G54SOD (Spring 2018)

Lecture 09
Experimentation

Peer-Olaf Siebers

pos@cs.nott.ac.uk

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Motivation

- Experimentation Preparation
  - Understanding the need for obtaining accurate results
  - Looking at methods for supporting obtaining accurate results
  - Improving the understanding of the model (Sensitivity Analysis)
  - Improving the efficiency of the experimentation process
- Experiments
  - Running experiments in AnyLogic
  - SimHeuristics
    - Tackling combinatorial optimisation problems with simulation
- Guest Speakers
  - Get an idea about PhD projects

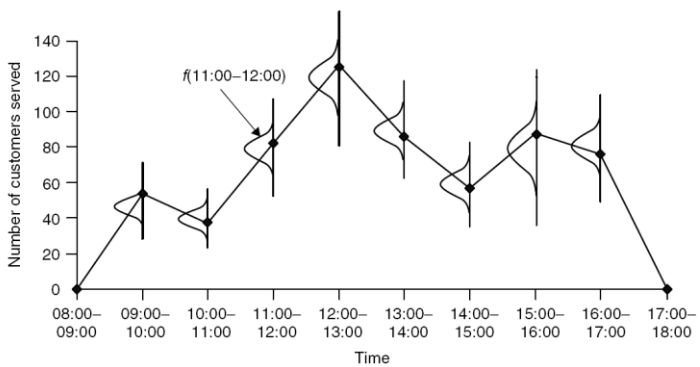
Experimentation Preparation

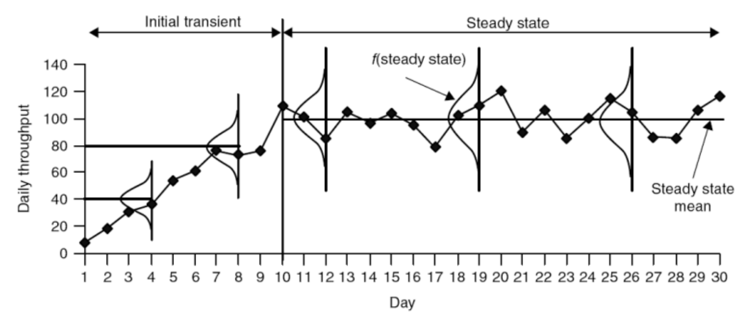
Experimentation Preparation

- What is experimentation preparation about?
  - Obtaining accurate results (for stochastic simulation)
    - Obtaining accurate results on the performance of the model (estimate of average performance and its variability).
    - This says nothing about how accurately the model predicts the performance of the real system
      - Use black box validation for this

Types of Systems

- Types of Systems
  - **Terminating** (natural end point that determines the length of a run) vs. **non terminating** (there is no specific reason why the simulation experiment should terminate)
  - **Transient** (the distribution of the output is constantly changing; true for most terminating simulations) vs. **steady state** (the output is varying to some fixed (steady-state) distribution; true for most non-terminating simulations)





Robinson (2004)

Experimentation Preparation

1. Dealing with initialisation bias (non-terminating simulations only)

– Solutions:
  a) Run model for a warm-up period: Running the model until it reaches a realistic condition and only collect results from the model after this point
  b) Set initial conditions in the model: Place the model in realistic conditions at the start of the simulation run [not practical]

2. Obtaining sufficient output data (terminating and non terminating simulations)

– Solutions:
  a) Multiple replications (terminating or non-terminating simulations): Equivalent to taking multiple samples in statistics; multiple runs of the simulation model with different random number streams
  b) Single long run (non-terminating simulations): Equivalent to taking one large sample in statistics [not practical; some statistical concerns]

Dealing with Initialisation Bias

1a. Running the model for a warm-up period

- Needs to be long enough to ensure that the model is in a realistic condition (the difficulty lies in determining whether the model is in a realistic condition)

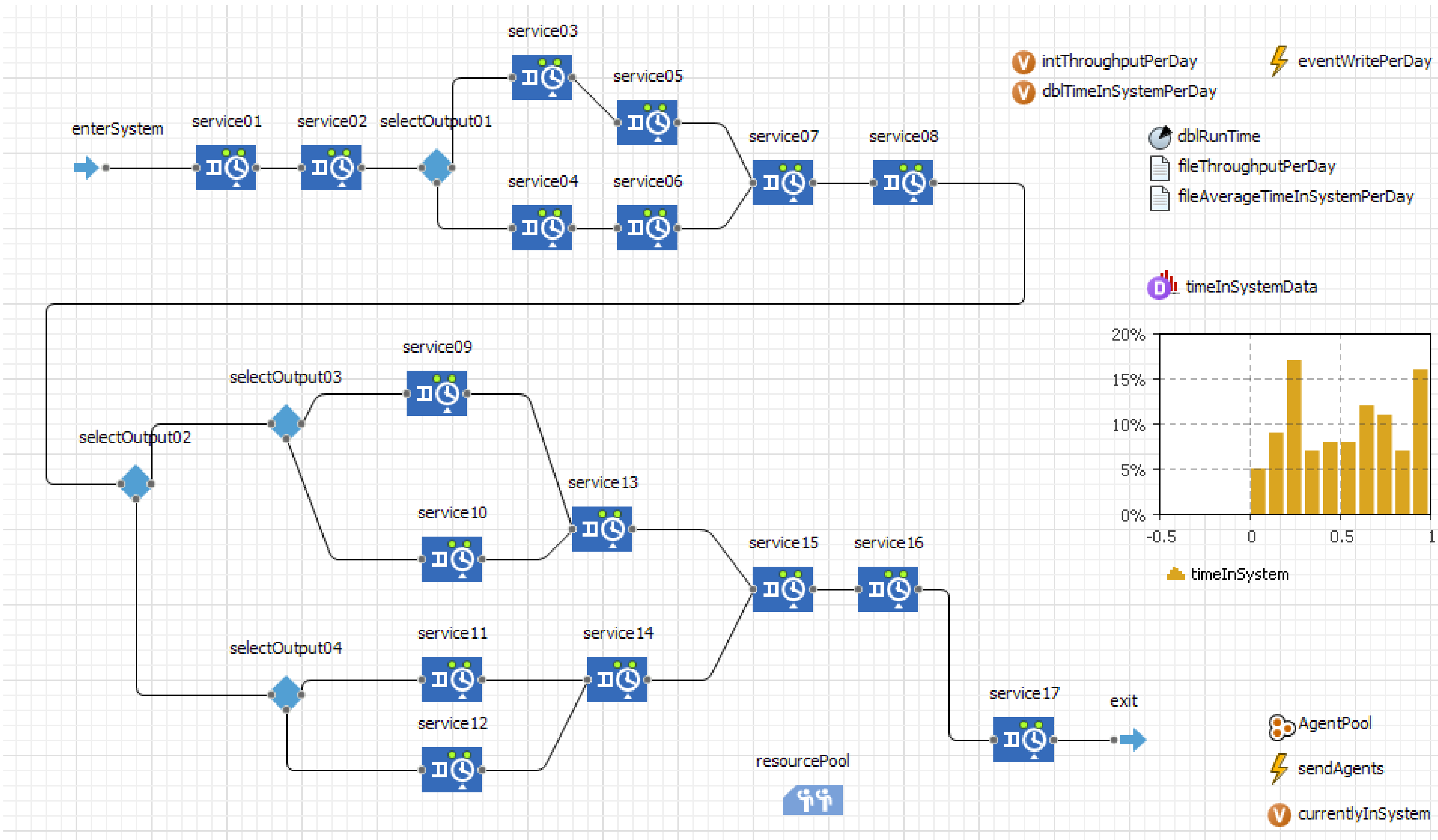
- Method categories for determining the warm-up period length (Robinson, 2002)
  - Graphical methods; heuristic approaches; statistical methods; initialisation bias tests; hybrid methods

