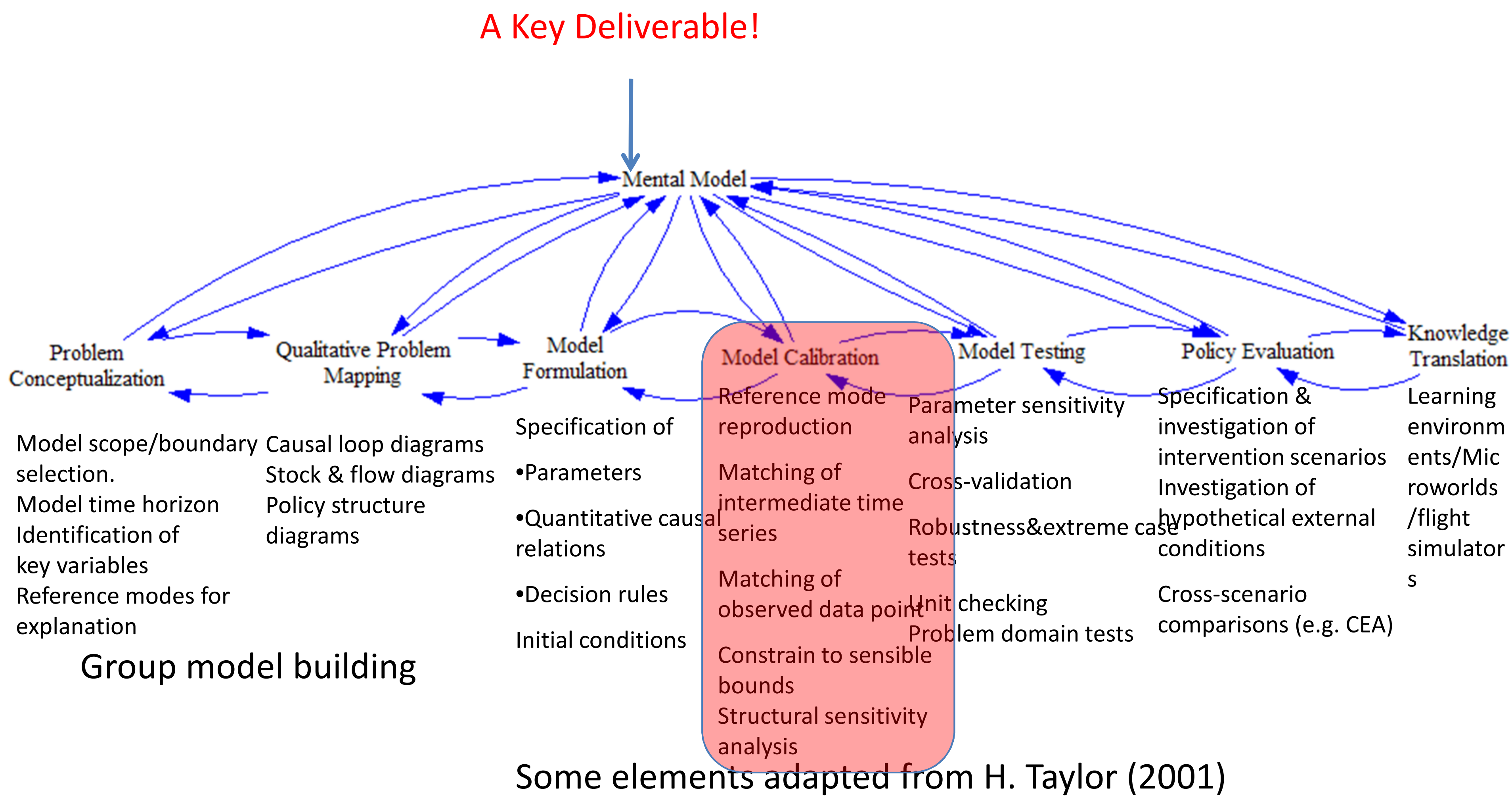
Dealing with Data Gradients: Calibration 2

Nathaniel Osgood

CMPT 858

March 31, 2011


A Key Deliverable!
Mental Model
Problem Conceptualization
Qualitative Problem Mapping
Model Formulation
Model Calibration
Model Testing
Policy Evaluation
Knowledge Translation
Model scope/boundary selection.
Model time horizon
Identification of key variables
Reference modes for explanation
Causal loop diagrams
Stock & flow diagrams
Policy structure diagrams
Specification of
•Parameters
•Quantitative causal relations
•Decision rules
Initial conditions
Reference mode reproduction
Matching of intermediate time series
Matching of observed data points
Constrain to sensible bounds
Structural sensitivity analysis
Parameter sensitivity analysis
Cross-validation
Robustness&extreme case tests
Unit checking
Problem domain tests
Specification & investigation of intervention scenarios
Investigation of hypothetical external conditions
Cross-scenario comparisons (e.g. CEA)
Learning environments/Microworlds/flight simulators
Group model building
Some elements adapted from H. Taylor (2001)

Recall: Dealing with Data Gradients

- Often we don't have reliable information on *some* parameters, but do have other data
  - Some parameters may not be observable, but some closely related observable data is available
  - Sometimes the data doesn't have the detailed breakdown needed to specifically address one parameter
    - Available data could specify sum of a bunch of flows or stocks
    - Available data could specify some function of several quantities in the model (e.g. prevalence)
- Some parameters may implicitly capture a large set of factors not explicitly represented in model
- There are two big ways of dealing with this: manually "backing out", and automated calibration

Recall: Calibration: “Triangulating” from Diverse Data Sources

- Calibration involves “tuning” values of less well known parameters to best match observed data
  - Often try to match against *many* time series or pieces of data at once
  - Idea is trying to get the software to answer the question: “What must these (less known) parameters be in order to explain all these different sources of data I see”
- Observed data can correspond to complex combination of model variables, and exhibit “emergence”
- Frequently we learn from this that our model structure just can’t produce the patterns!

Recall: Calibration: A Bit of the How

- Calibration uses a (global) optimization algorithm to try to adjust unknown parameters so that it automatically matches an arbitrarily large set of data
- The data (often in the form of time series) forms constraints on the calibration
- The optimization algorithm will run the model many (minimally, thousands, typically 100K or more) times to find the "best" match for all of the data

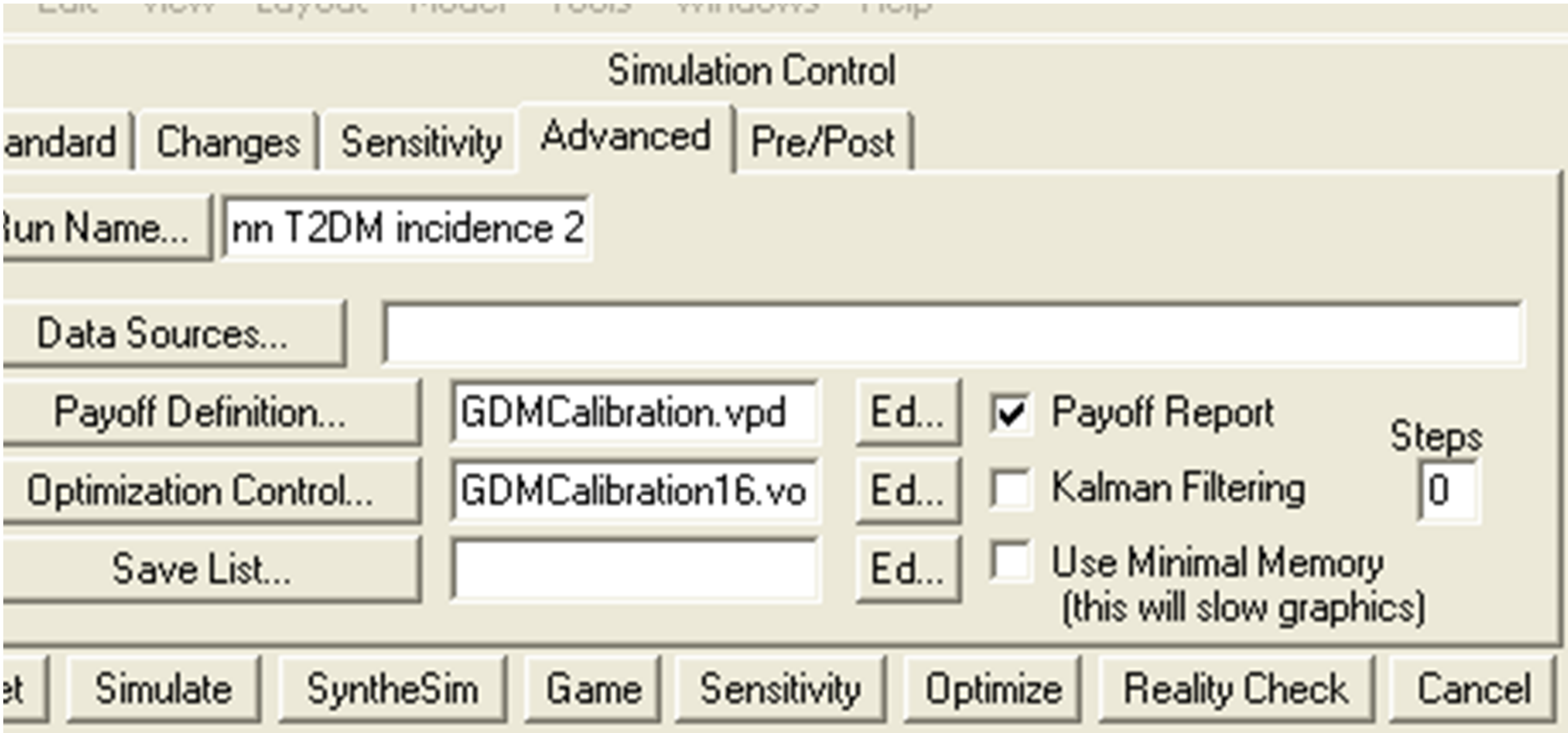
Recall: Required Information for Calibration

- Specification of what to match (and how much to care about each attempted match)
  - Involves an “error function” (“penalty function”, “energy function”) that specifies “how far off we are” for a given run (how good the fit is)
  - Alternative: specify “payoff function” (“objective function”)
- A statement of what parameters to vary, and over what range to vary them (the “parameter space”)
- Characteristics of desired tuning algorithm
  - Single starting point of search?

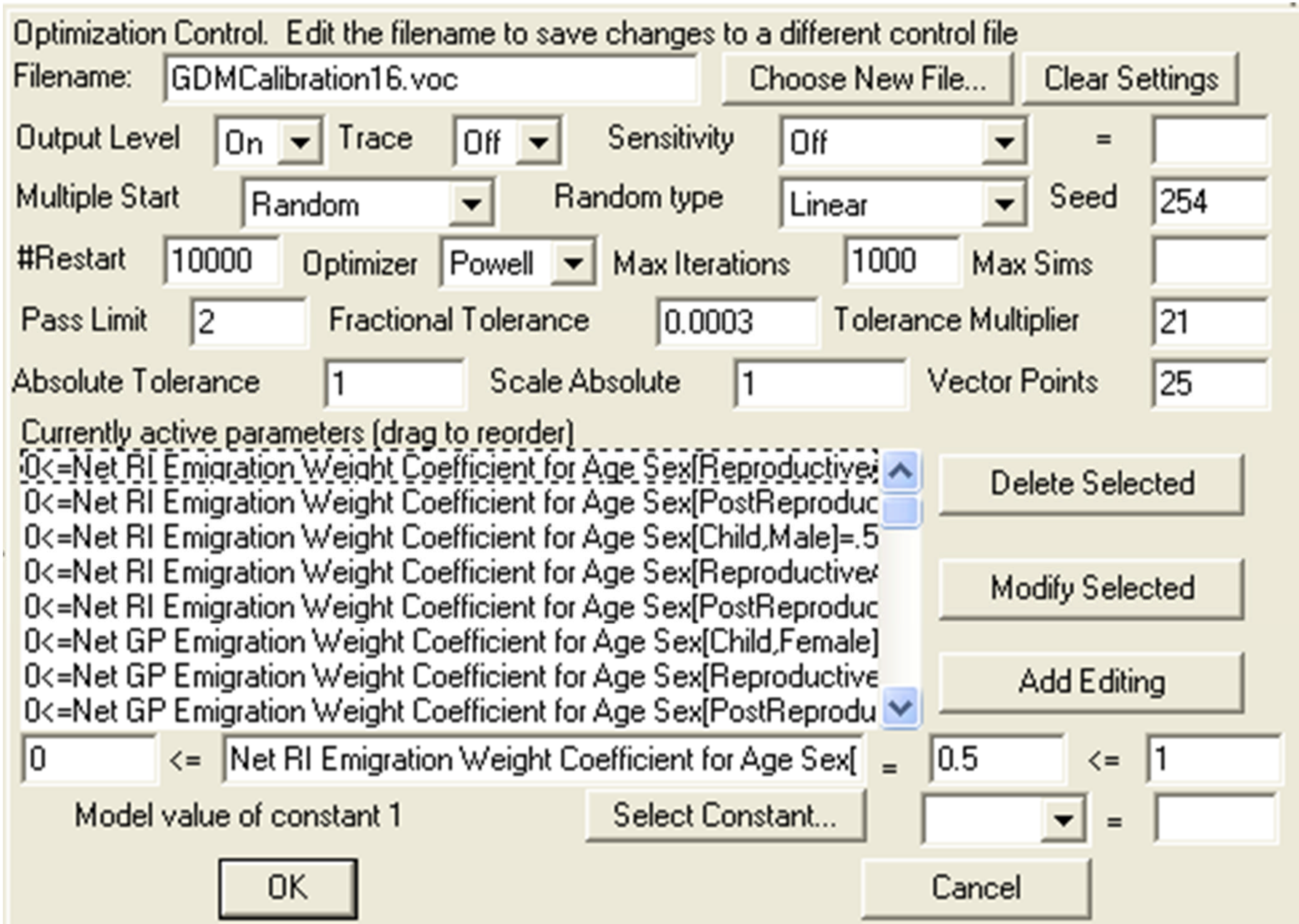
Recall: Example Global Optimization Algorithm

- Starts at random position, tries to improve match (minimize error) by
  - Adjusting parameters
  - Running Model
  - Recording error function
- Keeps on improving until reaches “local minimum” in error of fit
  - May add some randomness to knock out of local minima

