Building Valid, Credible, and Appropriately Detailed Simulation Models

Chapter 5

5.1 Introduction and Definitions

- We would like to determine whether:
  - A simulation model is an accurate representation of the actual system
  - There is sufficient detail in the model
- Verification
  - Determining whether the assumptions document has been correctly translated into a computer program

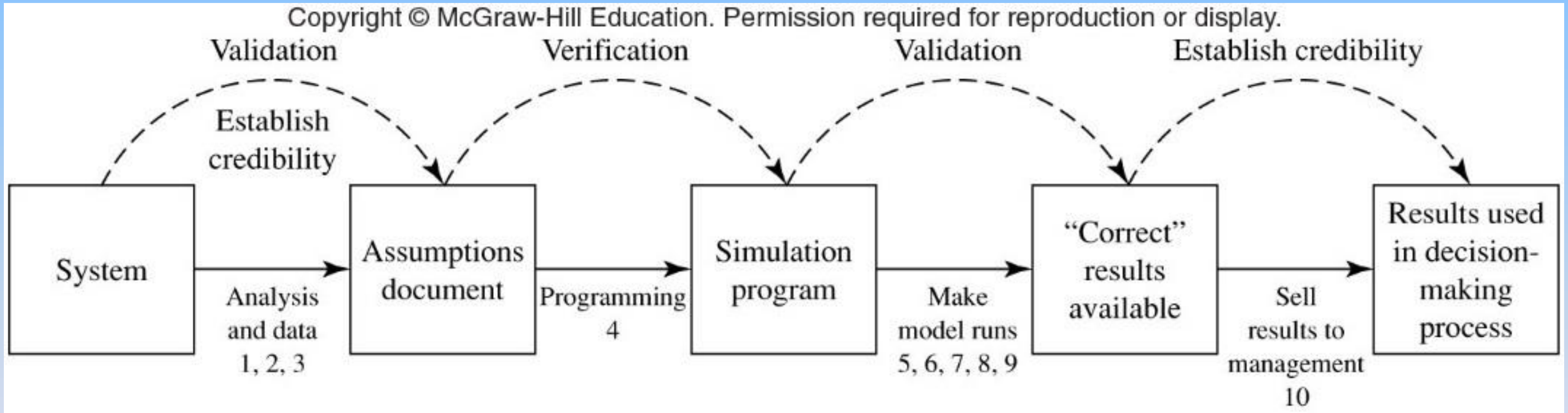
Validation

- Process of determining whether a simulation model is a valid representation of a system
  - For particular study objectives
- Difficulty depends on the system complexity and existence
- Can only be approximate
  - Most valid may not be most cost effective
- Valid for one purpose/application may not be valid for another
- Performance measurement should be the same as that for system
- Should be done during model development

Credibility and Accreditation

- Credibility
  - The extent to which a model is accepted by the manager as "correct"
    - Depends on manager's understanding and knowledge, developer reputation, demonstration
- Accreditation
  - Official certification that a simulation model is acceptable for a particular purpose
    - Has to do with quality of data and doc, model dev. and use history, known limitations
- Both requires validation and verification

Timing and relationships of validation, verification, and Establishing Credibility



Validation vs. Output Analysis

- Validation
  - Make the model as close to the system as possible
- Output analysis
  - Simulation run length, warmup period length, number of independent model runs etc.
  - Make the the estimate of statistics based on output as close to the model as possible
- Example: mean
  - Mean of the system, mean of the model, estimate of the mean

5.2 Guidelines for Determining the Level of Model Detail

- Define specific issues to be investigated by the study
  - Do not include more detail than necessary to address the issues of interest
- Define performance measures that will be used for evaluation
- Use subject-matter experts to help determine the level of model detail
- Consider time and money constraints

Guidelines for Determining the Level of Model Detail

- Entity moving through the simulation model does not have to be the same as entity moving through the system
- Level of detail should be consistent with the type of data available
- If number of factors is large, use a coarse simulation model or an analytic model to learn about critical factors and use them for final modeling

5.3 Verification of Simulation Computer Programs

- Verify by debugging
- Eight techniques for debugging simulation model computer program
  - Some are general

Techniques for debugging simulation model computer program

- Write code in modules or subprograms
  - Main program and few key subprograms should be written and debugged first
- Have more than one person review the program
  - Structured walk-through of program’s code