- Most commonly used methods for estimating the warm-up period
  - **Time series inspection**
  - **Welch's method**

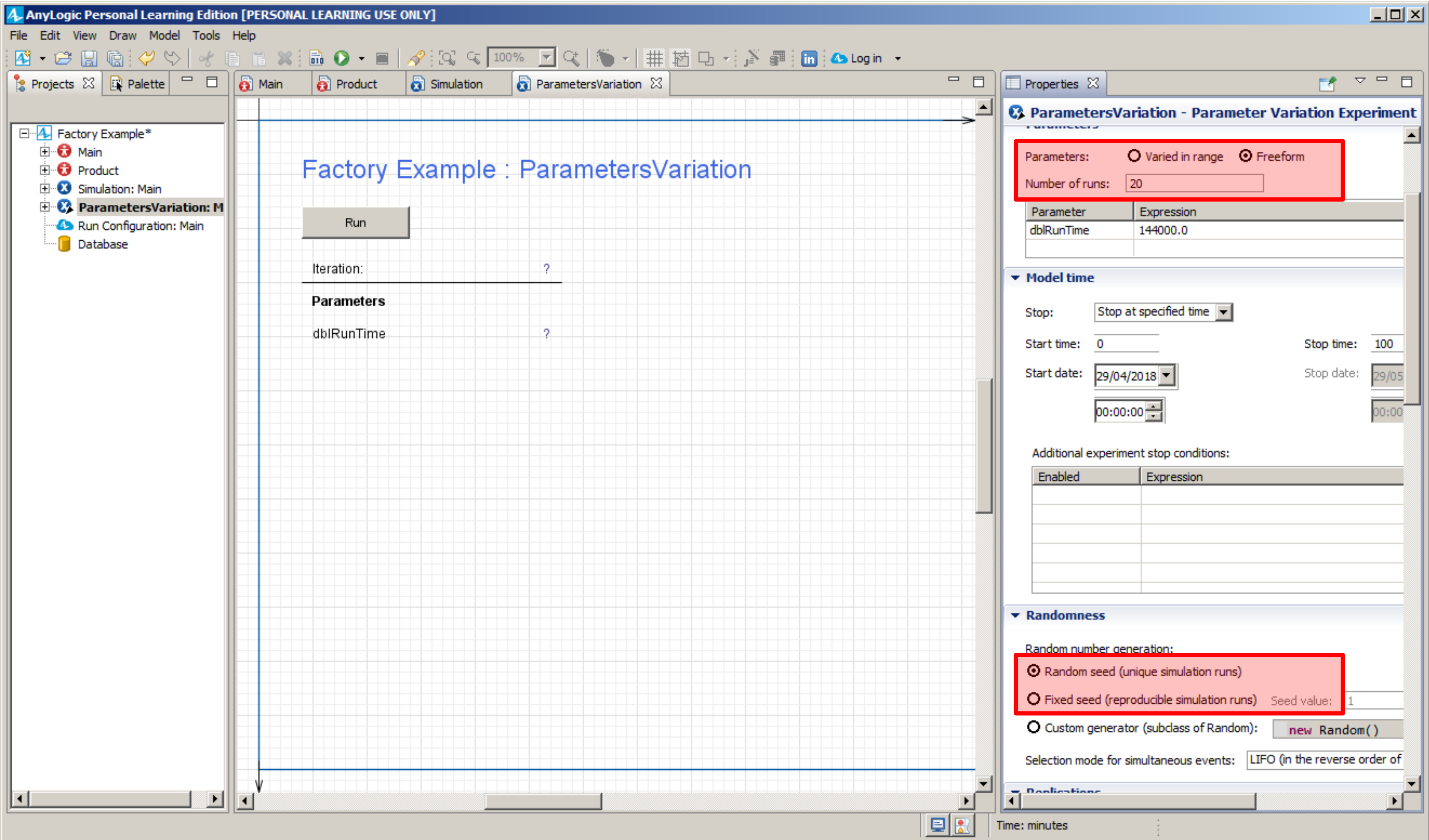
Dealing with Initialisation Bias

1a. Time series inspection (see www.wileyeurope.com/go/robinson - warmup.xls)

- Problem: Data from a single run can be very noisy
- Solution: Conduct multiple runs and inspect the mean averages of replications; the more replications you run the more the time series will be smoothed
- Shortcomings:
  - Relies on subjective assessment
  - If data are particularly noisy, subtle patterns in the data might go unnoticed



The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA


AnyLogic Personal Learning Edition [PERSONAL LEARNING USE ONLY]
File Edit View Draw Model Tools Help
Projects
Palette
Main
Product
Simulation
ParametersVariation
Factory Example*
Main
Product
Simulation: Main
ParametersVariation: M
Run Configuration: Main
Database
Factory Example : ParametersVariation
Run
Iteration:
Parameters
dblRunTime
Properties
ParametersVariation - Parameter Variation Experiment
Parameters:
Varied in range
Freeform
Number of runs:
20
Parameter
Expression
dblRunTime
144000.0
Model time
Stop:
Stop at specified time
Start time:
0
Stop time:
100
Start date:
29/04/2018
Stop date:
00:00:00
Additional experiment stop conditions:
Enabled
Expression
Randomness
Random number generation:
Random seed (unique simulation runs)
Fixed seed (reproducible simulation runs)
Seed value:
1
Custom generator (subclass of Random):
new Random()
Selection mode for simultaneous events:
LIFO (in the reverse order of
Time: minutes