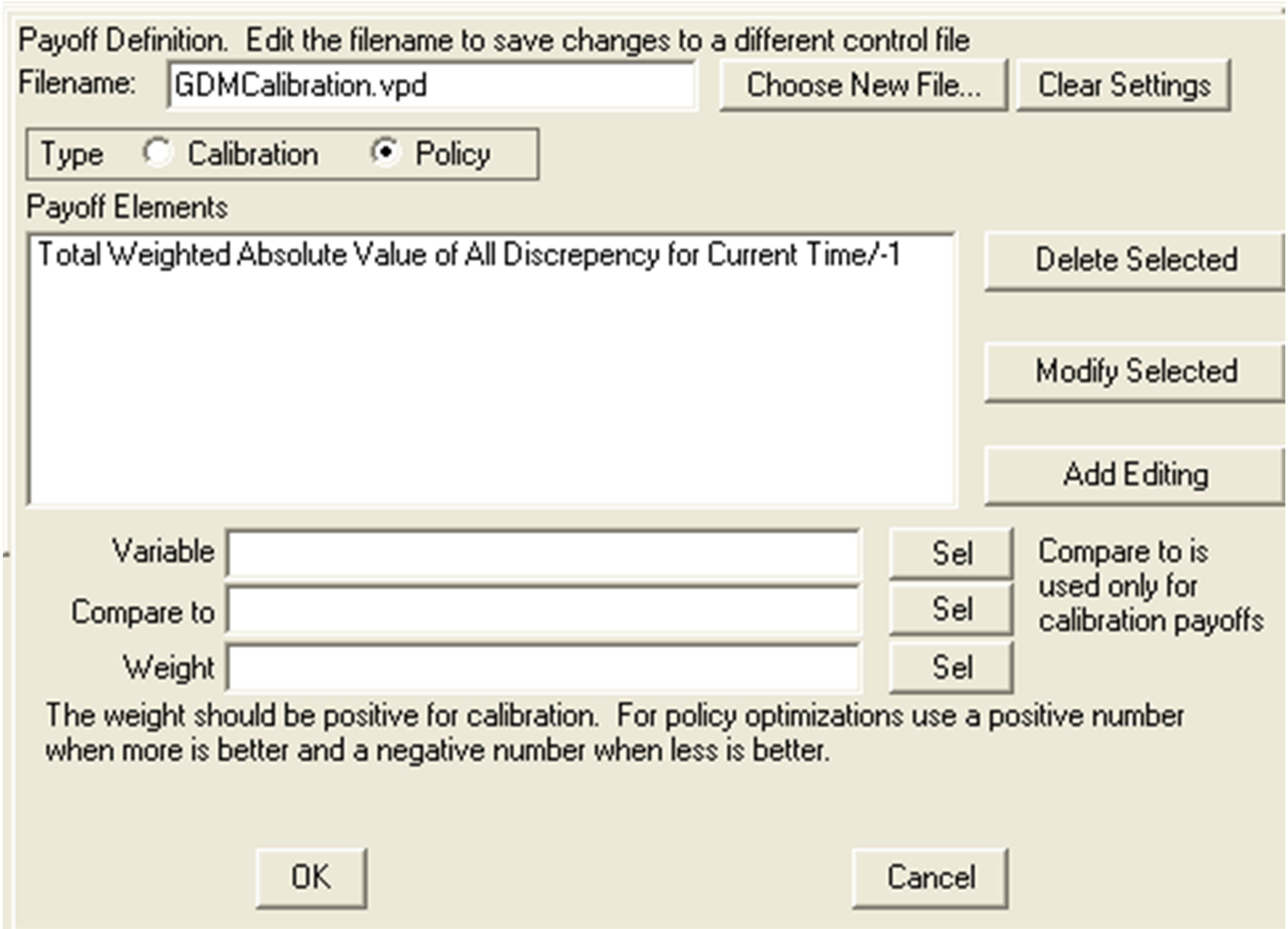
Running Calibrations in Vensim: (Under Model/Simulate Commands)



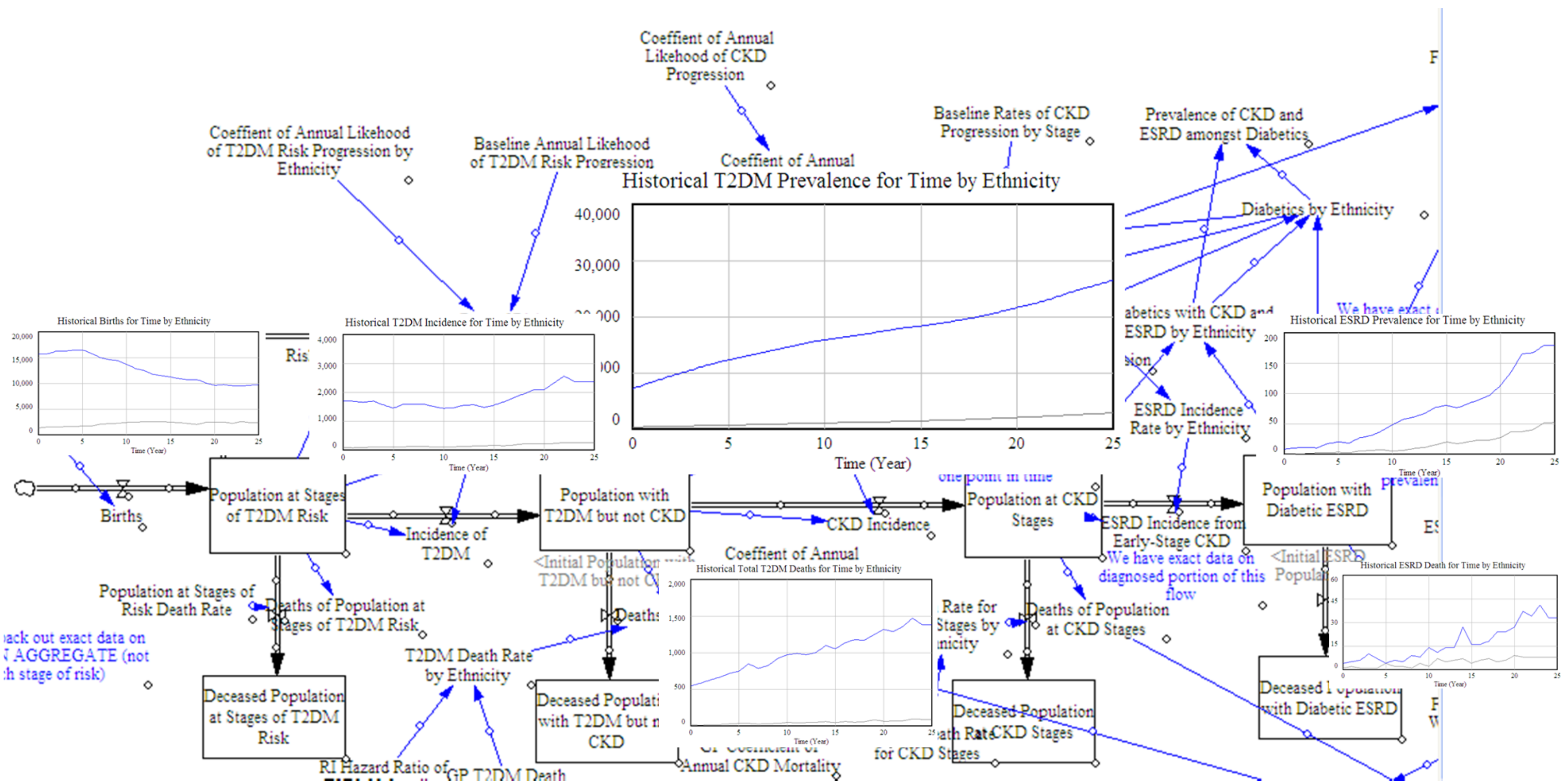
Optimization Control



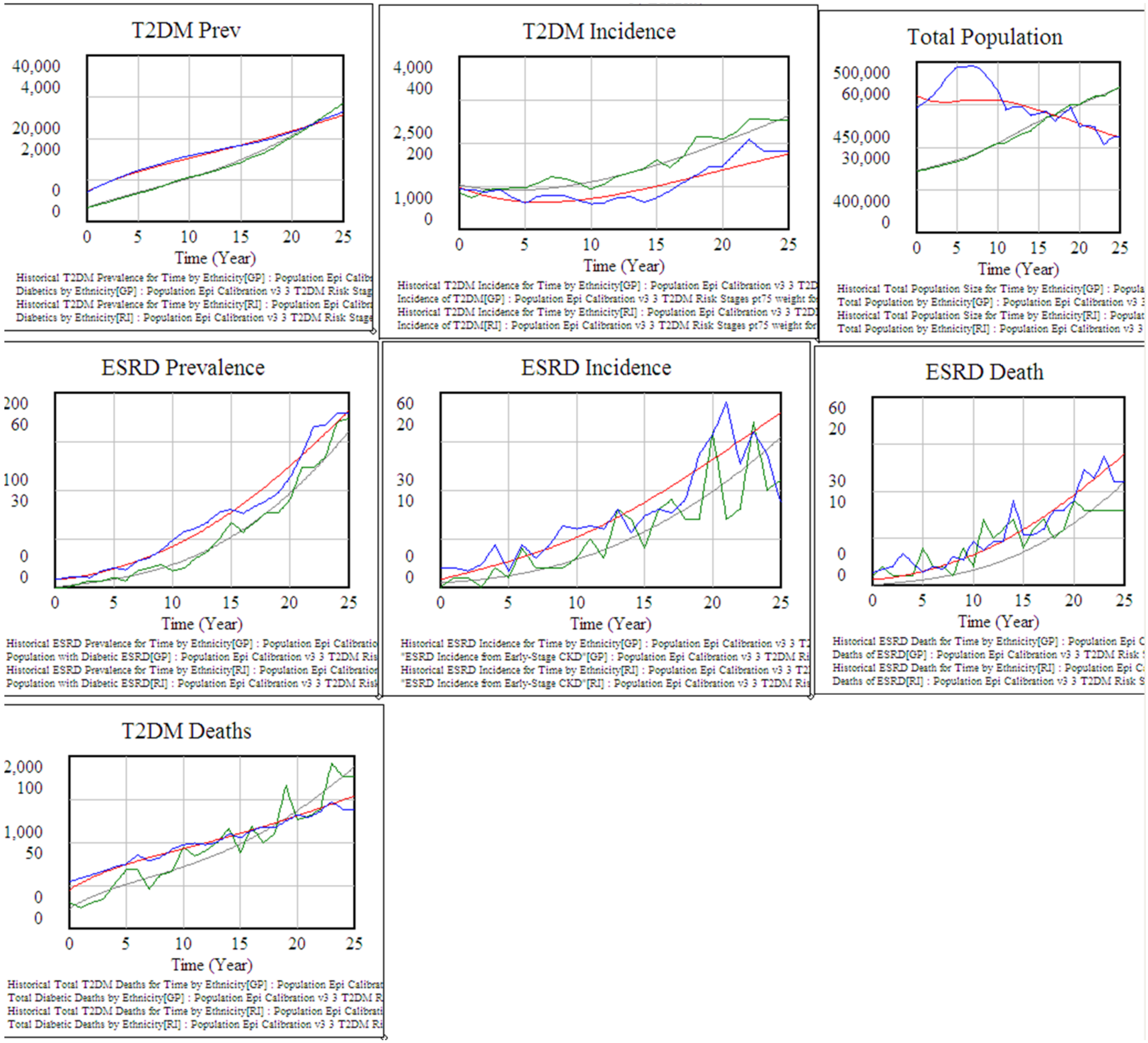
Payoff Definition



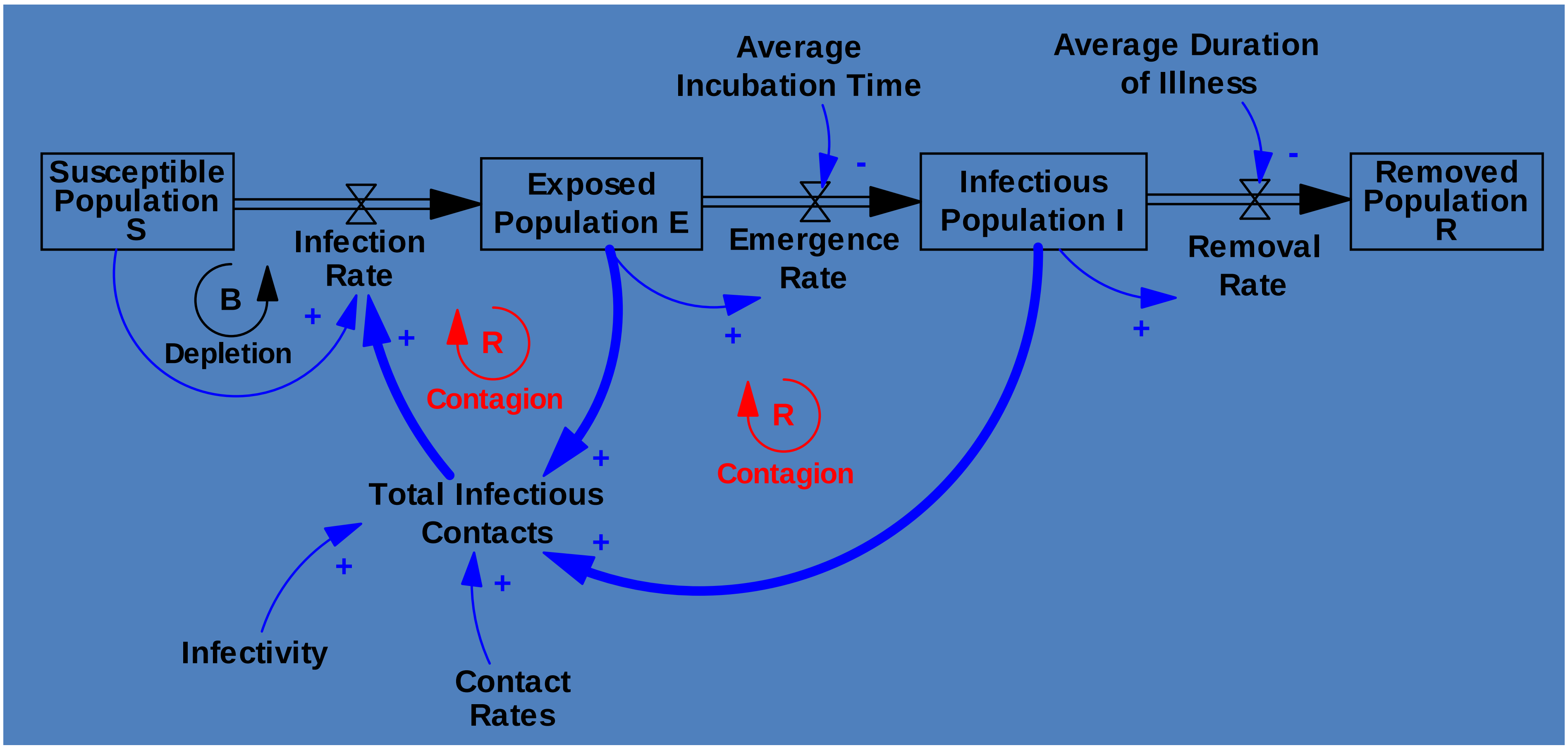
The Pieces of the Elephant
Example Model of Underlying Process & Time Series It Must Match