Techniques for debugging simulation model computer program

- Run the simulation under a variety of settings of the input parameters
  - Check to see that output is reasonable
  - Compute for some simple cases and compare
- Run the model under simplified assumptions for which true characteristics are known
  - Compute analytically and compare

Average number in queue		Average utilization		Average delay in queue	
T	S	T	S	T	S
0.676	0.685	0.600	0.604	0.563	0.565

Techniques for debugging simulation model computer program (cont'd)

- Trace state variables, statistical counters, etc.
  - Or use an interactive debugger
    - Can change variable values to "force" the occurrence of certain types of errors
  - Special input may be needed to trigger certain events

Trace Example (With Error)

Event	Clock	Server status	Number in queue	Times of arrival	Event list: Arrive	Event list: Depart
Initialization	0	0	0		0.4	∞
Arrival	0.4	1	0		1.6	2.4
Arrival	1.6	1	1	1.5	2.1	2.4
Arrival	2.1	1	2	1.5, 2.1	3.8	2.4
Departure	2.4	1	1	2.1	3.8	3.1
Departure	3.1	1	0		3.8	3.3
Departure	3.3	0	0		3.8	∞

Event	Number of customers delayed	Total delay	Area under number-in-queue function	Area under busy function
Initialization	0	0	0	0
Arrival	1	0	0	0
Arrival	1	0	0	1.1
Arrival	1	0	0.5	1.7
Departure	2	0.8	1.1	2.0
Departure	3	1.8	1.8	2.7
Departure	3	1.8	1.8	2.9

Techniques for debugging simulation model computer program (cont'd)

- Observe an animation of the simulation output
  - E.g. A simulation model of a network of automobile traffic intersections, cars colliding
- Compute sample mean and variance values and compare them with desired mean and variance
- Use a commercial simulation package
  - Most recent release might have bugs

5.4 Techniques for Increasing Model Validity and Credibility

- Make use of all available information and data
- Interacting with the manager on a regular basis
- Maintain a written assumptions document
- Use quantitative techniques to validate model components
- Validate output from overall simulation model
- Animation

Use All Available Info. & Data

- Conversations with subject-matter experts
  - Ex: Who to talk with for a cellular network?
- Observations of the system (history or realtime)
  - Modelers need to understand the process producing the data
  - Potential difficulties with data
    - Not representative, e.g. one combat differs from another
    - Inappropriate type or format, e.g. UTC instead of EST
    - Measurement, recording or rounding errors
    - Biased, e.g. because of self-interest
    - Inconsistent units

Use All Available Info. & Data (cont'd)

- Existing theory
  - E.g. Customer arrival rate is constant => inter-arrival time probably IID, arrival process is Poisson
- Relevant results from similar simulation studies
- Modelers' experience and intuition

Interacting w/ the Manager on a Regular Basis

- Do so throughout the course of the simulation study.
  - Maintain manager's interest and involvement
  - Allow manager to contribute knowledge, and to reformulate the objectives (as things may get clearer)
  - More likely for manager to "sign off" since s/he understand more

Maintain a Written Assumptions Document

- The doc is about how the system should be modeled
  - Not about how the system works
  - What to write depends on insight, knowledge and experience, usually includes
    - Overview of goals, addressed issues, inputs, performance measurement
    - Description of subsystems and their interaction
    - Simplifying assumptions
    - Limitations
    - Summaries of data sets
    - Information sources
    - Etc.
  - Should have enough detail to be a program "blue-print"

Maintain a Written Assumptions Document (cont'd)

- Perform a structured walk-through of this document with both SMEs and managers
  - Correct missing or invalid assumptions
- The doc can
  - Reduce communication errors
  - Enhance credibility