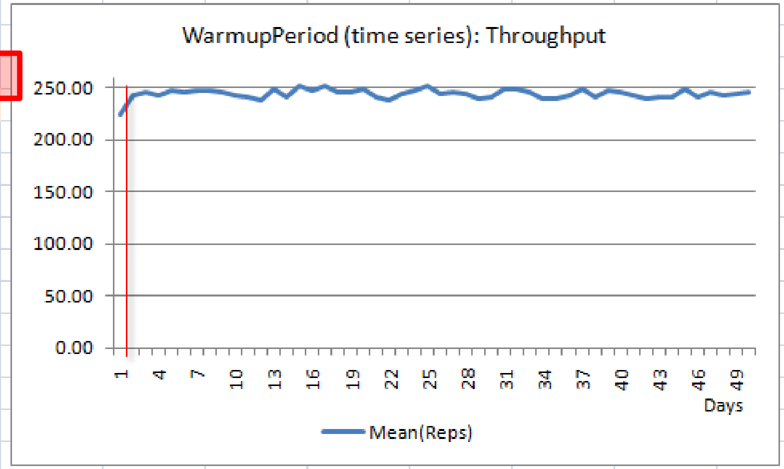
ThroughputPerDay

Day	Rep 1	Rep 2	Rep 3	Rep 4	Rep 5	Rep 6	Rep 7	Rep 8	Rep 9	Rep 10	Rep 11	Rep 20	Mean(Reps)	StDev(Reps)
1	254	243	228	225	214	220	226	205	193	244	231	250	223.65	19.62
2	261	228	255	262	260	237	232	237	230	257	247	231	242.35	14.06
3	230	263	240	252	239	247	246	241	223	208	251	246	245.45	20.44
4	241	235	244	242	267	239	246	242	269	246	248	217	243.00	16.09
5	252	261	235	231	253	245	271	238	245	244	236	251	247.30	11.65
6	220	218	244	266	244	276	246	255	230	258	237	244	246.30	17.99
7	251	263	255	238	234	237	260	274	235	250	255	221	247.05	12.75
8	257	241	256	236	281	257	238	241	249	241	242	258	247.05	12.10
9	250	256	246	257	248	223	228	232	264	254	264	242	245.55	13.16
10	255	242	250	260	218	245	239	222	269	250	237	233	242.60	14.47
11	207	266	231	247	254	241	246	250	232	243	249	262	240.90	17.23
12	265	215	239	235	282	254	256	225	222	247	229	245	238.70	20.81
13	235	262	236	251	253	253	262	231	269	269	223	225	249.30	14.26
100	234	250	234	240	233	236	257	235	225	286	236			
Mean(Days)	246.02	243.51	244.79	243.99	245.65	243.66	243.52	241.05	242.98	245.94	245.24			
StDev(Days)	16.03	17.59	13.00	14.82	14.91	15.12	16.76	14.48	15.49	15.70	14.63			

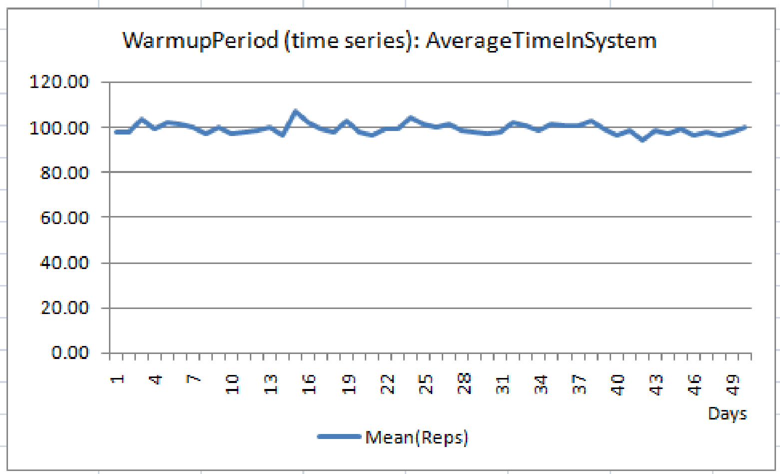
AverageTimeInSystemPerDay

Day	Rep 1	Rep 2	Rep 3	Rep 4	Rep 5	Rep 6	Rep 7	Rep 8	Rep 9	Rep 10	Rep 11	Rep 20	Mean(Reps)	StDev(Reps)
1	134.73	94.10	94.13	96.27	92.23	93.07	91.30	85.83	94.98	115.86	92.93	138.76	98.05	14.68
2	98.81	90.92	100.21	105.72	115.85	94.37	107.07	89.87	100.44	89.84	96.70	99.22	98.23	7.90
3	87.67	112.62	92.11	115.71	92.60	96.68	88.49	91.93	90.69	86.49	92.98	92.96	103.90	29.50
4	92.21	88.46	105.28	91.99	108.48	90.01	97.76	93.40	113.16	98.45	103.44	98.60	99.27	9.30
5	100.70	107.43	95.88	88.65	96.72	130.39	139.82	90.52	92.04	91.25	88.74	99.30	101.78	15.35
6	89.07	92.50	102.44	101.06	95.62	137.93	90.06	126.58	87.30	94.91	95.71	91.60	101.10	13.59
7	99.45	96.56	101.53	88.83	91.92	88.74	108.22	122.74	94.15	120.29	110.42	87.29	100.34	12.67
8	98.88	102.23	94.27	89.76	123.28	97.78	96.01	95.10	94.56	102.31	91.87	99.32	97.13	8.10
9	95.21	94.99	113.02	96.21	95.53	85.89	96.73	89.09	155.85	123.31	99.98	90.66	99.70	15.96
10	93.84	90.74	95.84	103.43	88.66	102.15	91.69	87.60	121.74	134.60	94.71	90.39	97.19	11.97
11	86.88	111.63	91.06	95.33	120.87	102.23	98.13	102.78	91.09	103.32	91.93	128.31	97.58	11.44
12	94.73	89.80	102.91	93.86	129.82	97.87	119.70	89.07	98.30	103.04	94.18	101.25	98.46	10.53
13	90.22	100.24	96.45	103.32	96.72	98.48	109.35	87.10	103.86	126.29	93.32	93.48	100.10	9.82
100	92.95	98.20	94.62	91.02	88.80	97.61	97.57	88.19	86.64	207.19	90.65			
Mean(Days)	100.66	97.68	99.38	98.89	99.41	97.81	99.88	95.88	98.50	101.50	99.23			
StDev(Days)	13.49	10.43	11.29	10.86	11.59	11.39	16.19	8.60	11.81	17.30	11.80			

WarmupPeriod	
ThroughputPerDay	
Day	Mean(Reps)
1	223.65
2	242.35
3	245.45
4	243.00
5	247.30
6	246.30
7	247.05
8	247.05
9	245.55
10	242.60
11	240.90
12	238.70



WarmupPeriod	
AverageTimeInSystem	
Day	Mean(Reps)
1	98.05
2	98.23
3	103.90
4	99.27
5	101.78
6	101.10
7	100.34
8	97.13
9	99.70
10	97.19
11	97.58
12	98.46



Dealing with Initialisation Bias

1a. Welch's method (see www.wileyeurope.com/go/robinson - warmup.xls)