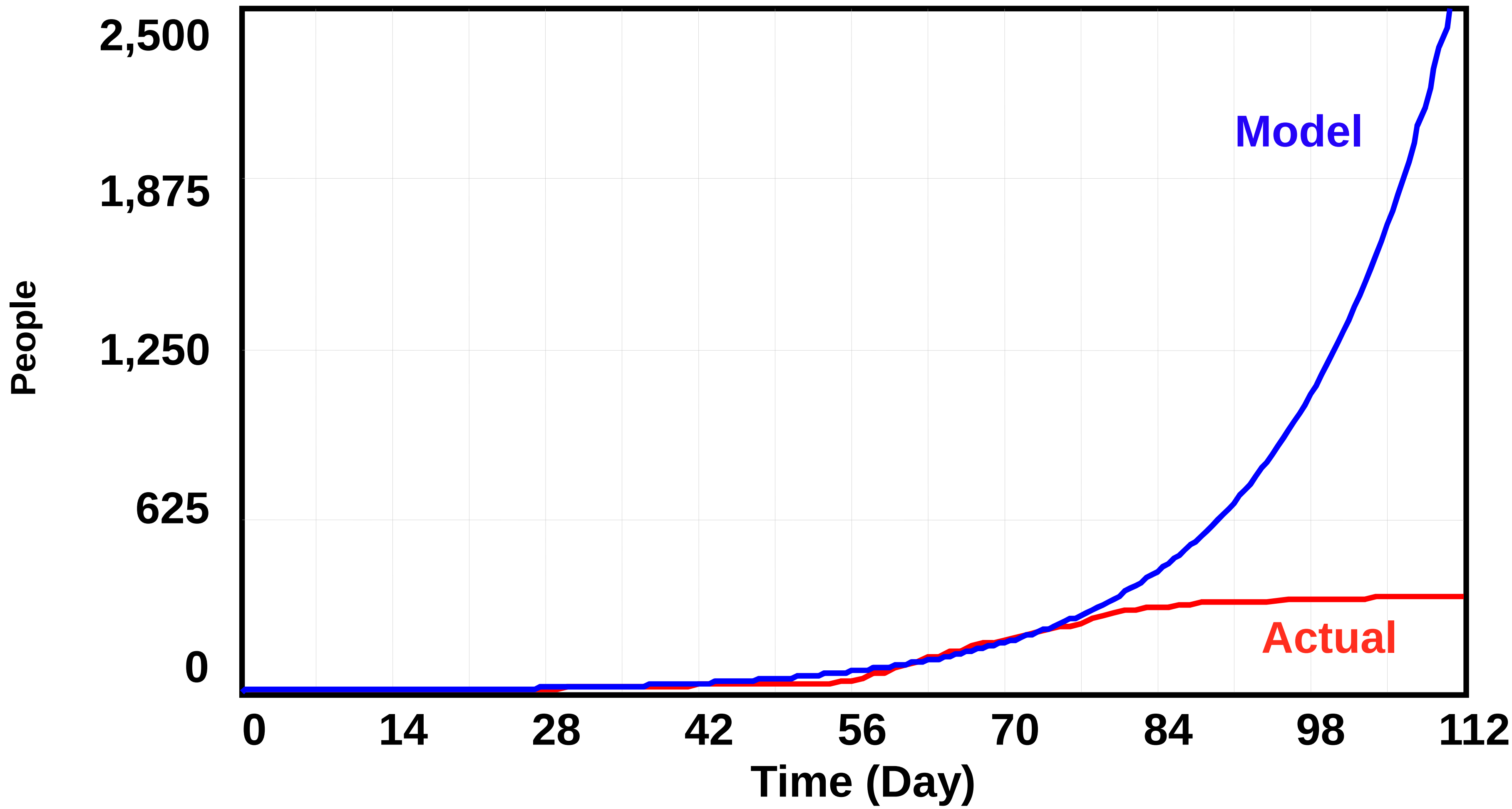
Single Model Matches Many Data Sources



Example: Iteration & Calibration

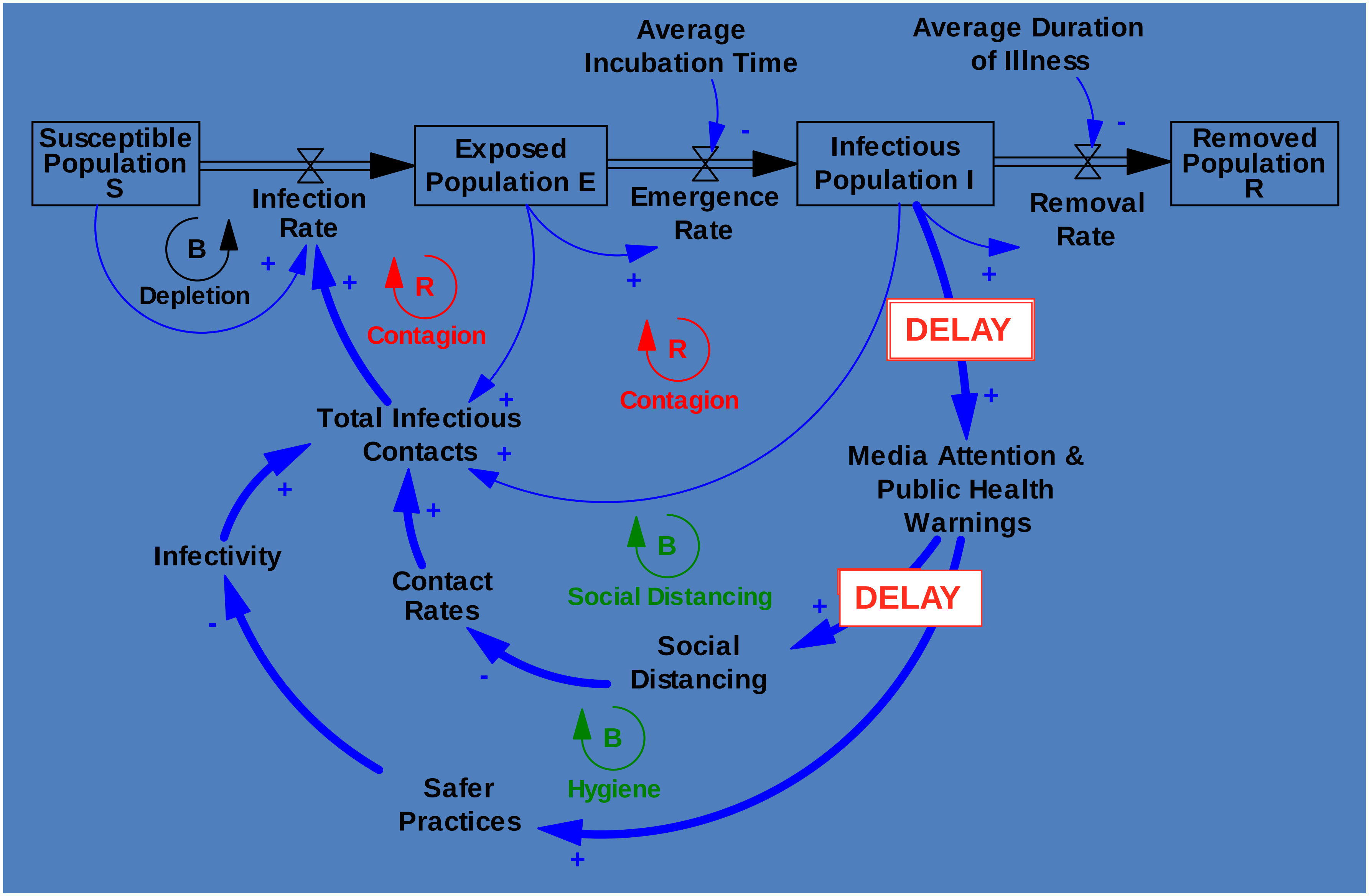


From Sterman

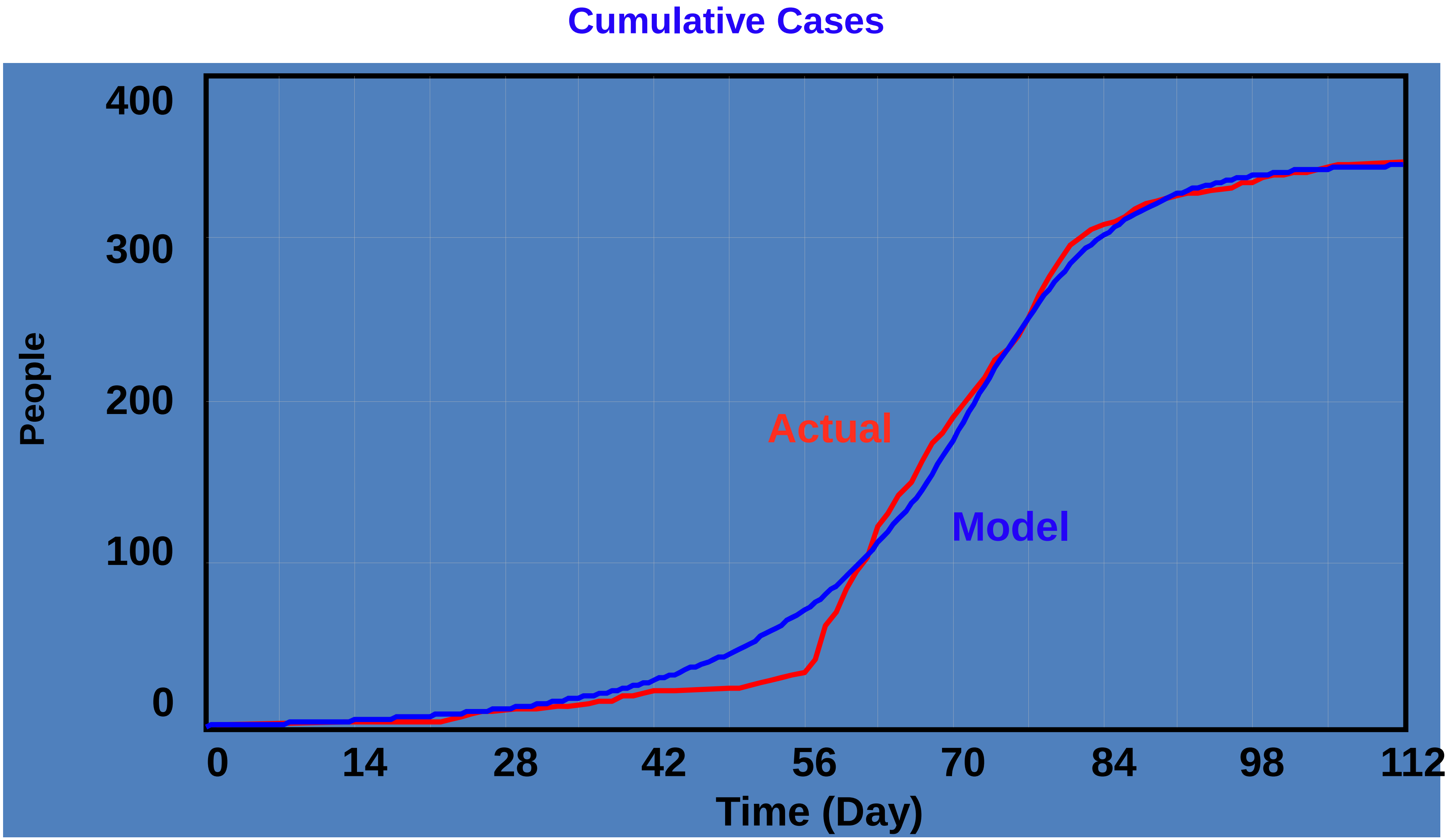


From Sterman

Expanding the Boundary: Behavioral Feedbacks

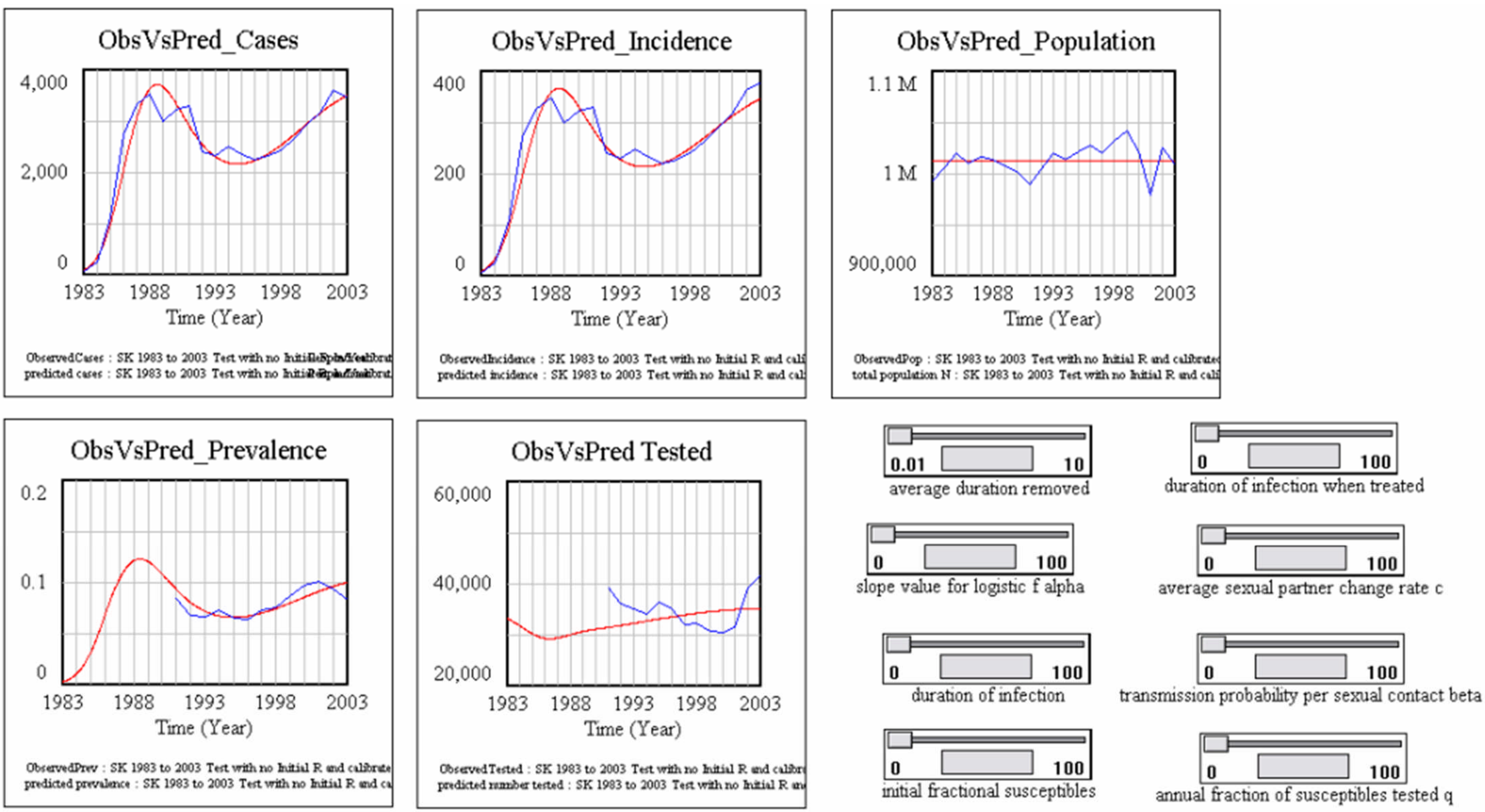


From Sterman

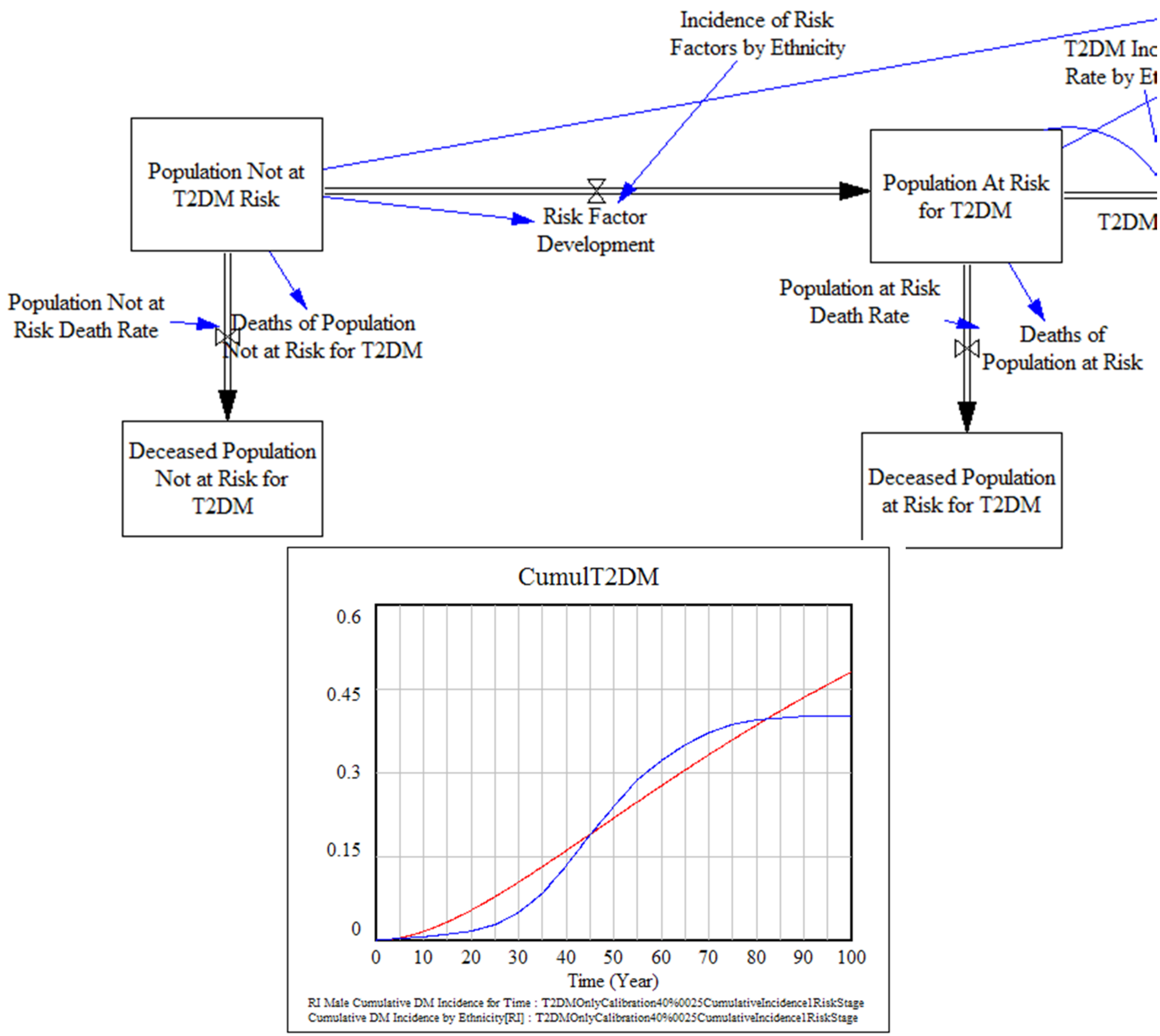


Cumulative Cases
400
300
200
100
0
People
Actual
Model
0
14
28
42
56
70
84
98
112
Time (Day)

Pieces of the Elephant: STI

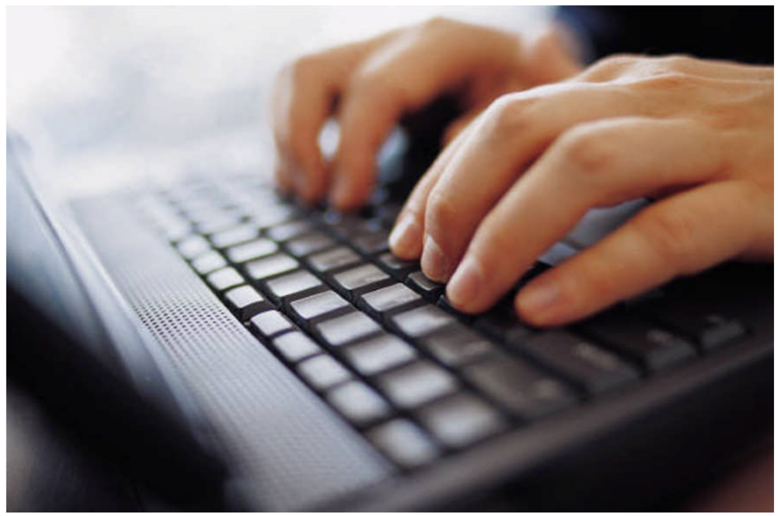


Science



Incidence of Risk Factors by Ethnicity
Population Not at T2DM Risk