Use Quantitative Techniques to Validate Model Components

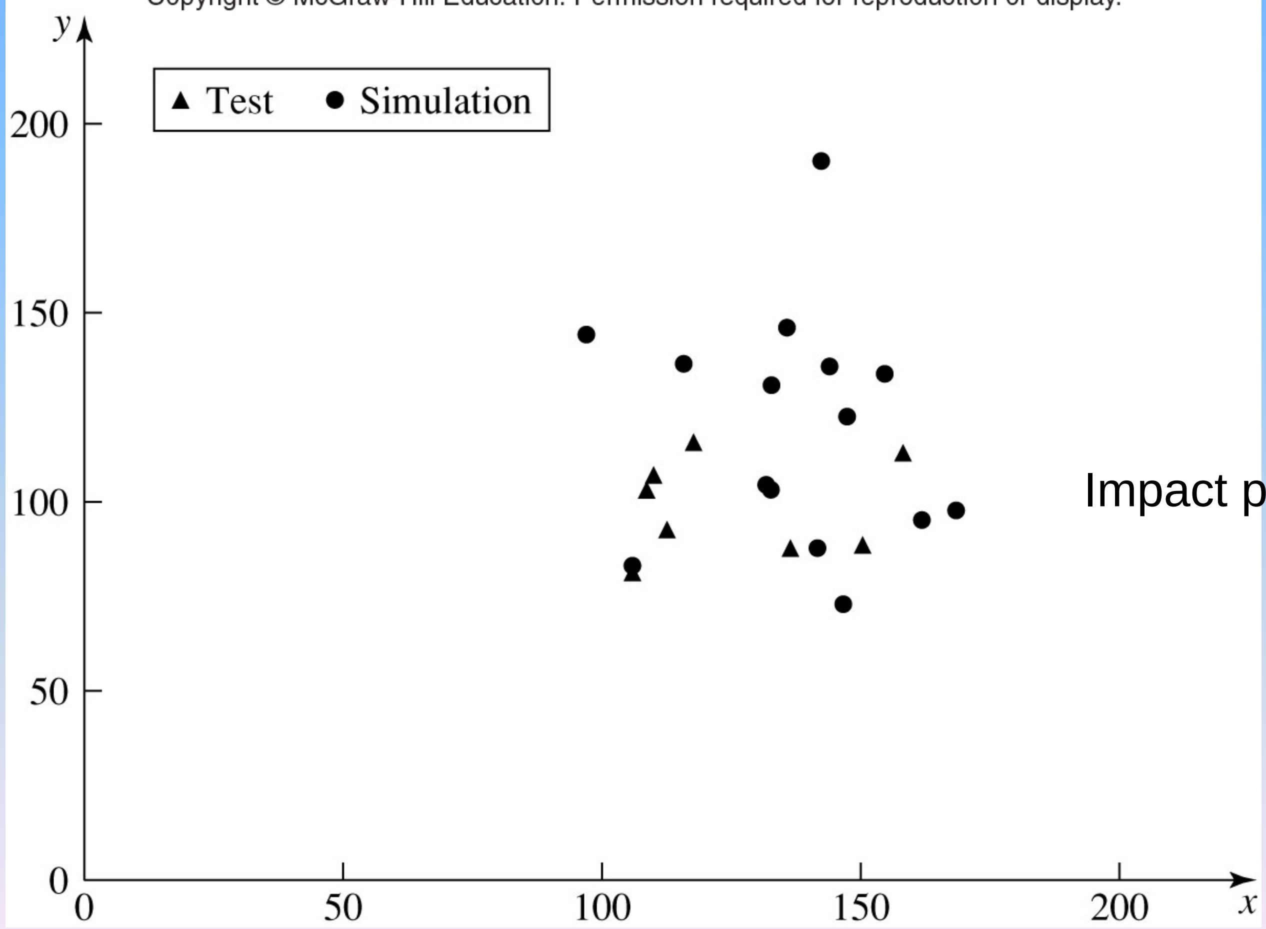
- Fit probability distribution to observed data
  - Use graphical plots and goodness-of-fit tests (Ch. 6)
- Use Kruskal-Wallis test to check the homogeneity of real data for modelling (Ch. 6)
- Sensitivity analysis: determine which factors significantly impact performance
  - Parameter values, distribution, entity, detail level, etc.
  - Factor changes result in little changed results -> non-sensitive
  - Avoid inadvertent impact (e.g. random values different between simulation runs)

Validate Output from Overall Simulation Model

- Compare with an existing system
  - Compare data from the existing system with simulation output using numerical statistics (e.g. mean, variance, correlation) and/or graphical plots (e.g. histograms, box plots, spider-web plots)