- Steps:
  - Perform a series of replications (at least 5)
  - Calculate mean of the output data across the replications for each period
  - Calculate the moving average based on a window size w (start with 5)
  - Plot the moving average on a time series
    - If the data are smooth identify the warm up period as the point where the time series becomes flat
    - If the data is not smooth increase window size w and calculate new moving averages
- Advantage: Calculation of moving averages smoothes out noise
- Shortcomings:
  - Relies on subjective assessment
  - Conservative; tends to overestimate the warm up period

Dealing with Initialisation Bias

The moving averages are calculated using the following formula:

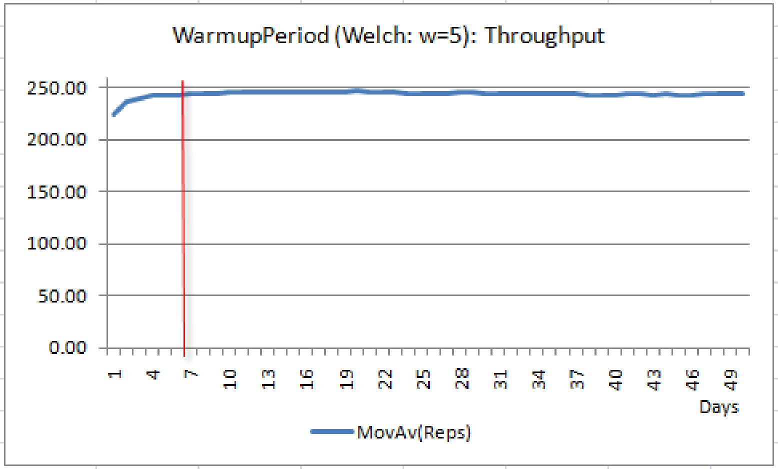
$$\overline{Y}_i(w) = \begin{cases} \dfrac{\sum_{s=-(i-1)}^{i-1} \overline{Y}_{i+s}}{2i-1} & \text{if } i = 1, \ldots, w \\[2ex] \dfrac{\sum_{s=-w}^{w} \overline{Y}_{i+s}}{2\,w+1} & \text{if } i = w+1, \ldots, m-w \end{cases}$$

where:

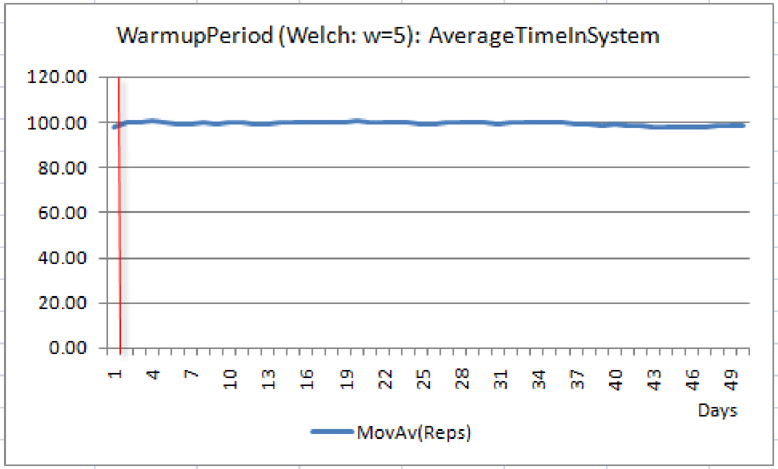
$\overline{Y}_i(w)$ = moving average of window size w
$\overline{Y}_i$ = time-series of output data (mean of the replications)
i = period number
m = number of periods in the simulation run

Robinson (2004)

WarmupPeriod		
ThroughputPerDay: w=5		
Day	Mean(Reps)	MovAv(Rep)
1	223.65	223.65
2	242.35	237.15
3	245.45	240.35
4	243.00	242.16
5	247.30	243.08
6	246.30	242.84
7	247.05	244.20
8	247.05	244.84
9	245.55	244.50
10	242.60	245.27
11	240.90	245.22
12	238.70	245.71



WarmupPeriod		
AverageTimeInSystem: w=5		
Day	Mean(Reps)	MovAv(Rep)
1	98.05	98.05
2	98.23	100.06
3	103.90	100.24
4	99.27	100.38
5	101.78	99.94
6	101.10	99.48
7	100.34	99.52
8	97.13	99.69
9	99.70	98.99
10	97.19	99.72
11	97.58	99.72
12	98.46	99.56



Dealing with Initialisation Bias

1a. Running the model for a warm-up period: Important

- If the model has more than one key response the initial transient should be investigated for each one

- In theory, the warm-up period should be determined separately for every experimental scenario; in practice it is only done for the base scenario!

Run length

- Warm-up period generally takes less than 10 percent of the run length
  - Rule of thumb:
    - The experiment run length should be 10 x warm-up period
    - This implies a total run length of 11 x warm-up period (including 1 x for the warm up period that is removed)

Experimentation Preparation

1. Dealing with initialisation bias
   - Solutions:
     a) Run model for a warm-up period: Running the model until it reaches a realistic condition and only collect results from the model after this point
     b) Set initial conditions in the model: Place the model in realistic conditions at the start of the simulation run [not practical]

2. Obtaining sufficient output data
   - Solutions:
     a) Multiple replications (terminating or non-terminating simulations): Equivalent to taking multiple samples in statistics; multiple runs of the simulation model with different random number streams
     b) Single long run (non-terminating simulations): Equivalent to taking one large sample in statistics [not practical; some statistical concerns]

Obtaining Sufficient Output Data

2a. Multiple replications

- Multiple replications are performed by changing the random number stream
- Producing multiple samples in order to get better estimates for the mean performance
- Most commonly used methods for estimating the number of required replications:
  - **Plotting cumulative means**
  - **Confidence interval method**

Obtaining Sufficient Output Data

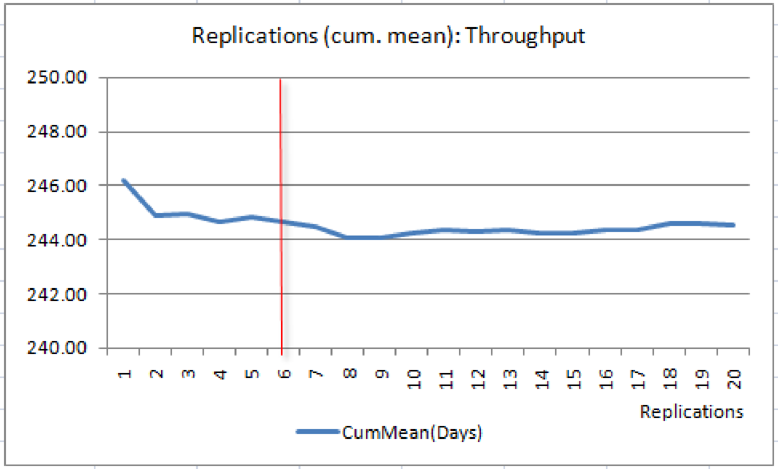
2a. Plotting cumulative means (see www.wileyeurope.com/go/robinson - replications.xls)