Risk Factor Development
Population At Risk for T2DM
Population Not at Risk Death Rate
Deaths of Population Not at Risk for T2DM
Population at Risk Death Rate
Deaths of Population at Risk
Deceased Population Not at Risk for T2DM
Deceased Population at Risk for T2DM
CumulT2DM
0.6
0.45
0.3
0.15
0
0
10
20
30
40
50
60
70
80
90
100
Time (Year)
RI Male Cumulative DM Incidence for Time : T2DMOnlyCalibration40%0025CumulativeIncidence1RiskStage
Cumulative DM Incidence by Ethnicity[RI] : T2DMOnlyCalibration40%0025CumulativeIncidence1RiskStage

Hands on Model Use Ahead

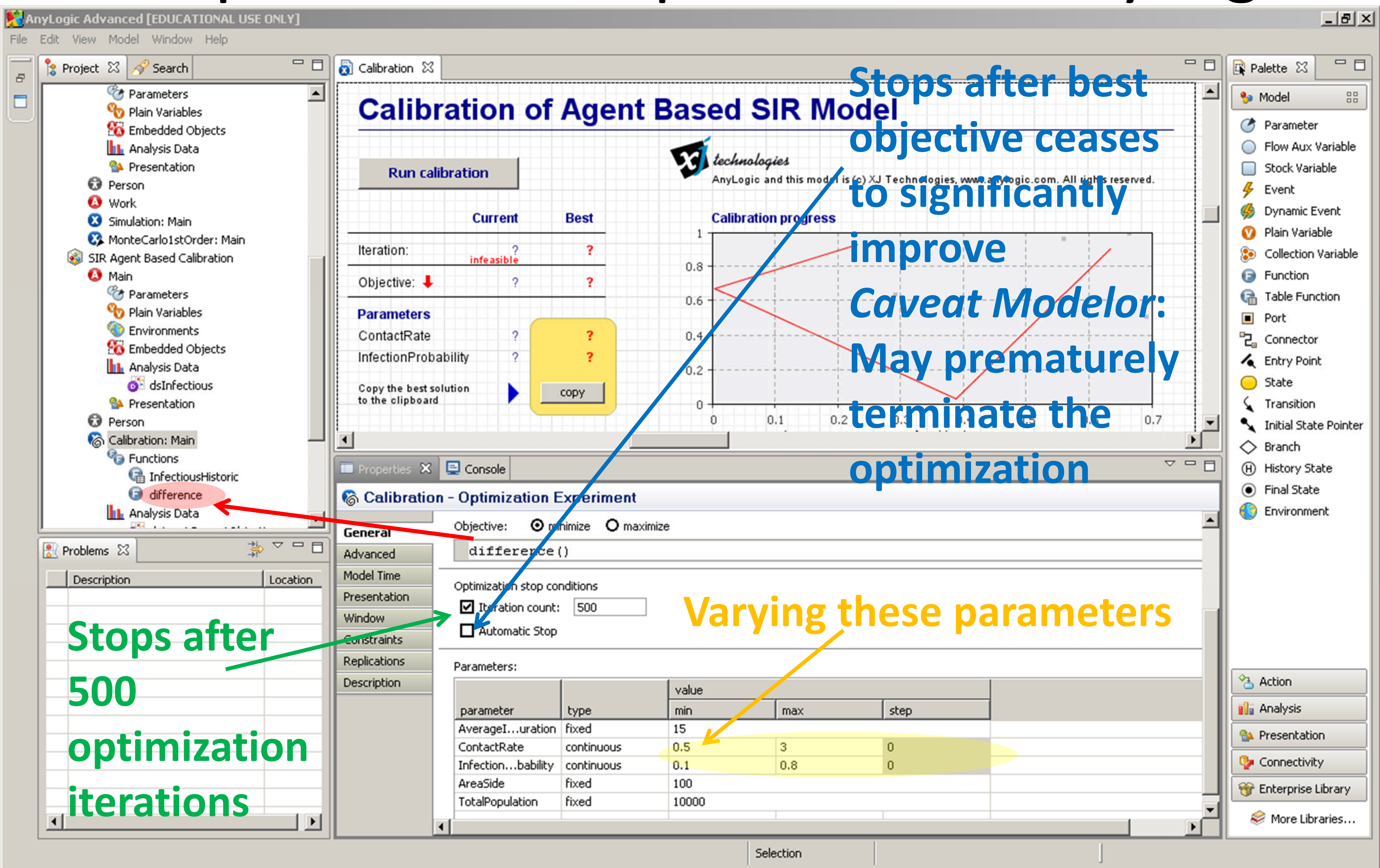


Load Sample Model:

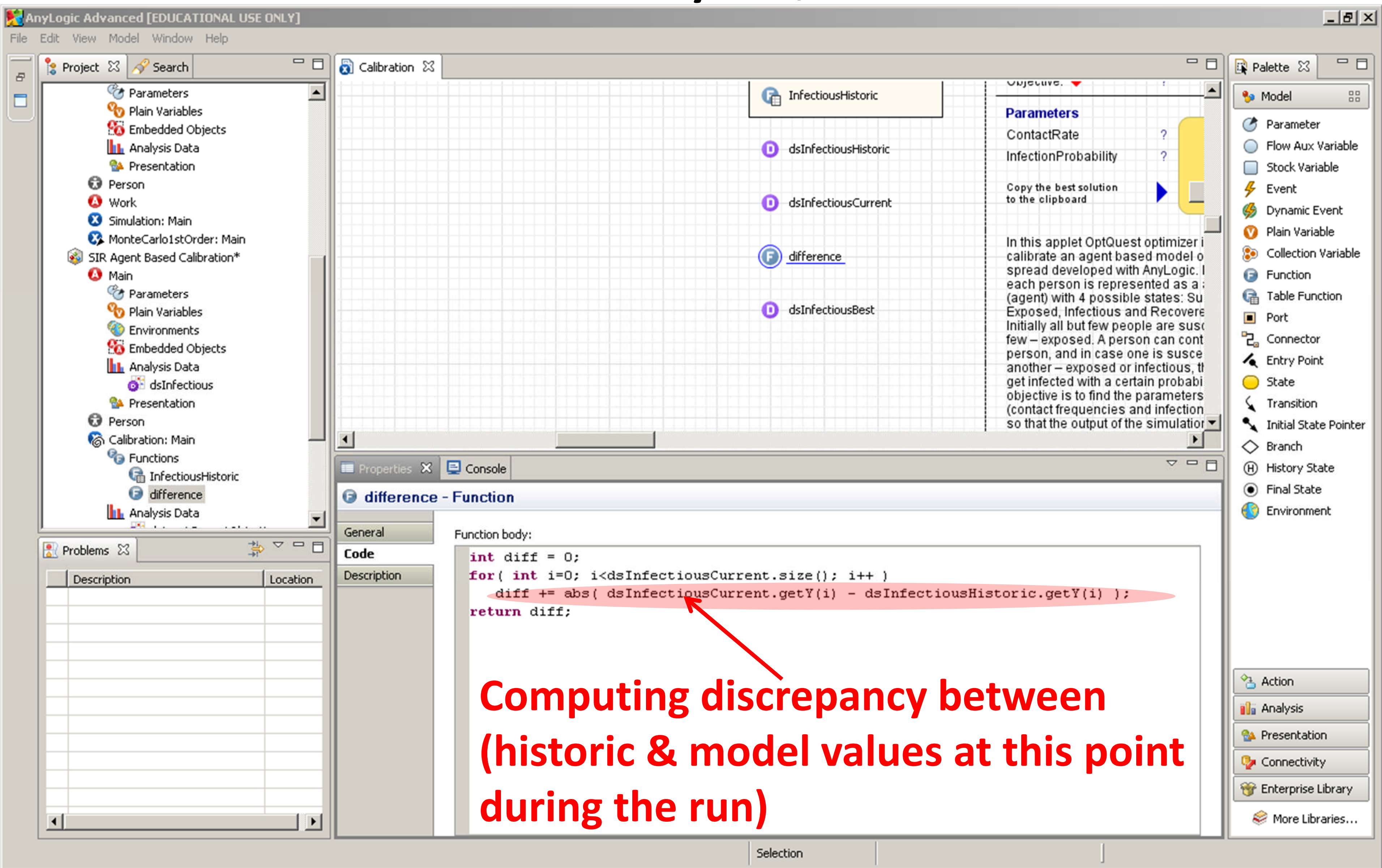
SIR Agent Based Calibration

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

An Optimization Experiment in AnyLogic

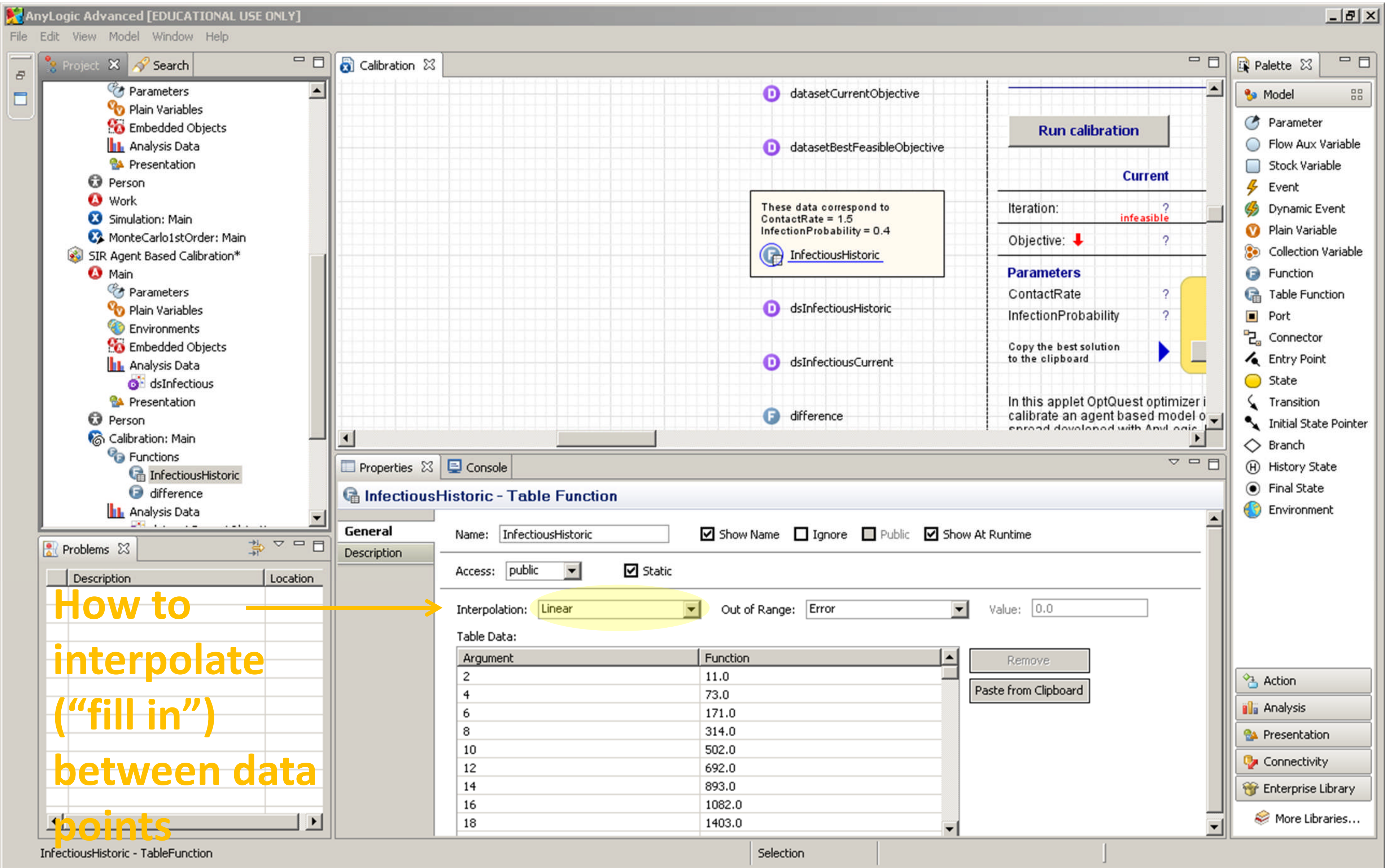


Defining a Payoff Function
Caveat: Non-Analytic, Non-Concave



Computing discrepancy between (historic & model values at this point during the run)

Historic Data Captured via Table Function



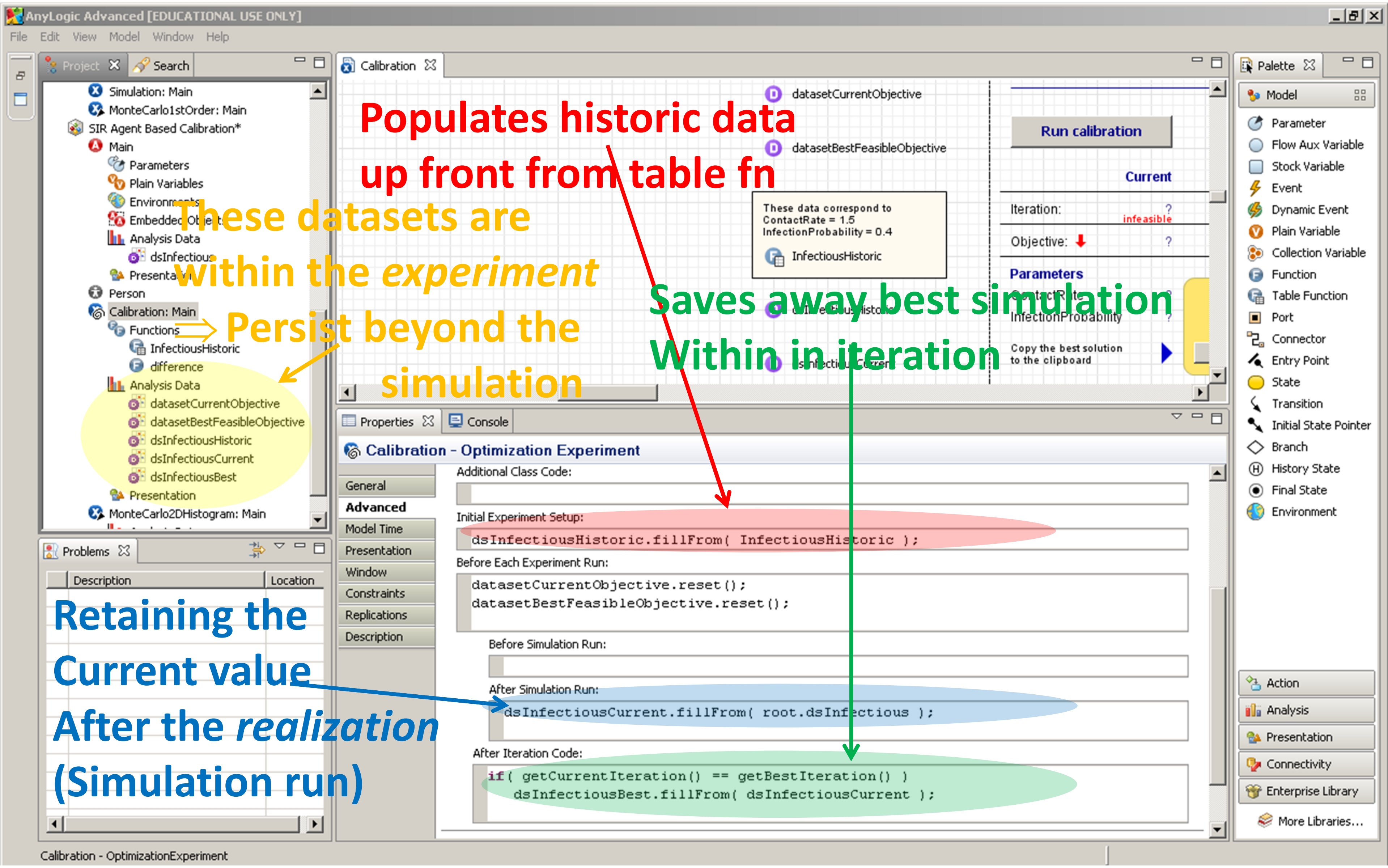
Stochastics in Agent-Based Models

- Recall that ABMs typically exhibit significant stochastics
  - Event timing within & outside of agents
  - Inter-agent interactions
- When calibrating an ABM, we wish to avoid attributing a good match to a particular set of parameter values simply due to chance
- To reliably assess fit of a given set of parameters, we need to repeatedly run model realizations
  - We can take the mean fit of these realizations

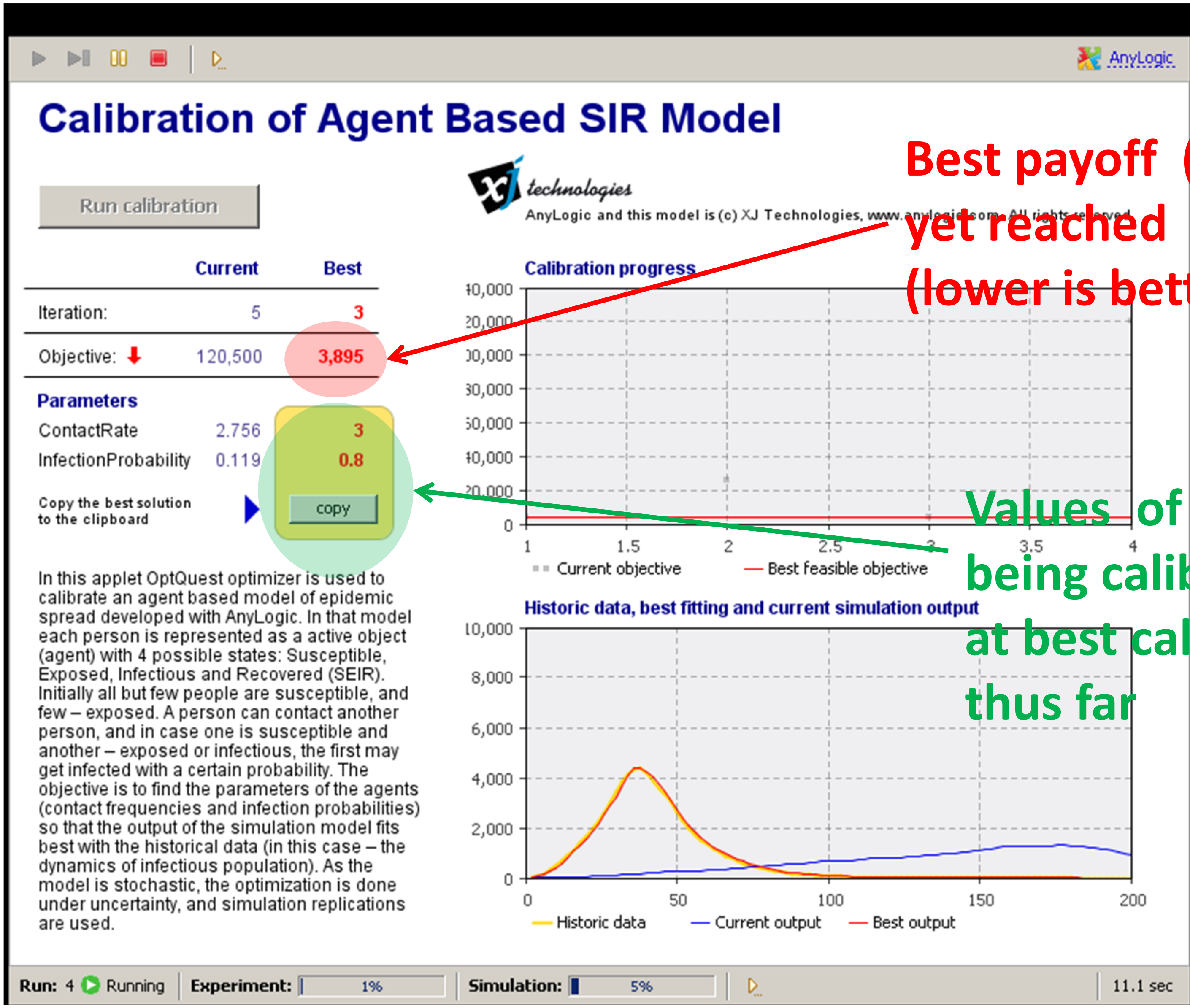
Distinction

- Replication/”Run”: One realization
  - Particular random number seed
- Iteration: Evaluation of a particular parameter set
  - This can contain many realizations (“replications”)
- Confusingly, the term “simulation” appears to sometimes be used for either of the above