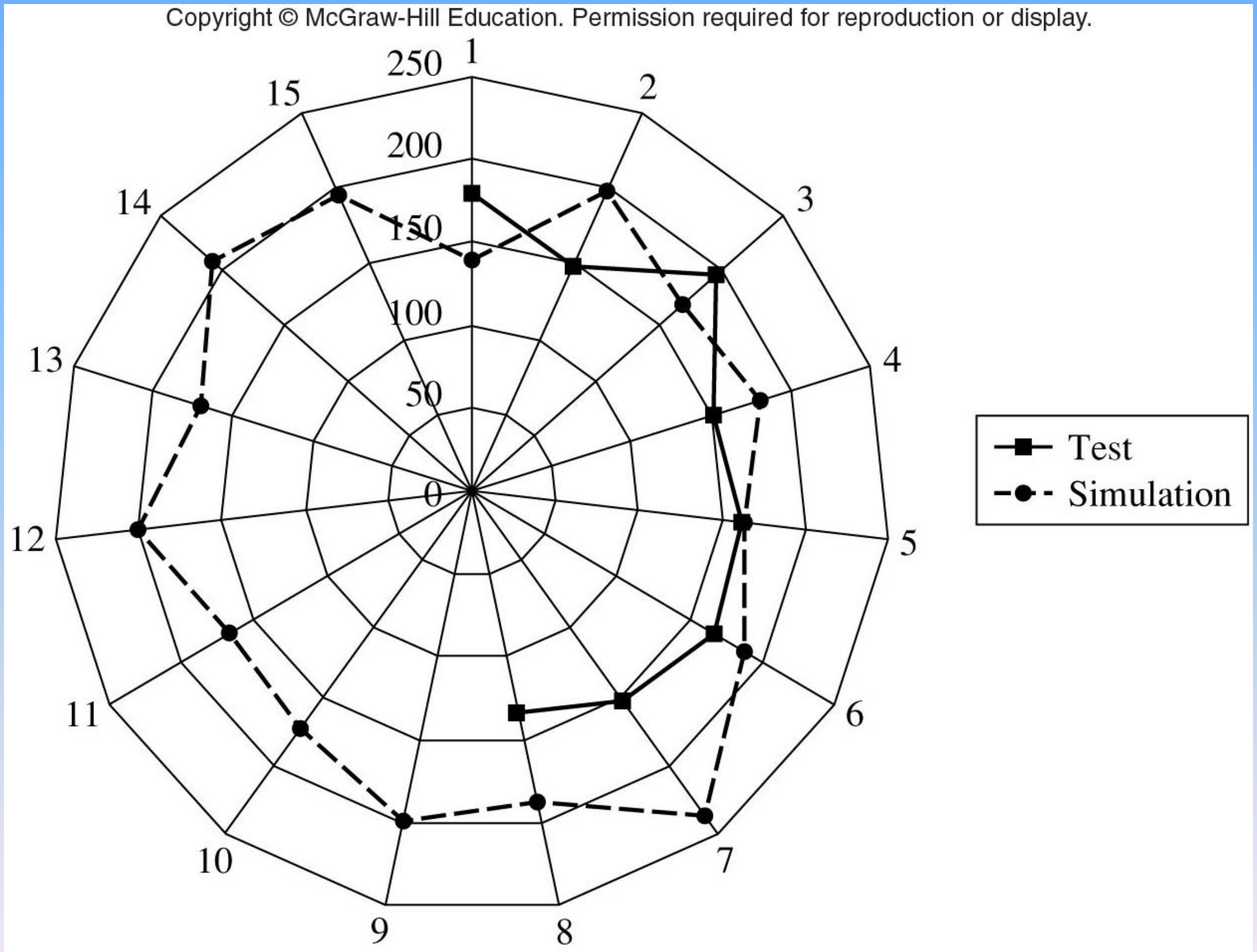
Output Validation Example (1)