- Plotting the cumulative mean of the output data from a series of replications
- As more replications are performed the graph should become a flat line (minimal variability, no upwards and downwards trends)
- Number of replications is defined by the point at which the graph becomes flat

Number of replications

Throughput

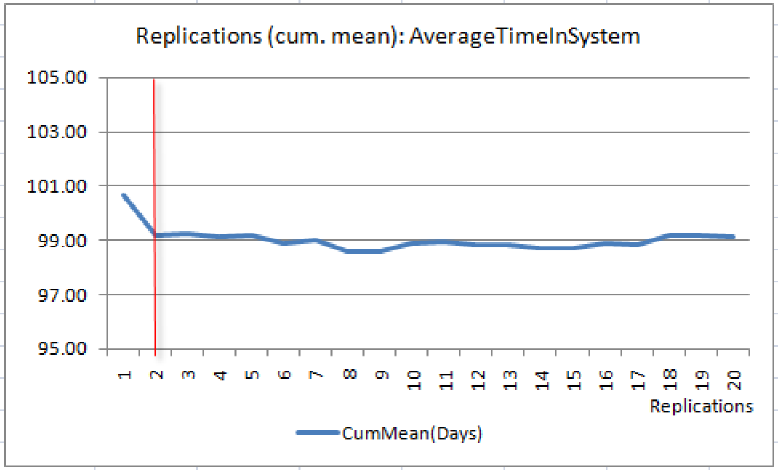
Rep	Mean(Days)	CumMean(Days)
1	246.21	246.21
2	243.65	244.93
3	245.03	244.96
4	243.84	244.68
5	245.62	244.87
6	243.64	244.66
7	243.46	244.49
8	241.35	244.10
9	243.70	244.06
10	246.14	244.26
11	245.47	244.37
12	243.36	244.29



Number of replications

AverageTimeInSystem

Rep	Mean(Days)	CumMean(Days)
1	100.67	100.67
2	97.68	99.18
3	99.45	99.27
4	98.82	99.16
5	99.36	99.20
6	97.22	98.87
7	99.72	98.99
8	95.85	98.60
9	98.63	98.60
10	101.84	98.92
11	99.50	98.98
12	97.39	98.84



Obtaining Sufficient Output Data

2a. Confidence intervals (see www.wileyeurope.com/go/robinson - replications.xls)

- Statistical means for showing how accurately the mean average of a value is estimated
- The narrower the interval the more accurate the estimate (i.e. the smaller the deviation between the upper and lower limit)
- Standard applications: Use 95% confidence interval (sign. level α=5%)
  - This gives a 95% probability that the value of the true mean (obtained if the model is run for an infinite period) lies within the confidence interval
- Critical application: Use 99% confidence interval (sign. level α=1%)

Obtaining Sufficient Output Data

2a. Confidence intervals (see www.wileyeurope.com/go/robinson - replications.xls)

- The modeller needs to decide which %deviation between the upper and the lower limit is acceptable and choose the required number of replications to stay below this %deviation from the statistics

Obtaining Sufficient Output Data

When analysing simulation output data a confidence interval is calculated as follows:

$$CI = \overline{X} \pm t_{n-1,\alpha/2} \frac{S}{\sqrt{n}}$$

where:

$\overline{X}$ = mean of the output data from the replications
S = standard deviation of the output data from the replications (see equation below)
n = number of replications
$t_{n-1,\alpha/2}$ = value from Student's t-distribution with $n-1$ degree of freedom and a significance level of $\alpha/2$

The formula for the standard deviation is:

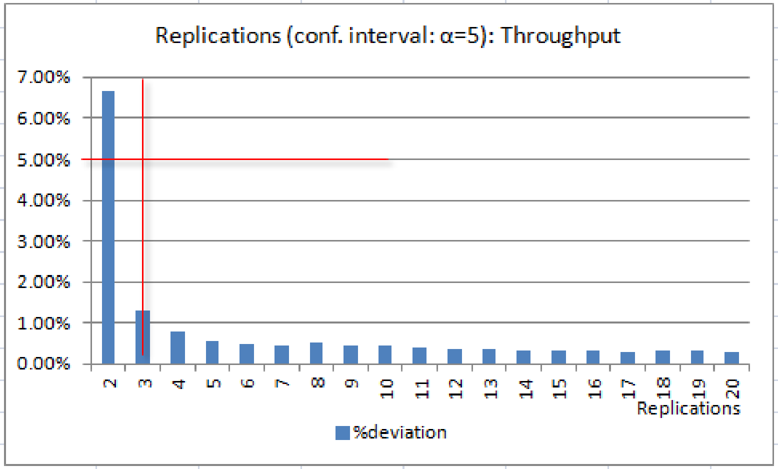
$$S = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}}$$

where:

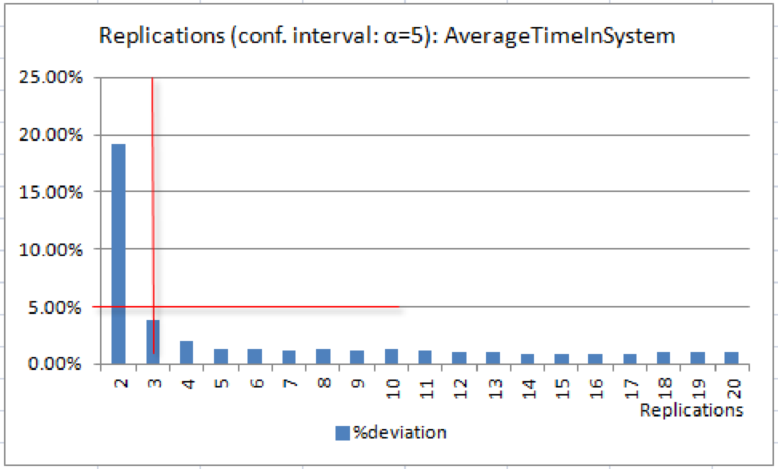
X_i = the result from replication i

Robinson (2004)

Number of replications			alpha	5%		
Throughput						
Rep	Mean(Days)	CumMean(Days)	StdDev(Days)	Lower interval	Upper interval	%deviation
1	246.21	246.21	n/a	n/a	n/a	n/a
2	243.65	244.93	1.81	228.64	261.22	6.65%
3	245.03	244.96	1.28	241.78	248.15	1.30%
4	243.84	244.68	1.19	242.79	246.58	0.77%
5	245.62	244.87	1.11	243.49	246.25	0.56%
6	243.64	244.66	1.11	243.50	245.83	0.48%
7	243.46	244.49	1.11	243.46	245.52	0.42%
8	241.35	244.10	1.52	242.83	245.37	0.52%
9	243.70	244.06	1.42	242.96	245.15	0.45%
10	246.14	244.26	1.50	243.19	245.33	0.44%
11	245.47	244.37	1.46	243.39	245.36	0.40%
12	243.36	244.29	1.43	243.38	245.20	0.37%