Populating the Appropriate Datasets



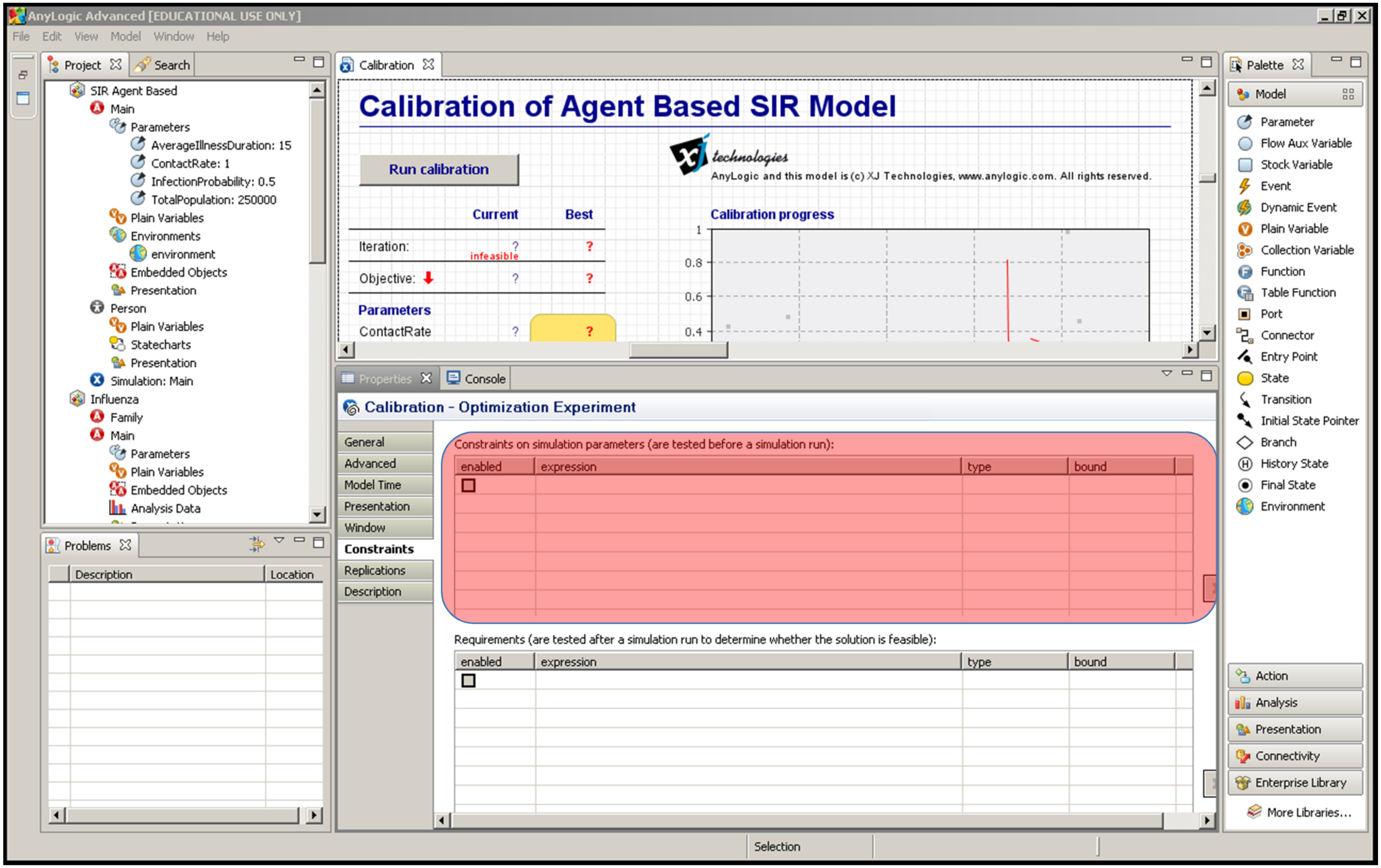
Running Calibration in AnyLogic



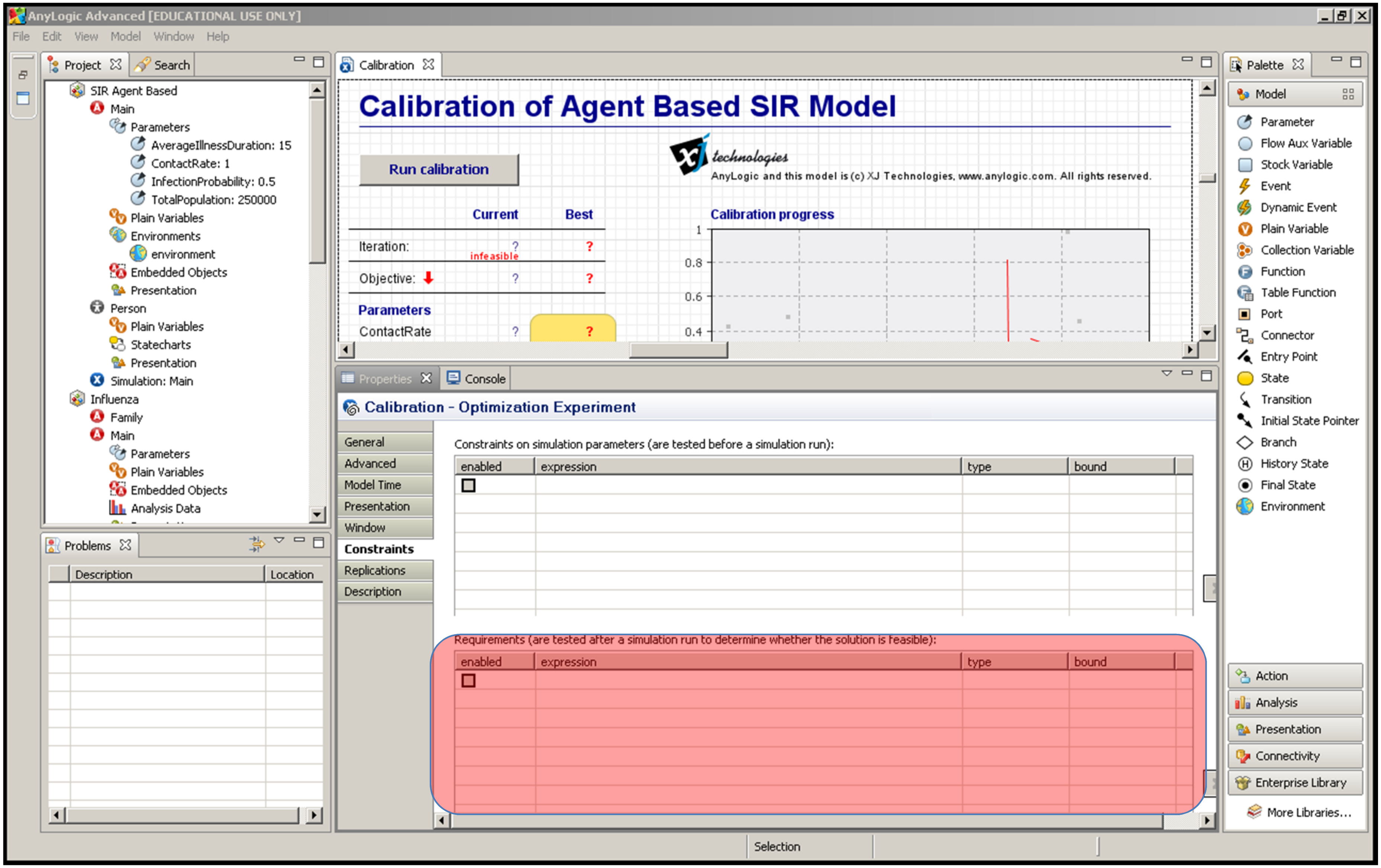
Best payoff (objective) yet reached (lower is better)

Values of parameters being calibrated at best calibration thus far

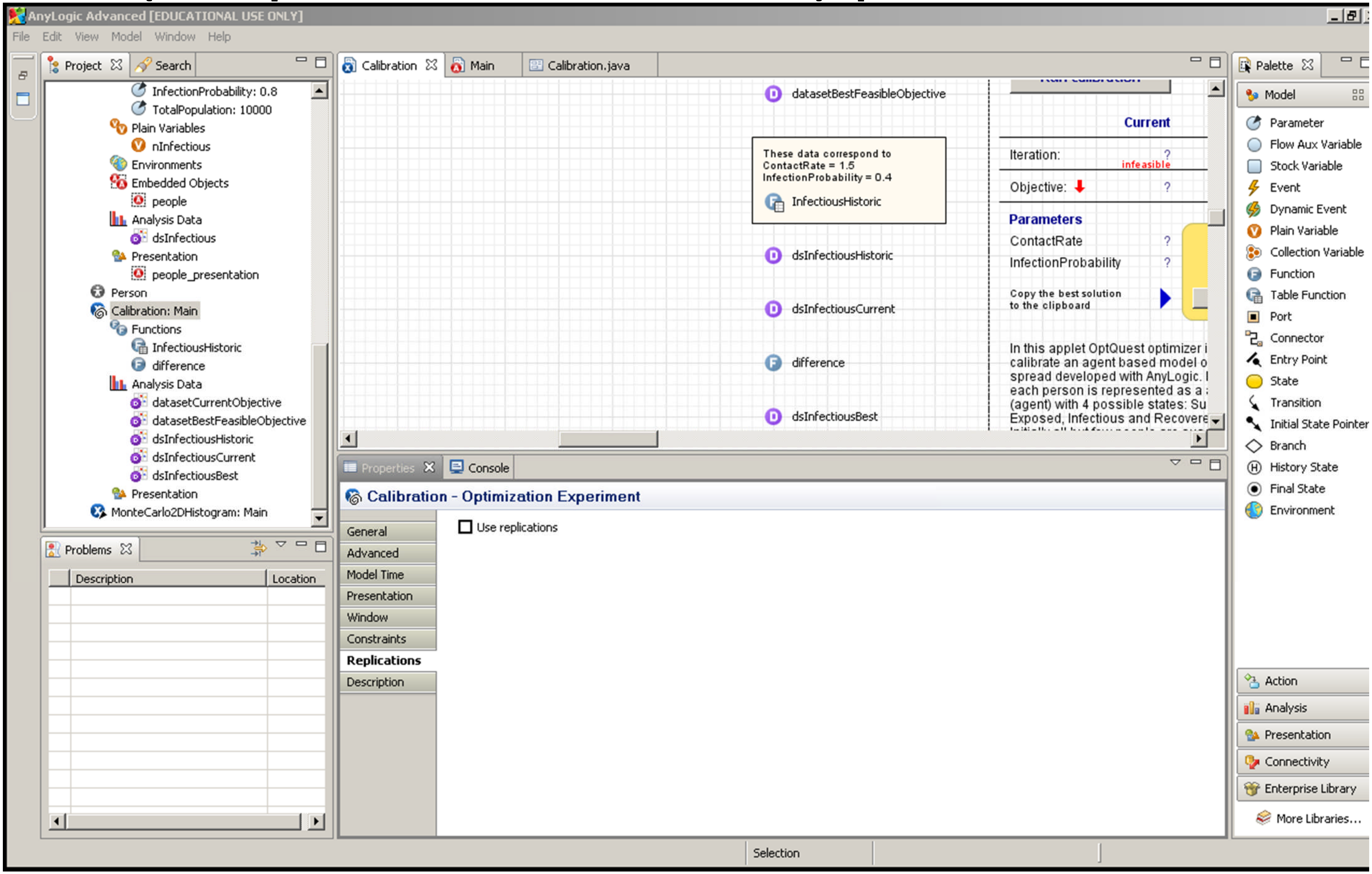
Optimization Constraints – Tests on Legitimacy of Parameter Values



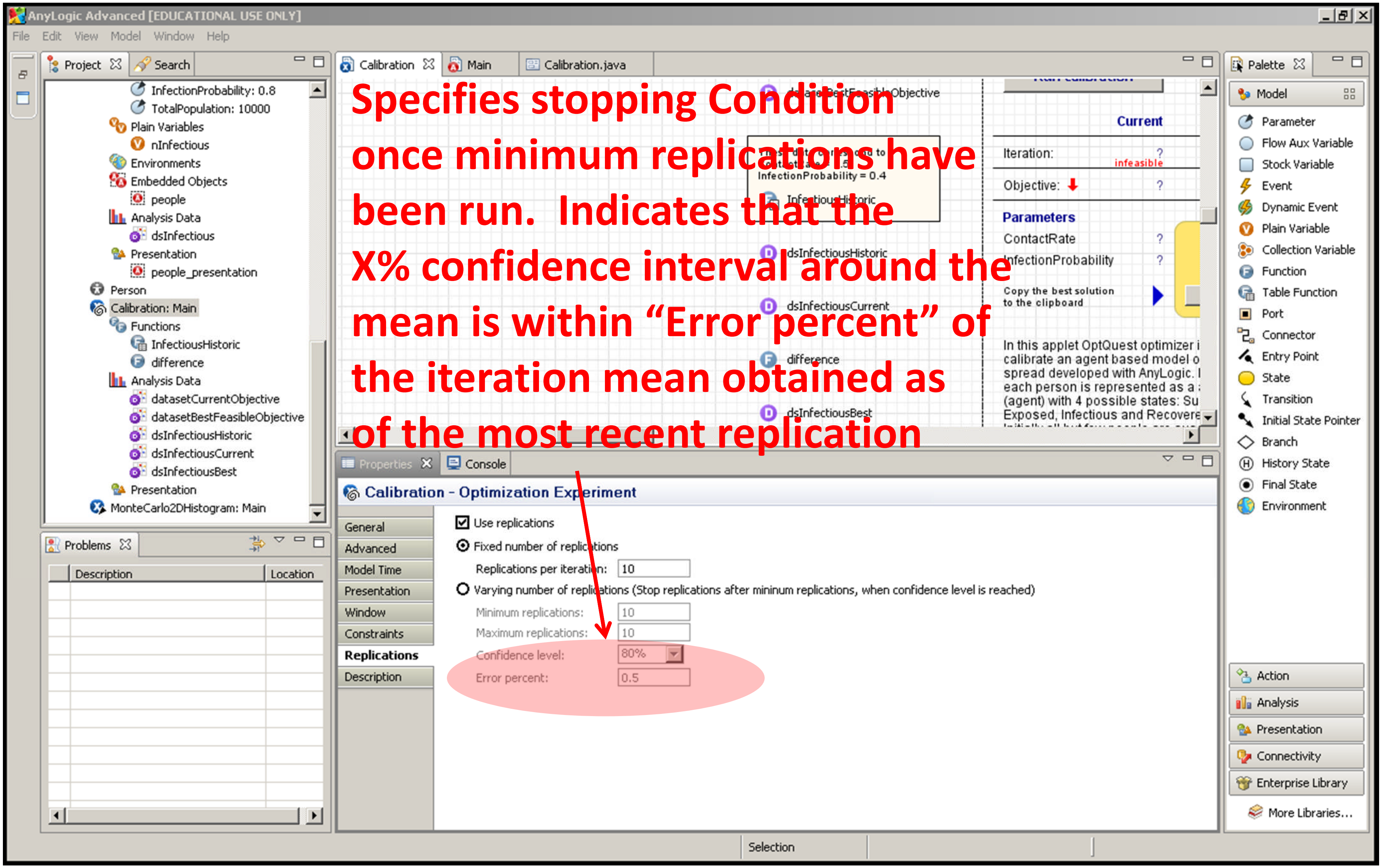
Optimization Requirements – Tests on Emergent results to Sense Validity