- Missile simulation

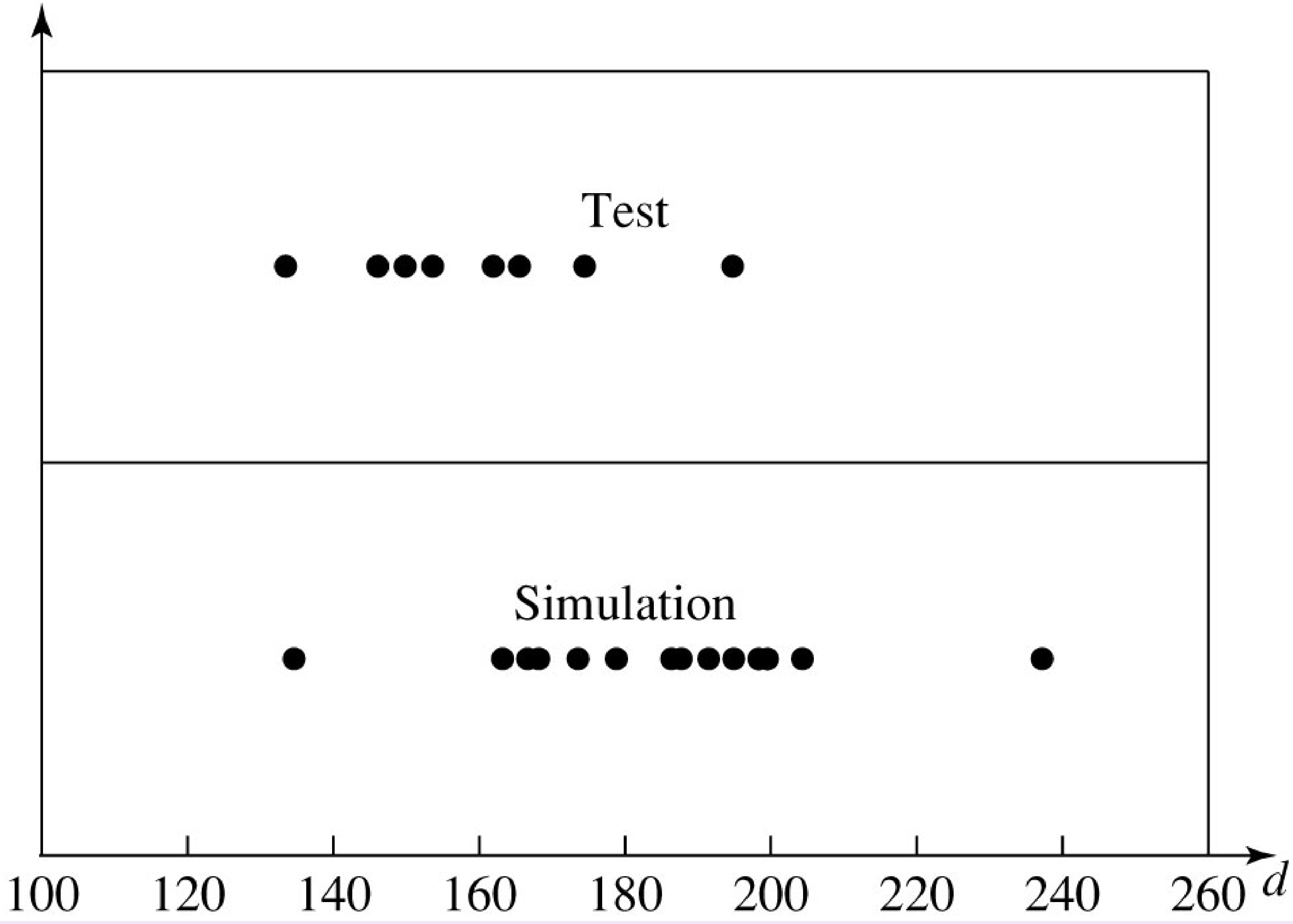


Impact points

Output Validation Example (2)



Output Validation Example (3)



Output Validation Example (4)

Missile number	Test miss distance	Simulation miss distance
1	174.45	134.60
2	146.09	194.73
3	194.72	168.14
4	149.84	178.82
5	161.93	163.78
6	165.52	186.39
7	153.62	237.20
8	133.46	187.73
9	—	197.90
10	—	173.55
11	—	166.64
12	—	199.10
13	—	168.17
14	—	204.32
15	—	191.48
Sample mean	159.95	183.50
Sample variance	355.75	545.71

Validate Output from Overall Simulation Model (cont'd)