Number of replications			alpha	5%		
AverageTimeInSystem						
Rep	Mean(Days)	CumMean(Days)	StdDev(Days)	Lower interval	Upper interval	%deviation
1	100.67	100.67	n/a	n/a	n/a	n/a
2	97.68	99.18	2.12	80.17	118.18	19.16%
3	99.45	99.27	1.50	95.53	103.00	3.76%
4	98.82	99.16	1.25	97.17	101.14	2.00%
5	99.36	99.20	1.08	97.85	100.54	1.36%
6	97.22	98.87	1.26	97.54	100.19	1.34%
7	99.72	98.99	1.20	97.88	100.09	1.12%
8	95.85	98.60	1.57	97.28	99.91	1.33%
9	98.63	98.60	1.47	97.47	99.73	1.14%
10	101.84	98.92	1.72	97.69	100.15	1.24%
11	99.50	98.98	1.64	97.87	100.08	1.11%
12	97.39	98.84	1.63	97.81	99.88	1.05%



Obtaining Sufficient Output Data

2a. Multiple replications: Important

- Run length:
  - Transient models have a defined run length (e.g. one day)
  - Steady state models should be run at least 10x the warm-up period
- Remember to delete the warm-up period data before conducting any further analysis if you have a non-terminating simulation!

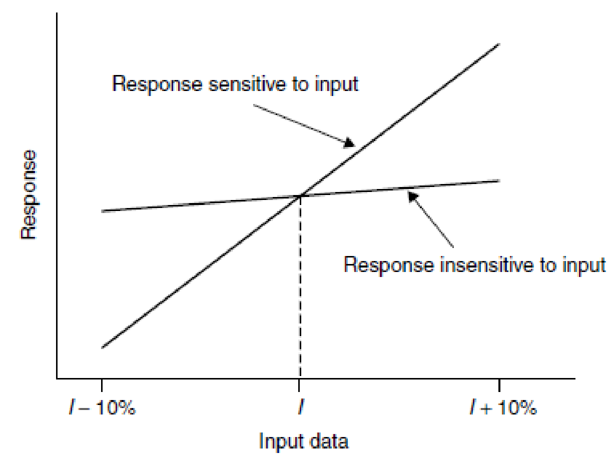
- In theory, the number of replications should be determined separately for every experimental scenario; in practice it is only done for the base scenario!
- If the model has more than one key response the number of replications should be chosen on the basis of the response that requires the most replications

Obtaining Sufficient Output Data

- Best solution:
  - Multiple replications of long runs!
    - The more output data can be obtained, the larger the sample, and the more certainty there can be in the accuracy of the results

Experimentation Preparation

- Sensitivity Analysis
  - Improving the understanding of the model
  - Assessing the consequences of changes in model inputs
  - Model inputs: experimental factors and model data
  - Process:
    - Vary input (I)
    - Run the simulation
    - Measure the effect on the response
  - Result:
    - Significant shift in response?
      - Response sensitive to input change
    - No significant shift in response?
      - Response insensitive to input change



Robinson (2004)

Experimentation Preparation

- Sensitivity Analysis
  - Useful for improving the understanding of the model:
    - Assessing the effect of uncertainties in the data
    - Understanding how changes to the experimental factors affect responses
    - Assessing the robustness of a solution
  - Sensitivity analysis can be time consuming and should be restricted to key inputs
    - Inputs about which there is greatest uncertainty
    - Inputs which are believed to have the greatest impact on the response

Experimentation Preparation

- Experimental design
  - Improving the efficiency of the experimentation process
    - In search experimentation it is often not possible to try all scenarios (i.e. all factor/level combinations)
  - Identifying the experimental factors that are most likely to lead to significant improvements, therefore reducing the total factor/level combinations to be investigated

Experimentation Preparation

- Experimental design
  - How do we identifying the experimental factors that have the greatest impact (i.e. give the greatest improvement towards meeting the objectives of the simulation study)?
    - Data analysis:
      - By analysing data in a model it is sometimes possible to draw conclusions about the likely impact of a change to an experimental factor
    - Expert knowledge
      - Subject matter experts often have a good understanding of the system and the factors that are likely to have a big impact on the responses
    - Preliminary experimentation
      - Interactive simulation can be used to quickly try out different levels based of what is observed during the simulation run
      - Preliminary (simple) sensitivity analysis

Experimentation Preparation

- Experimental design
  - Problem with identifying important experimental factors
    - When factors are changed in isolation they may have a very different effect compared to when they are changed in combination
    - Factors might have a significant effect on simulation output but changes in the real system could have limited practical significance

Importance requires both, statistical and practical significance!

Experimentation Preparation

- Developing an understanding of the solution space
  - Simulating a limited number of scenarios often allows you already to form an opinion as to the likely outcomes of other scenarios
  - How does it work? What methods can you apply?
    - Thinking about the likelihood of a scenario to give the desired output and only simulating the ones likely to lead to success
    - Using linear interpolation (assumes that the solution space is linear)
    - Using unconstrained models (e.g. removing queue limits) – in this way maximum requirements can be established [this is very useful!]
    - Perform some experiments with factor levels that are far apart

Experimentation Preparation

- Experiment design: 2^k factorial designs
  - k = number of experimental factors
  - Each factor is set to two levels (+ and -)
  - Example:
    - k = 3 → $2^3 = 8$ scenarios are simulated and the responses recorded
    - This allows to calculate the mean average effect on the response of changing a factor from its – to its + level → **Main Effect**
    - Main effect

Scenario	Factor 1	Factor 2	Factor 3	Response
1	−	−	−	R_1
2	+	−	−	R_2
3	−	+	−	R_3
4	+	+	−	R_4
5	−	−	+	R_5
6	+	−	+	R_6
7	−	+	+	R_7
8	+	+	+	R_8

$$e_1 = \frac{(R_2 - R_1) + (R_4 - R_3) + (R_6 - R_5) + (R_8 - R_7)}{4}$$

$$e_1 = \frac{-R_1 + R_2 - R_3 + R_4 - R_5 + R_6 - R_7 + R_8}{4}$$

Experimentation Preparation

- Experiment design - 2^k factorial designs
  - Example (cont.)
    - Main effect
      - If main effect of factor is positive, changing the factor from – to + level increases the response by the value of the main effect (at average)
        - » Indicates the direction of change required to achieve a certain effect
        - » Identifies the most important factors
    - Interaction effect
      - If the interaction effect between two factors is positive then setting factor 1 and 2 at the same level (either – or +) will increase the response