Enabling Multiple Realizations ("Replications","Runs") per Iteration

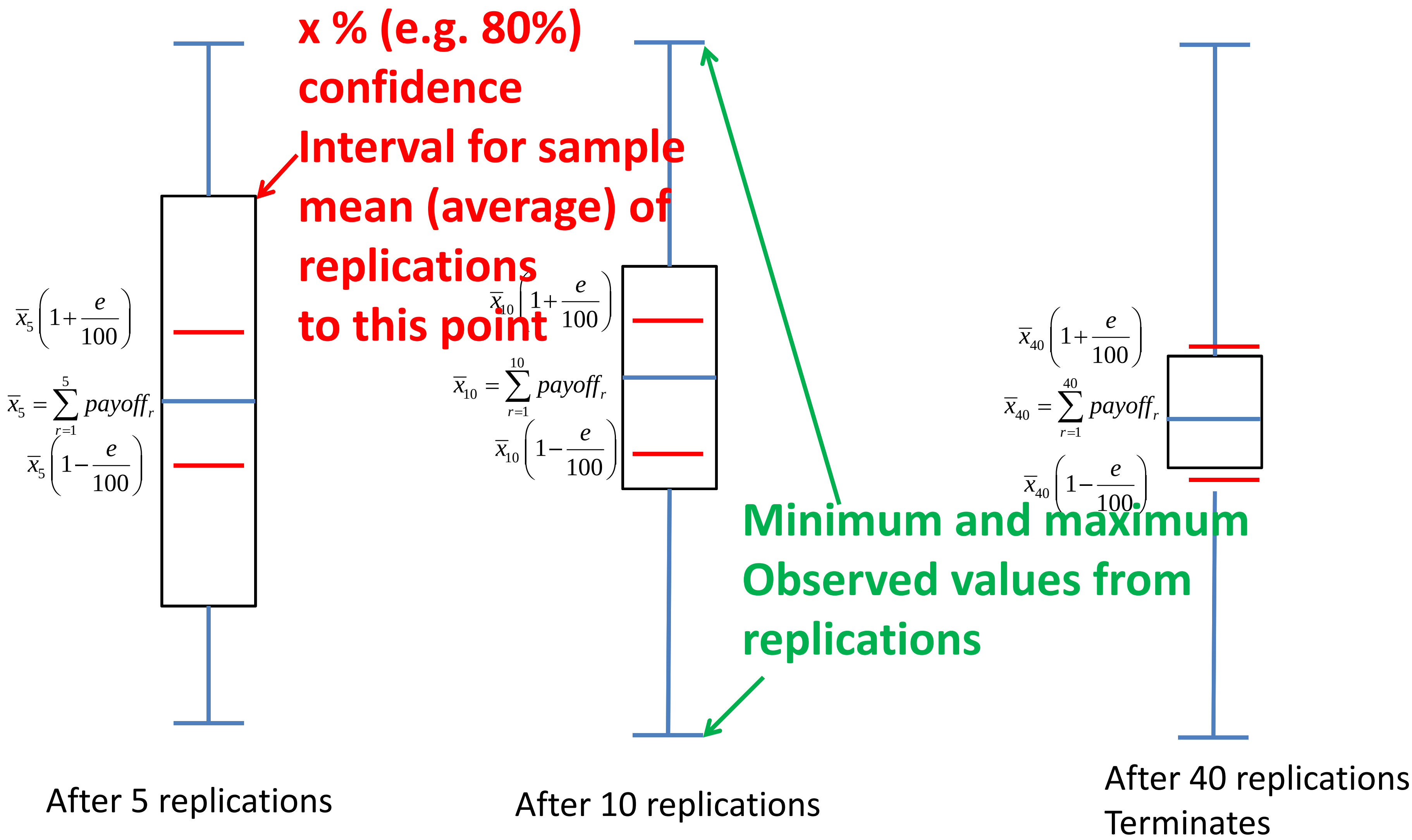


Fixed Number of Replications per Iteration

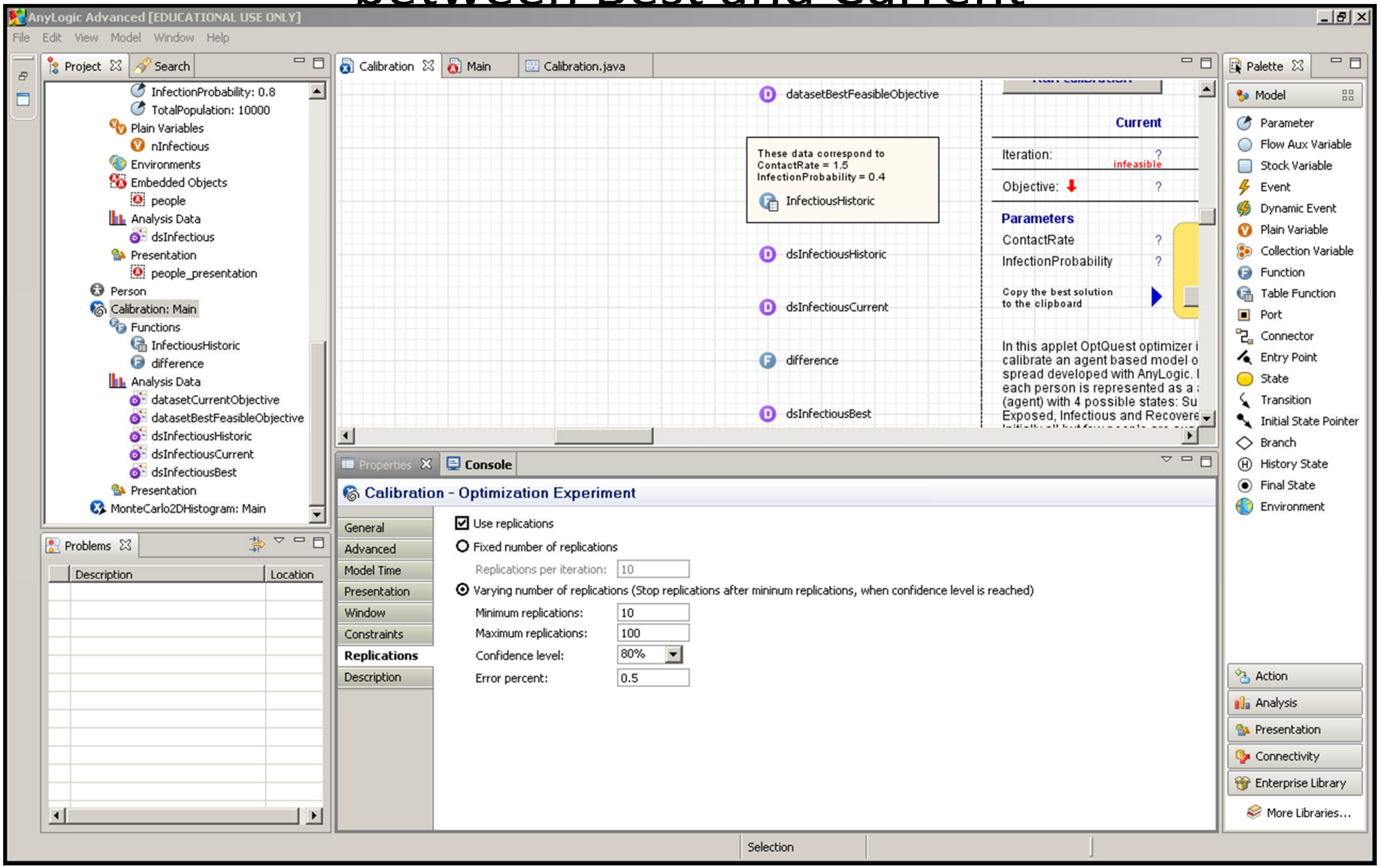


Specifies stopping Condition once minimum replications have been run. Indicates that the X% confidence interval around the mean is within “Error percent” of the iteration mean obtained as of the most recent replication

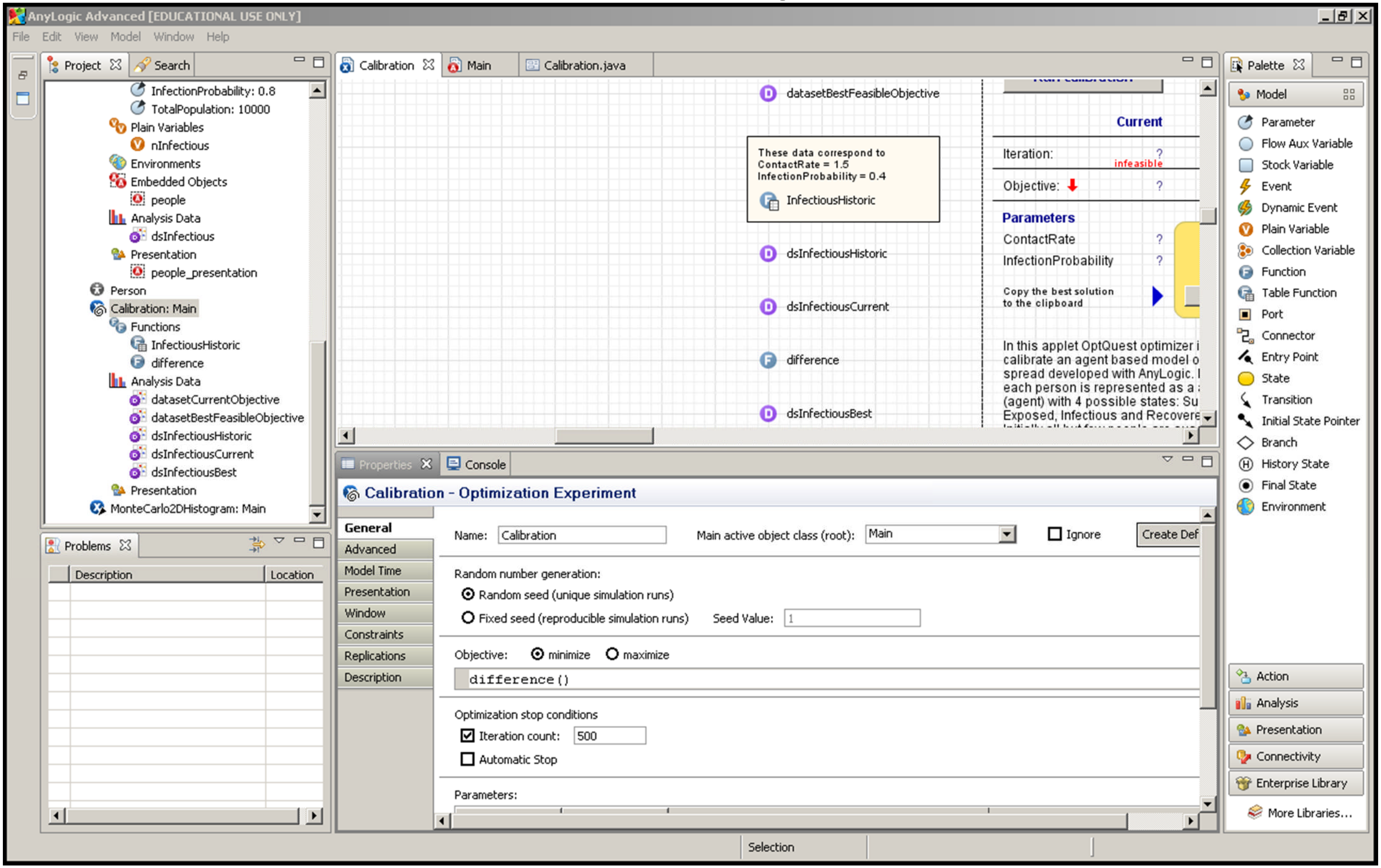
Example


x % (e.g. 80%) confidence Interval for sample mean (average) of replications to this point
Minimum and maximum Observed values from replications
$\bar{x}_5\left(1+\frac{e}{100}\right)$
$\bar{x}_5 = \sum_{r=1}^{5} payoff_r$
$\bar{x}_5\left(1-\frac{e}{100}\right)$
$\bar{x}_{10}\left(1+\frac{e}{100}\right)$
$\bar{x}_{10} = \sum_{r=1}^{10} payoff_r$
$\bar{x}_{10}\left(1-\frac{e}{100}\right)$
$\bar{x}_{40}\left(1+\frac{e}{100}\right)$
$\bar{x}_{40} = \sum_{r=1}^{40} payoff_r$
$\bar{x}_{40}\left(1-\frac{e}{100}\right)$
After 5 replications
After 10 replications
After 40 replications
Terminates

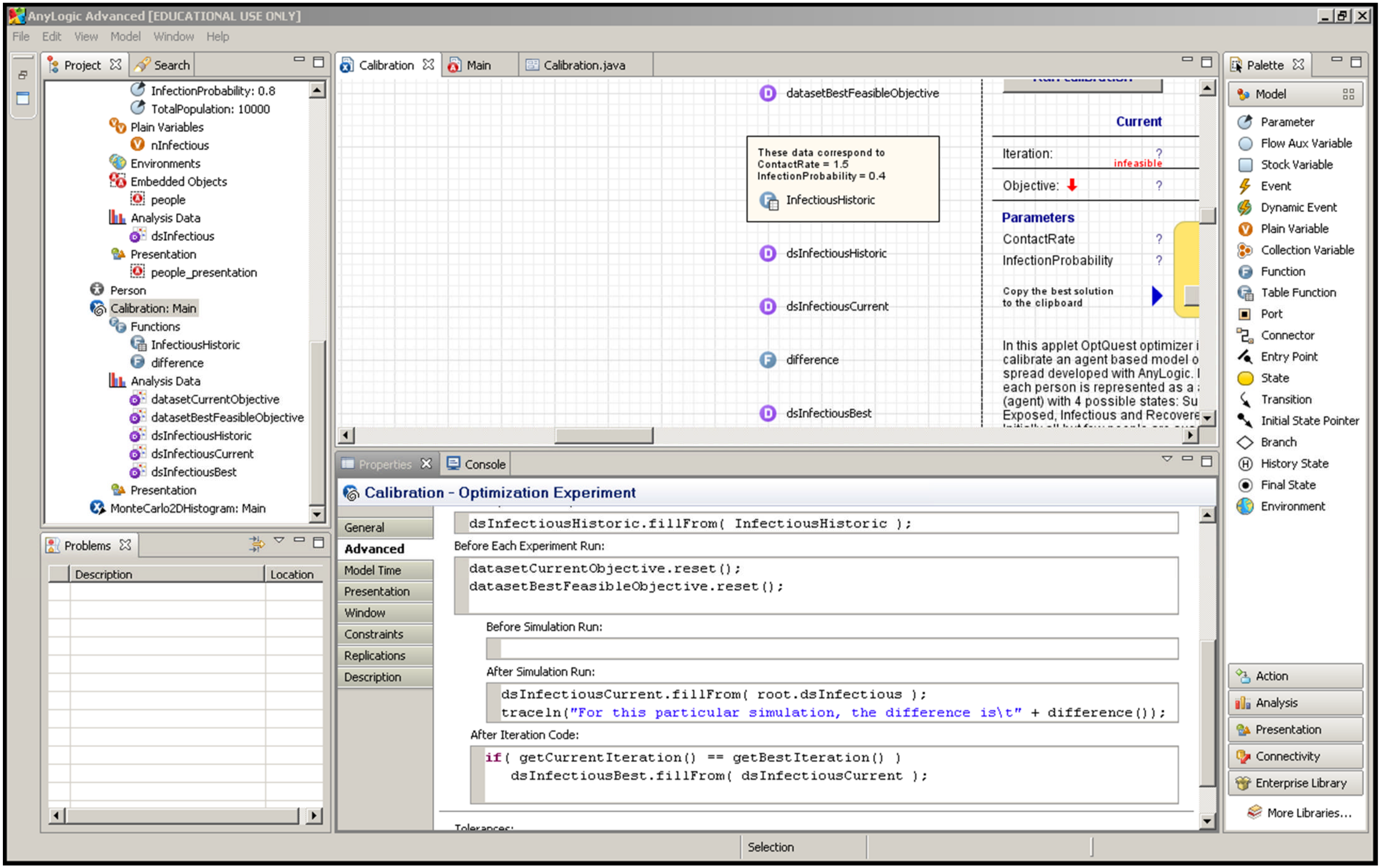
Automatic Throttling of Replications Based on Empirical Fractiles for the Average of the Differences between Best and Current