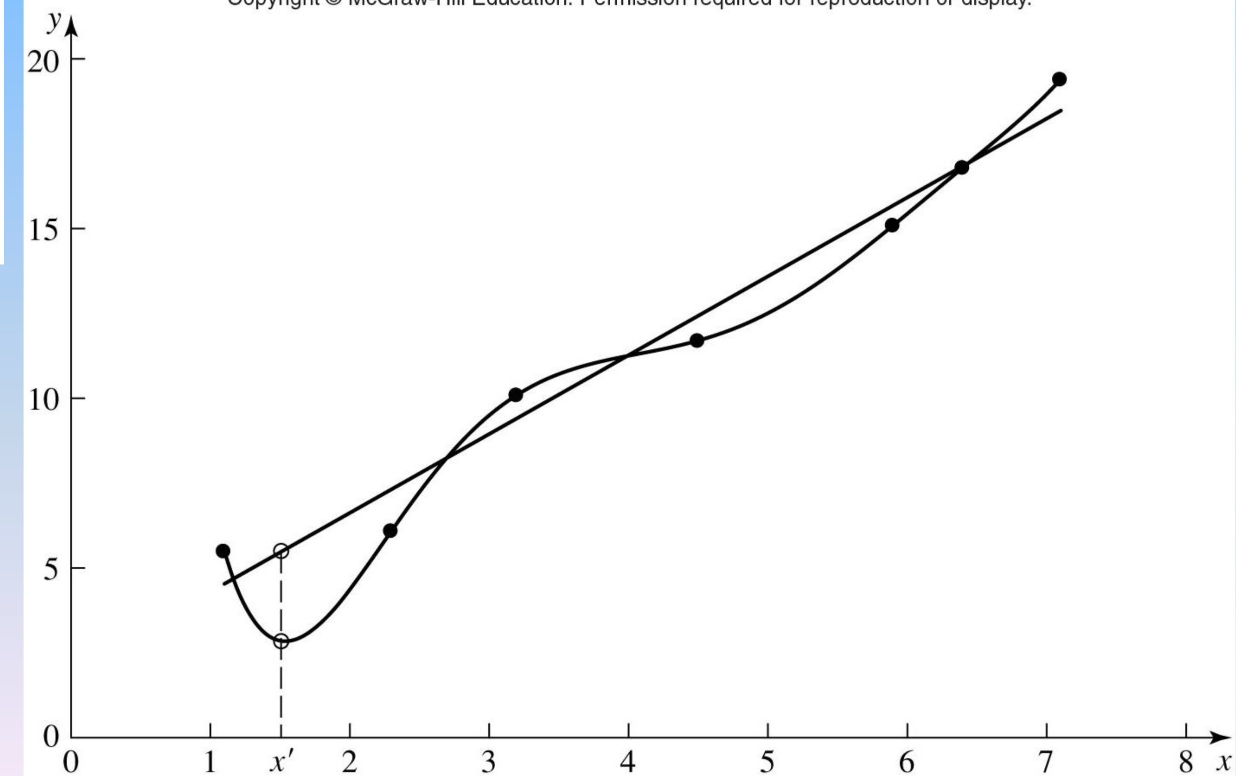
- Turing test
  - Ask SMEs to differentiate system data and simulation data without prior knowledge
  - Data may be of various forms, e.g. animation
- Prospective validation (check Model’s ability to predict future system behavior)
  - Data collected at a later time can be compared with predictions to validate the model
  - Use discrepancies to improve the model if suggested objectively by the results

Techniques for Increasing Model Validity and Credibility

- Given discrepancies, what to do?
- Calibration of the model
  - Parameters are tweaked until two data sets agree closely
  - Use another data set pairs to validate tweaking
    - The tweaked model might only work well on (or be representative of) the first data set

Calibration Example

x	y
1.1	5.5
2.3	6.1
3.2	10.1
4.5	11.7
5.9	15.1
6.4	16.8
7.1	19.4



Validate Output from Overall Simulation Model (cont'd)

- Compare results with expert opinion
  - If consistent with perceived system behavior, model is said to have face validity
- Compare with another model
  - Another model developed for the same system for a similar purpose and have been validated

Use Animation

- Provide visual and intuitive perception

5.5 Management's Role in the Simulation Process

- Responsibilities of the manager
  - Formulate problem objectives
  - Direct personnel to provide information and data to the simulation modeler
    - And to attend the structured walk-through
  - Interact with the simulation modeler on a regular basis
  - Use the simulation results as an aid in the decision-making process