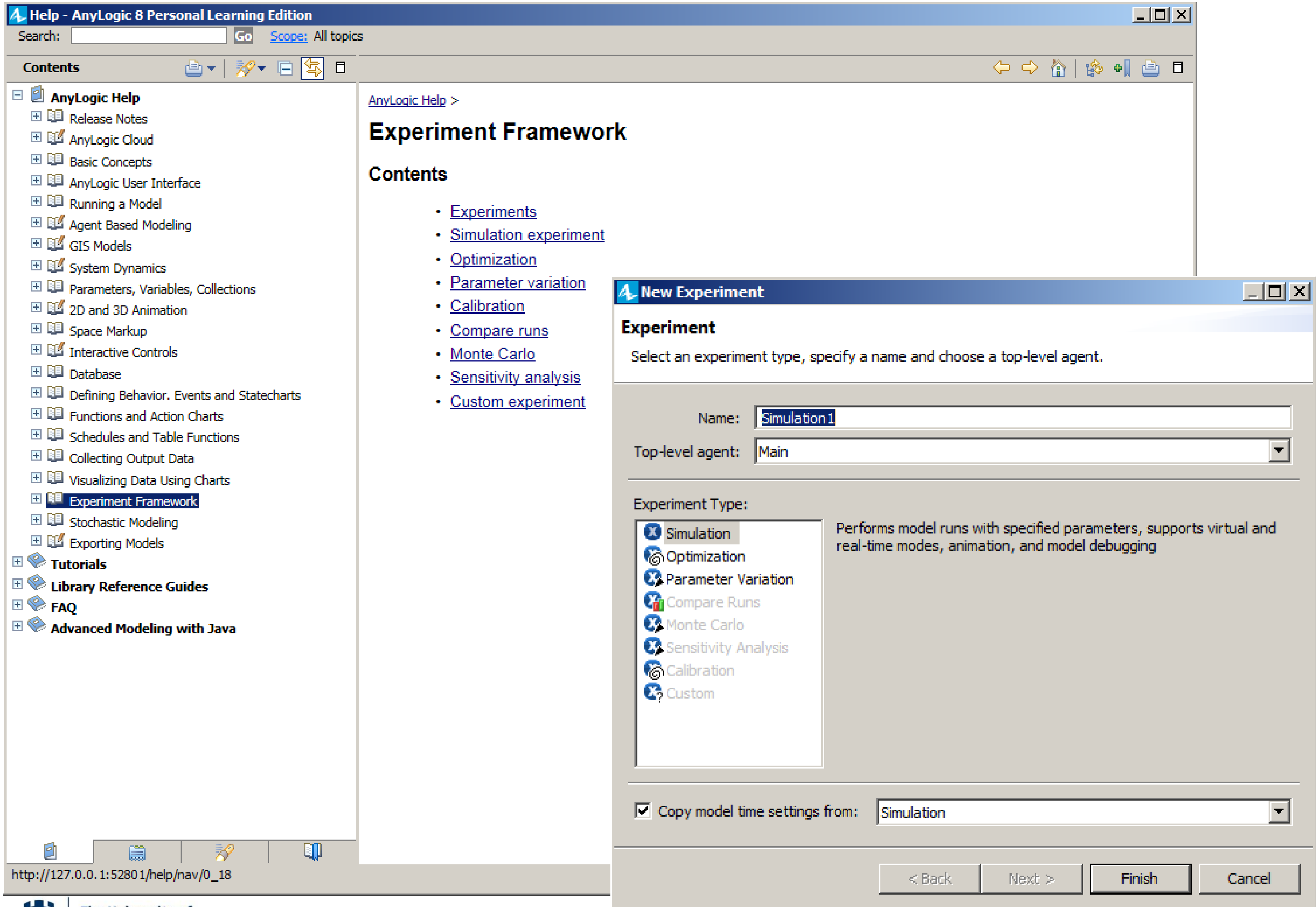
$$e_{12} = \frac{1}{2}\left(\frac{(R_4 - R_3) + (R_8 - R_7)}{2} - \frac{(R_2 - R_1) + (R_6 - R_5)}{2}\right)$$

$$e_{12} = \frac{R_1 - R_2 - R_3 + R_4 + R_5 - R_6 - R_7 + R_8}{4}$$

Experimentation Preparation

- Other approaches to Experimental Design
  - Fractional factorial design
    - When you have too many factors
  - ANOVA
    - Involves a series of hypothesis tests in which it is determined whether changes to the experimental factors have an effect on the response
  - SimHeuristics
    - We come back to this later ...

Experimentation


Help - AnyLogic 8 Personal Learning Edition
Search:
Go
Scope: All topics
Contents
AnyLogic Help
Release Notes
AnyLogic Cloud
Basic Concepts
AnyLogic User Interface
Running a Model
Agent Based Modeling
GIS Models
System Dynamics
Parameters, Variables, Collections
2D and 3D Animation
Space Markup
Interactive Controls
Database
Defining Behavior. Events and Statecharts
Functions and Action Charts
Schedules and Table Functions
Collecting Output Data
Visualizing Data Using Charts
Experiment Framework
Stochastic Modeling
Exporting Models
Tutorials
Library Reference Guides
FAQ
Advanced Modeling with Java
http://127.0.0.1:52801/help/nav/0_18
AnyLogic Help >
Experiment Framework
Contents
Experiments
Simulation experiment
Optimization
Parameter variation
Calibration
Compare runs
Monte Carlo
Sensitivity analysis
Custom experiment
New Experiment
Experiment
Select an experiment type, specify a name and choose a top-level agent.
Name: Simulation1
Top-level agent: Main
Experiment Type:
Simulation
Optimization
Parameter Variation
Compare Runs
Monte Carlo
Sensitivity Analysis
Calibration
Custom
Performs model runs with specified parameters, supports virtual and real-time modes, animation, and model debugging
Copy model time settings from: Simulation
< Back
Next >
Finish
Cancel

Important to Know

- What is the difference between Simulation, Parameter Variation, and Optimisation Experiments?
  - Simulation
    - Performs model runs with specified parameters
  - Parameter Variation
    - Performs multiple model runs varying non or one or more parameters
  - Optimisation
    - Searches for a parameter set corresponding to the best value of the provided objective function; a number of constraints on parameters or model variables can be specified