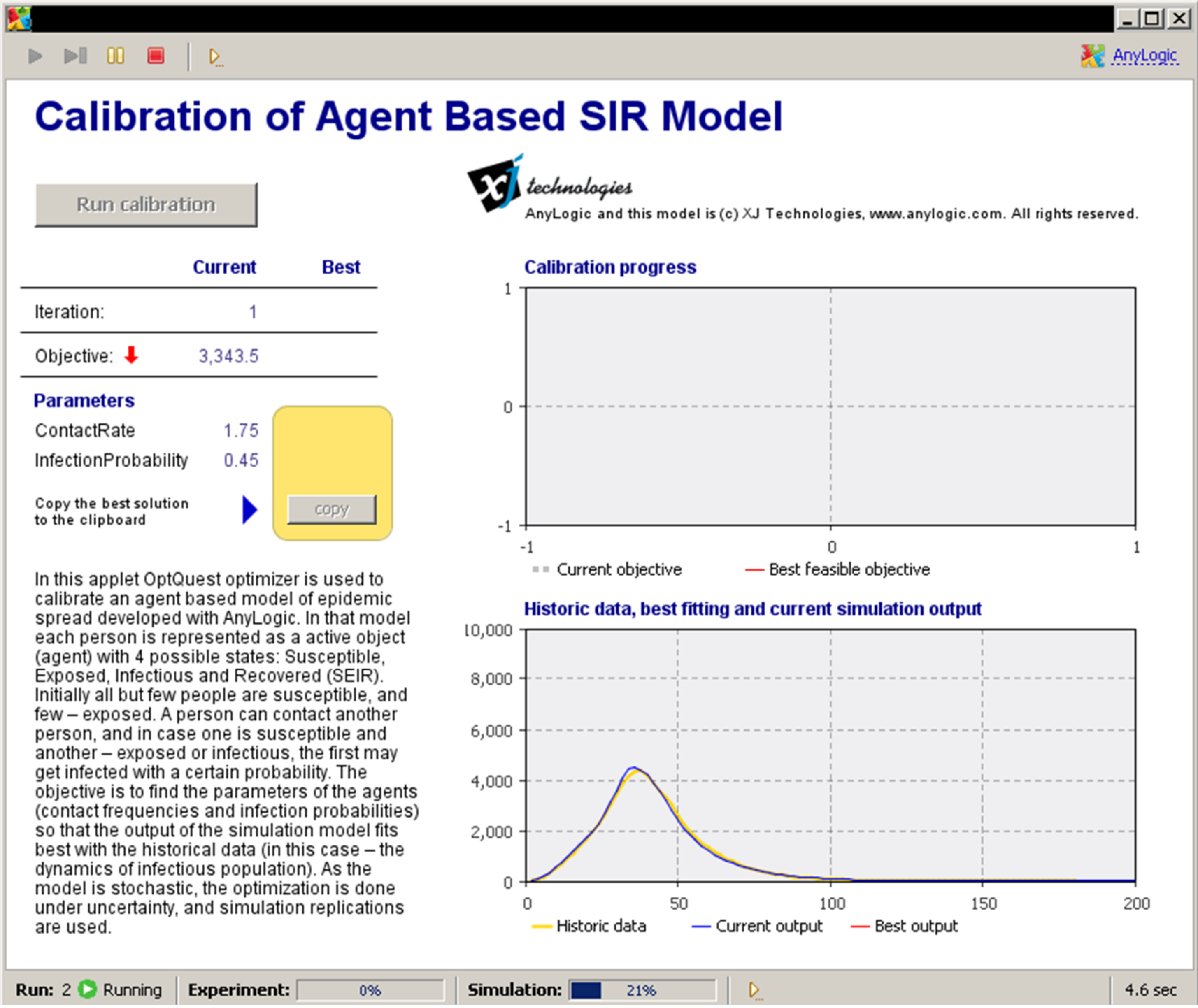
Enabling Random Variation Between Realizations ("Replications")



Understanding Replications: Report Results for Each Replication!

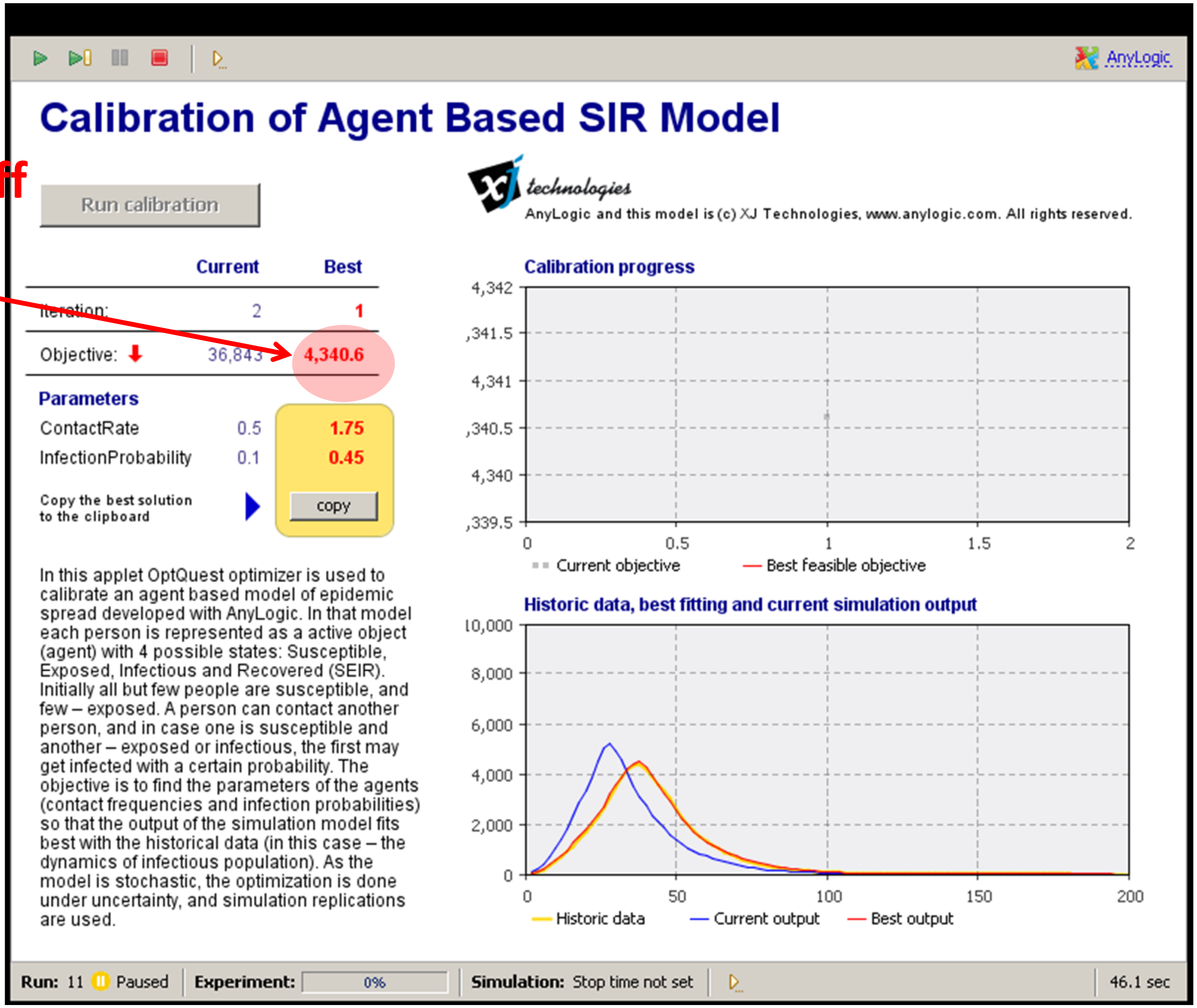


During First Several Realizations ("Replications", "Runs"), No Results Appear

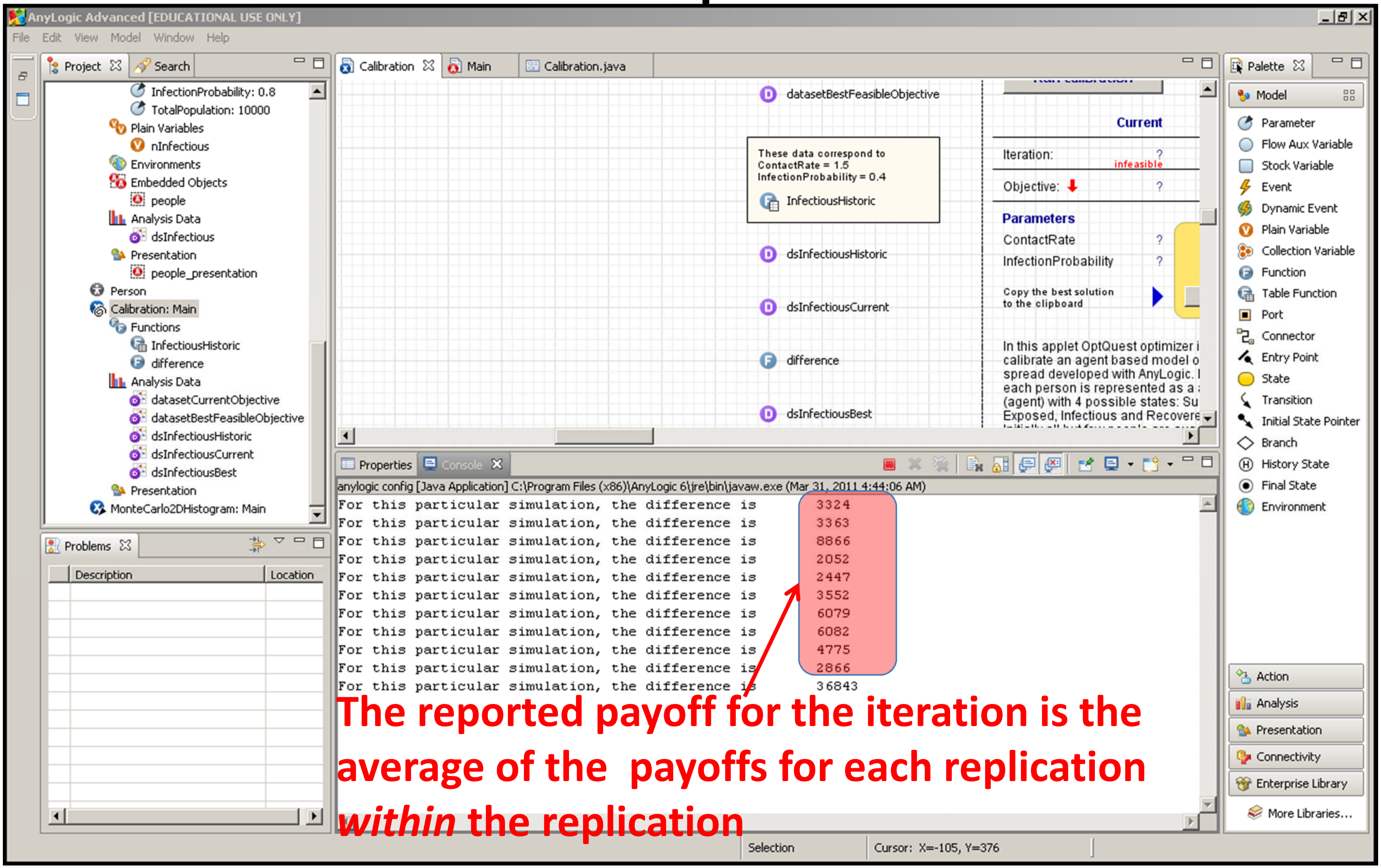


Report on Iteration 1 Appears after a Count of Runs Equal to Replications per Iteration

Reports best payoff (objective) yet reached (lower is better), but from where did this number Come?



Output


AnyLogic Advanced [EDUCATIONAL USE ONLY]
File Edit View Model Window Help
Project Search
InfectionProbability: 0.8
TotalPopulation: 10000
Plain Variables
nInfectious
Environments
Embedded Objects
people
Analysis Data
dsInfectious
Presentation
people_presentation
Person
Calibration: Main
Functions
InfectiousHistoric
difference
Analysis Data
datasetCurrentObjective
datasetBestFeasibleObjective
dsInfectiousHistoric
dsInfectiousCurrent
dsInfectiousBest
Presentation
MonteCarlo2DHistogram: Main
Problems
Description
Location
Calibration
Main
Calibration.java
datasetBestFeasibleObjective
These data correspond to ContactRate = 1.5 InfectionProbability = 0.4
InfectiousHistoric
dsInfectiousHistoric
dsInfectiousCurrent
difference
dsInfectiousBest
Current
Iteration:
infeasible
Objective:
Parameters
ContactRate
InfectionProbability
Copy the best solution to the clipboard
In this applet OptQuest optimizer i calibrate an agent based model o spread developed with AnyLogic. I each person is represented as a (agent) with 4 possible states: Su Exposed, Infectious and Recovere
Properties
Console
anylogic config [Java Application] C:\Program Files (x86)\AnyLogic 6\jre\bin\javaw.exe (Mar 31, 2011 4:44:06 AM)
For this particular simulation, the difference is 3324
For this particular simulation, the difference is 3363
For this particular simulation, the difference is 8866
For this particular simulation, the difference is 2052
For this particular simulation, the difference is 2447
For this particular simulation, the difference is 3552
For this particular simulation, the difference is 6079
For this particular simulation, the difference is 6082
For this particular simulation, the difference is 4775
For this particular simulation, the difference is 2866
For this particular simulation, the difference is 36843
Palette
Model
Parameter
Flow Aux Variable
Stock Variable
Event
Dynamic Event
Plain Variable
Collection Variable
Function
Table Function
Port
Connector
Entry Point
State
Transition
Initial State Pointer
Branch
History State
Final State
Environment
Action
Analysis
Presentation
Connectivity
Enterprise Library
More Libraries...
Selection
Cursor: X=-105, Y=376
The reported payoff for the iteration is the
average of the payoffs for each replication
within the replication

Average of Results for Replications is the Reported Score for the Iteration!