5.6 Statistical Procedures for Comparing Real-World Observations and Simulation Output Data

- Classical statistical tests are not applicable
  - Output processes are nonstationary and autocorrelated
- Inspection approach
- Confidence-interval approach
- Time-Series approach

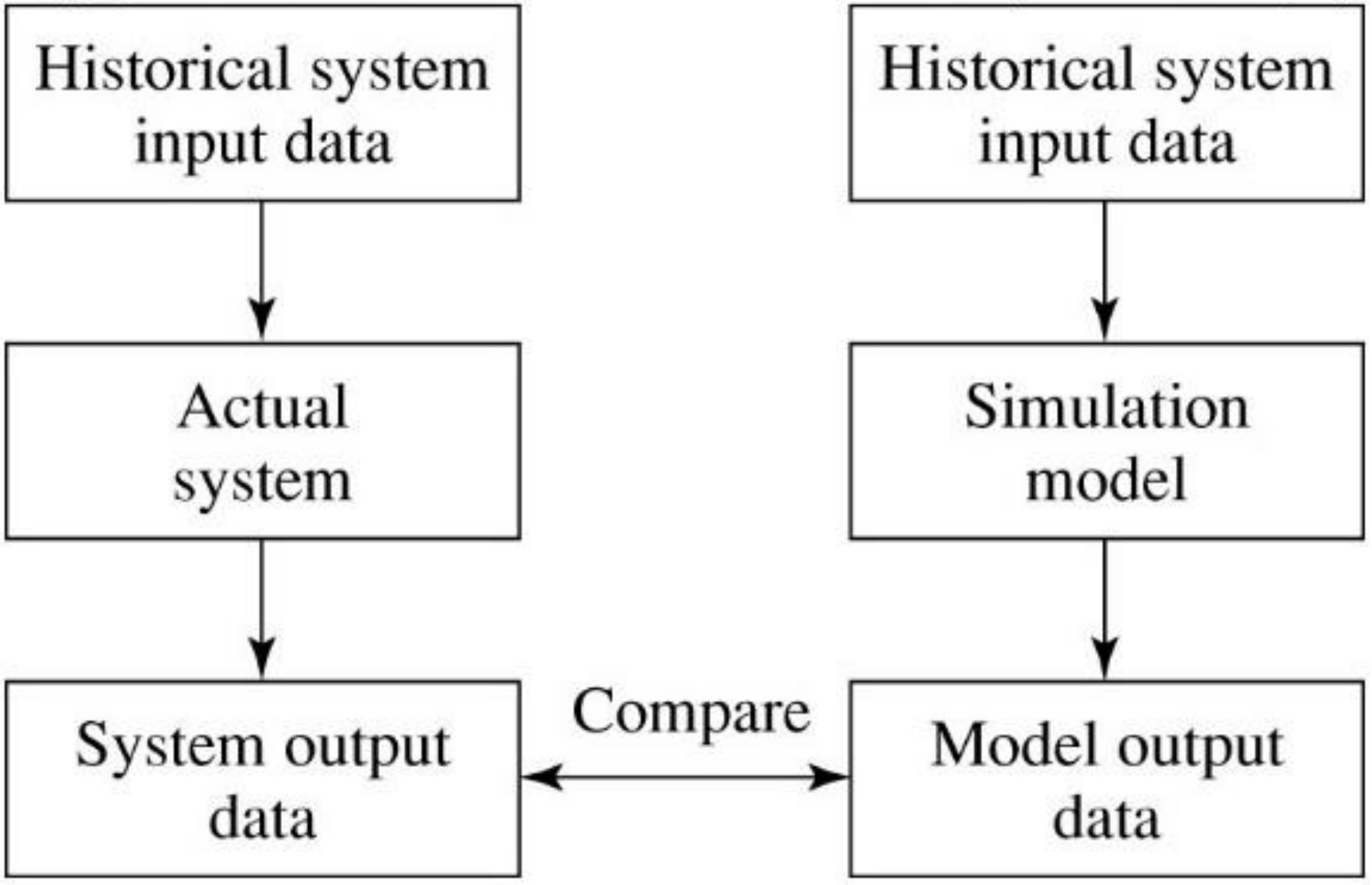
Inspection Approach

- Basic inspection approach
  - Compare two sets of statistics from real-world output and simulation output
  - Simulation uses input from randomly generated numbers based on estimated distribution

Inspection Approach (cont'd)

- Correlated inspection approach
  - “Drive” the model with historical system input data
    - Also called a trace-driven simulation
  - Then, compare the model and system outputs
  - The more definitive method to validate the assumptions of the simulation model

Correlated Inspection Approach



Correlated Inspection Approach Ex.