Optimisation Experiment

https://www.youtube.com/watch?v=3pM30wQfxV8

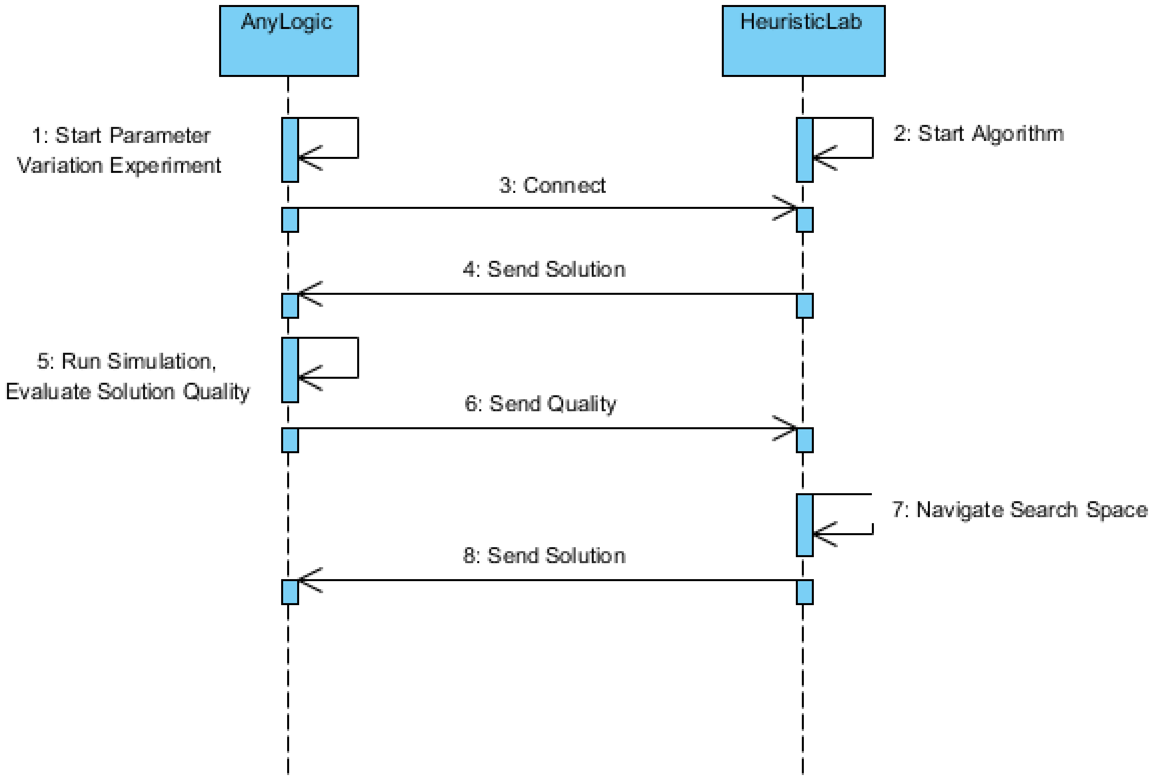
Perhaps also useful: Using Agents in a Process Flow

https://www.youtube.com/watch?v=88D95QYtduM

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

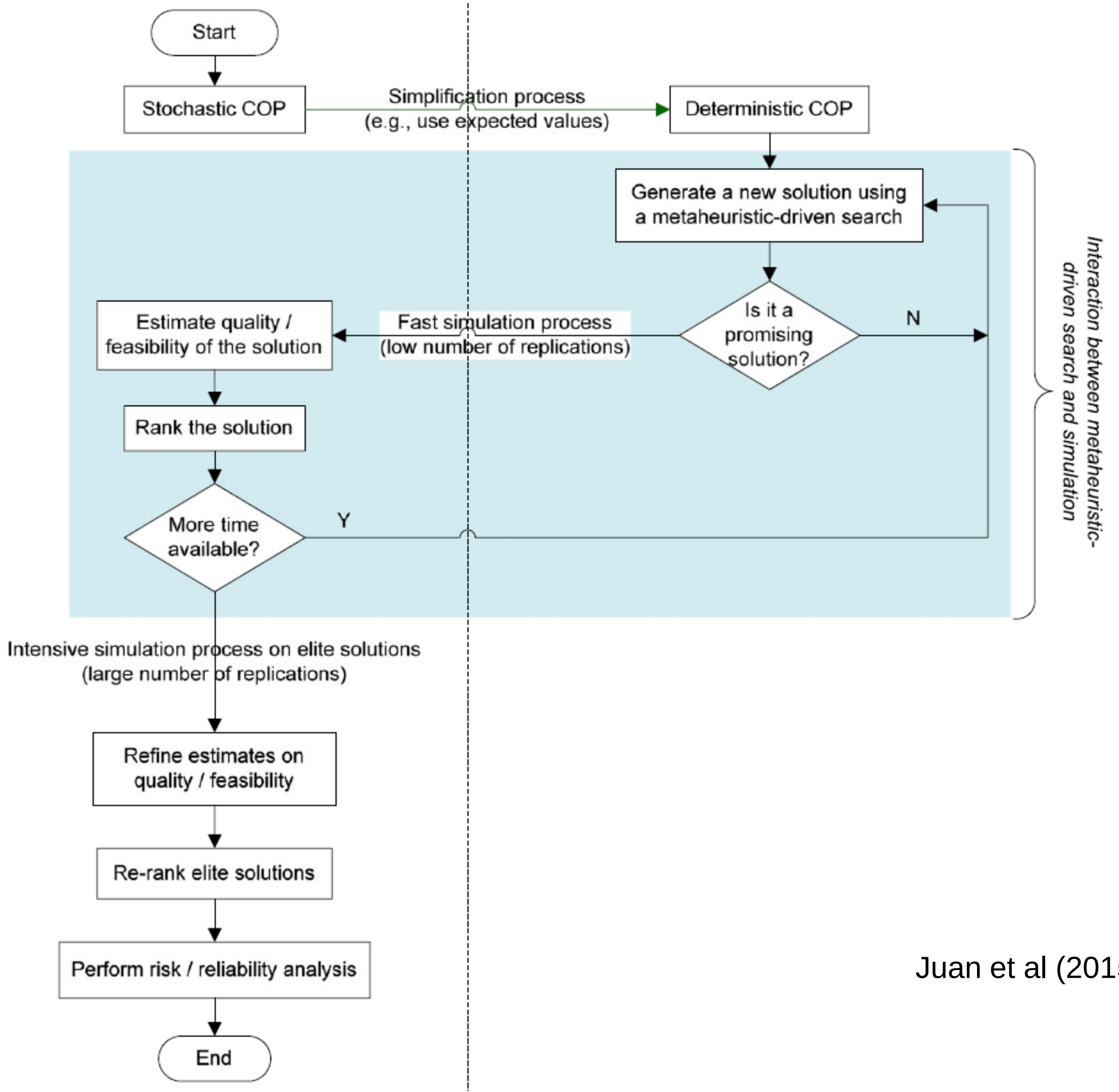
Important to Know

- Optimisation with AnyLogic and HeuristicLab



SimHeuristics

Combinatorial Optimisation Considering Uncertainty



Juan et al (2015)

Questions / Comments

References

- Juan et al (2015) A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems
- Robinson (2002). A statistical process control approach for estimating the warm-up period. In: Proceeding of the 2002 Winter Simulation Conference, San Diego, CA
- Robinson (2004). Simulation: The practice of model development and use. John Wiley & Sons: Chichester, UK