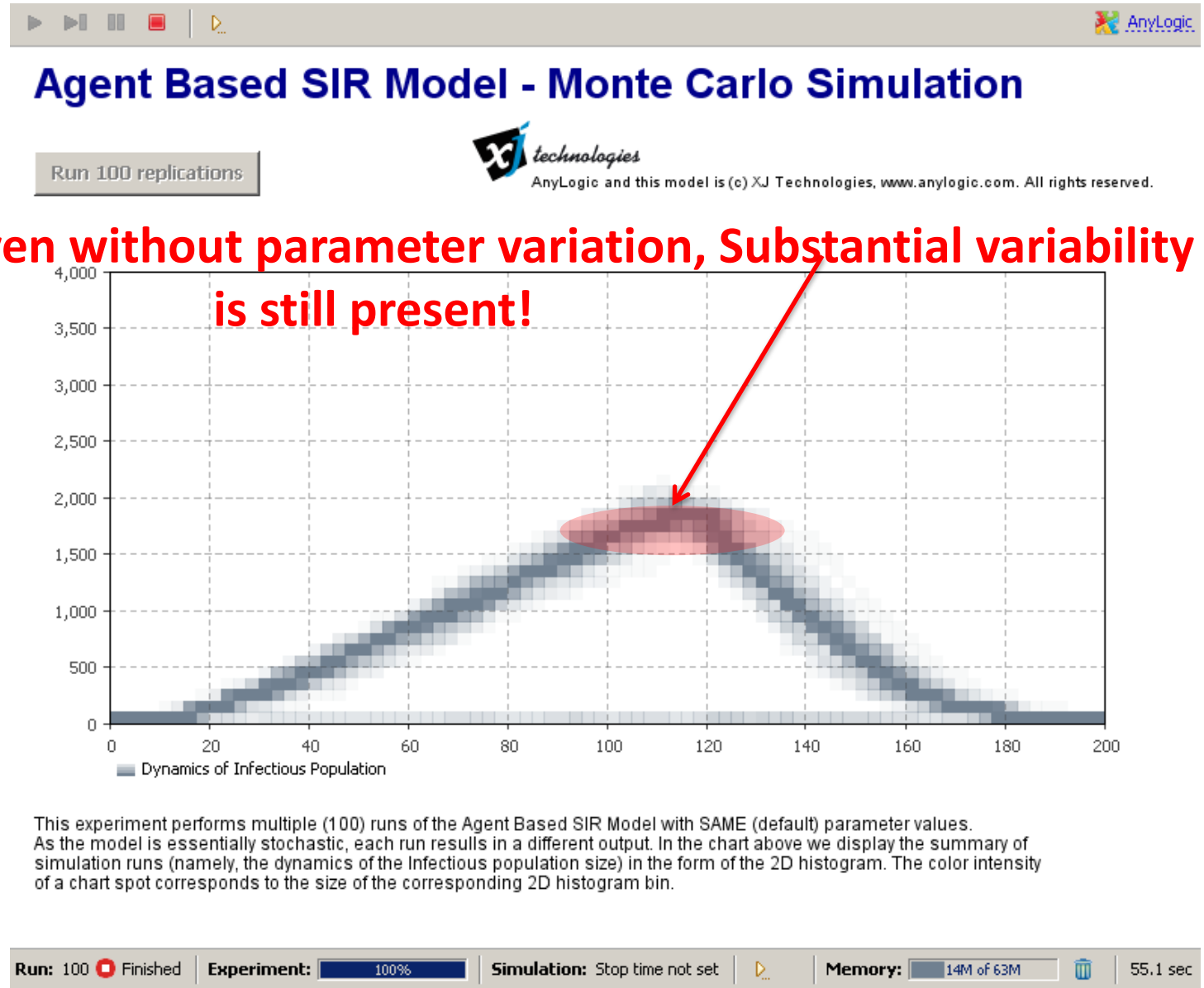
Setting the “Random Seed” to Differ between Simulations



Monte Carlo Analysis with Fixed Parameter Values



Results of Monte Carlo Simulation



Stochastic Processes in AnyLogic

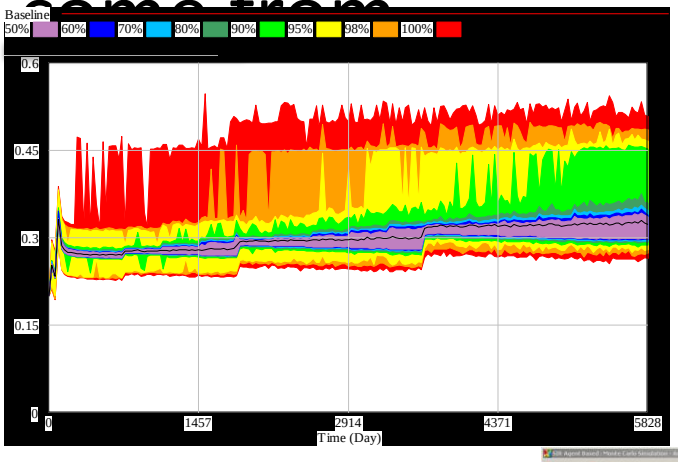
- In AnyLogic, ABM and Discrete Event Models ("Network-Based Modeling") are typically stochastic
  - Transitions between states
  - Event firing
  - Messages
    - (Frequent) timing of message send
    - Target of messages
  - Duration of a procedure
- As a result, there will be variation in the results from simulation to simulation

Summarizing Variability

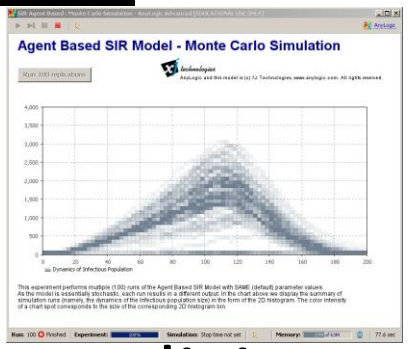
- To gain confidence in model results, typically need to run an ensemble of realizations
  - Deal with means, standard deviations, and empirical fractiles
  - As is seen here, there are typically still broad regularities between most runs (e.g. rise & fall)
- Need to reason over a population of realizations $\Rightarrow$ statistics are very valuable
  - Fractile within which historic value falls
  - Mean difference in results between interventions

Closing Question: How can we best adapt our policies to deal with ongoing uncertainty?

- We are dealing here with making decisions in an environment that changes over time
- This uncertainty could come from
  - Stochastic variability



  - Uncertainty regarding parameter value



- There is an incredibly vast # of possible policies

Stochastic Processes in Vensim