- Five-teller bank w/ or w/o jockeying

Experiment j	X_j	Y_j	Y'_j	$X_j - Y_j$	$X_j - Y'_j$
1	3.06	3.81	2.62	−0.75	0.44
2	2.79	3.37	2.05	−0.58	0.74
3	2.21	2.61	4.56	−0.40	−2.35
4	2.54	3.59	1.86	−1.05	0.68
5	9.27	11.02	2.41	−1.75	6.86
6	3.09	3.75	1.85	−0.66	1.24
7	2.50	2.84	1.13	−0.34	1.37
8	0.31	0.71	3.12	−0.40	−2.81
9	3.17	3.94	5.09	−0.77	−1.92
10	0.98	1.18	1.25	−0.20	−0.27
Sample mean of all 500	2.10	2.85	2.70	−0.75	−0.60
Sample variance of all 500	2.02	2.28	2.12	0.08	4.08

- Reason: Var(A-B) = Var(A) + Var(B) – 2Cov(A,B)

Confidence-Interval Approach Based on Independent Data

- Compare model to system by constructing a confidence interval for $\mu_X - \mu_Y$
- More reliable approach, when it is possible to collect large data sets from both system and model
  - H_0: $\mu_X = \mu_Y$ is false for most cases as models are approx.
  - C.I. indicates the magnitude of $\mu_X - \mu_Y$
- Use paired-*t* approach or Welch approach (Sec. 10.2)
  - Paired-t: Same data set size for real and model. May use the correlated inspection approach.
  - Welch: Can have diff. data set size (>= 2). Real and model data must be independent.

Confidence-Interval Approach Based on Independent Data (cont'd)

- If 0 is not in the C.I., $\mu_X - \mu_Y$ is statistically significant at level α
  - Equivalent to rejecting H_0: $\mu_X = \mu_Y$
  - The model may still in practically valid
  - When is $\mu_X - \mu_Y$ practically significant?
    - Subjective decision, depending on model purposes and utility function (that measures preferences)

Confidence-Interval Approach Ex. 1

- Five-teller bank w/ or w/o jockeying
  - $\bar{W}(10) = \bar{X}(10) - \bar{Y}(10) = 2.99 - 3.68 = -0.69$
  - $\widehat{Var}\left(\bar{W}(10)\right) = \frac{\sum_{j=1}^{10}\left[W_j - \bar{W}(10)\right]^2}{(10)(9)} = 0.02$
  - 90% c.i. of $\mu_X - \mu_Y$ (paired-t approach)

$$\bar{W}(10) \pm t_{9,0.95}\sqrt{\widehat{Var}\left(\bar{W}(10)\right)} = -0.69 \pm 0.26$$

  - $\mu_X - \mu_Y$ statistically significant
    - Practical significant?

Confidence-Interval Approach Ex. 2

- Missile simulation
- X_i, Y_i: Missile distance from target
- $\bar{X}(8) = 159.95, S_X^2(8) = 355.75$
- $\bar{Y}(15) = 183.5, S_Y^2(15) = 545.71$
- $\hat{f} = 17.34$ (estimated degree of freedom)
- 95% c.i. of $\mu_X - \mu_Y$ (Welch approach)
- $\bar{X}(8) - \bar{Y}(15) \pm t_{\hat{f},0.975}\sqrt{\frac{S_X^2(8)}{8} + \frac{S_Y^2(15)}{15}} = -23.55 \pm 18.97$

Time-series Approaches

- Require only one set of each type of output data
- Spectral-analysis approach
  - Compute the sample spectrum of each output process
  - Construct a confidence interval for the difference of the logarithms of the two spectra
  - Output processes must be covariance-stationary
  - Mathematical sophisticated

Other Approaches

- Hypothesis test for a trace-driven simulation model
  - Kleijnen, Bettonvil, and Van Groenendaal (1998)
- Distribution-free hypothesis test based on bootstrapping
  - Kleijnen, Cheng, and Bettonvil (2000, 2001)
- Both tests on statistics of the system and the model