Chapter 10
Verification and Validation of Simulation Models

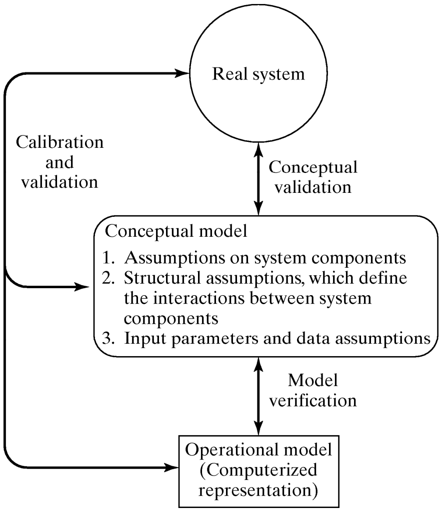
Banks, Carson, Nelson & Nicol

Discrete-Event System Simulation

Purpose & Overview

- The goal of the validation process is:
  - To produce a model that represents true behavior closely enough for decision-making purposes
  - To increase the model's credibility to an acceptable level
- Validation is an integral part of model development
  - Verification – building the model correctly (correctly implemented with the software)
  - Validation – building the correct model (an accurate representation of the real system)

Modeling-Building, Verification & Validation



Real system

Calibration and validation

Conceptual validation

Conceptual model

1. Assumptions on system components
2. Structural assumptions, which define the interactions between system components
3. Input parameters and data assumptions

Model verification

Operational model (Computerized representation)

Verification - Debugging

- Purpose: ensure the conceptual model is reflected accurately in the computerized representation.
- Many common-sense suggestions, for example:
  - Have someone else check the model.
  - Make a flow diagram that includes each logically possible action a system can take when an event occurs.
  - Closely examine the model output for reasonableness under a variety of input parameter settings. (Often overlooked!)
  - Print the input parameters at the end of the simulation, make sure they have not been changed inadvertently.

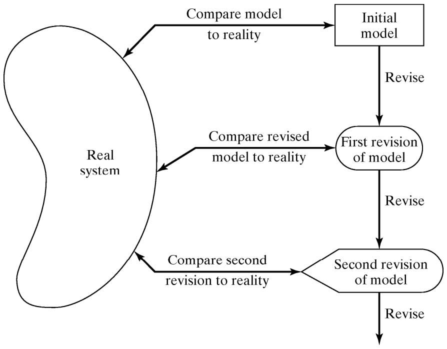
Other Important Tools

[Verification]

- Documentation
  - A means of clarifying the logic of a model and verifying its completeness
- Use of a trace
  - A detailed printout of the state of the simulation model over time.
- Animation

Calibration and Validation

- Validation: the overall process of comparing the model and its behavior to the real system.
- Calibration: the iterative process of comparing the model to the real system and making adjustments.



Calibration and Validation

- No model is ever a perfect representation of the system
  - ☐ The modeler must weigh the possible, but not guaranteed, increase in model accuracy versus the cost of increased validation effort.
- Three-step approach:
  - ☐ Build a model that has high face validity.
  - ☐ Validate model assumptions.
  - ☐ Compare the model input-output transformations with the real system's data.

High Face Validity

[Calibration & Validation]

- The model should appear reasonable to model users and others who are knowledgeable about the system.
  - ☐ Especially important when it is impossible to collect data from the system
- Ensure a high degree of realism: Potential users should be involved in model construction (from its conceptualization to its implementation).
- Sensitivity analysis can also be used to check a model's face validity.
  - ☐ Example: In most queueing systems, if the arrival rate of customers were to increase, it would be expected that server utilization, queue length and delays would tend to increase.

Validate Model Assumptions

[Calibration & Validation]

- General classes of model assumptions:
  - ☐ Structural assumptions: how the system operates.
  - ☐ Data assumptions: reliability of data and its statistical analysis.
- Bank example: customer queueing and service facility in a bank.
  - ☐ Structural assumptions, e.g., customer waiting in one line versus many lines, served FCFS versus priority.
  - ☐ Input data assumptions, e.g., interarrival time of customers, service times for commercial accounts.
    - Verify data reliability with bank managers.
    - Test correlation and goodness of fit for data (see Chapter 9 for more details).

Validate Input-Output Transformation

[Calibration & Validation]

- Goal: Validate the model's ability to predict future behavior
  - ☐ The only objective test of the model.
  - ☐ The structure of the model should be accurate enough to make good predictions for the range of input data sets of interest.
- One possible approach: use historical data that have been reserved for validation purposes only.
- Criteria: use the main system responses of interest.

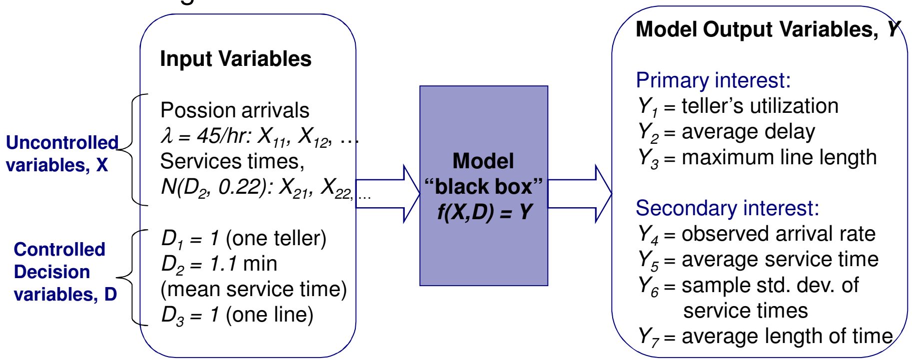
Bank Example [Validate I-O Transformation]

- Example: One drive-in window serviced by one teller, only one or two transactions are allowed.
  - ☐ Data collection: 90 customers during 11 am to 1 pm.
    - Observed service times $\{S_i, i = 1,2, \ldots, 90\}$.
    - Observed interarrival times $\{A_i, i = 1,2, \ldots, 90\}$.
  - ☐ Data analysis let to the conclusion that:
    - Interarrival times: exponentially distributed with rate $\lambda = 45$
    - Service times: $N(1.1, 0.2^2)$

The Black Box

[Bank Example: Validate I-O Transformation]

- A model was developed in close consultation with bank management and employees
- Model assumptions were validated
- Resulting model is now viewed as a "black box":



Input Variables

Uncontrolled variables, X

Possion arrivals
$\lambda = 45/hr: X_{11}, X_{12}, \ldots$
Services times,
$N(D_2, 0.22): X_{21}, X_{22}, \ldots$

Controlled Decision variables, D

$D_1 = 1$ (one teller)
$D_2 = 1.1$ min (mean service time)
$D_3 = 1$ (one line)

Model "black box" $f(X,D) = Y$

Model Output Variables, Y

Primary interest:
Y_1 = teller's utilization
Y_2 = average delay
Y_3 = maximum line length

Secondary interest:
Y_4 = observed arrival rate
Y_5 = average service time
Y_6 = sample std. dev. of service times
Y_7 = average length of time

Comparison with Real System Data

[Bank Example: Validate I-O Transformation]

- Real system data are necessary for validation.
  - Average delays should have been collected during the same time period (from 11am to 1pm on the same Friday.)
- Compare the average delay from the model Y with the actual delay Z:
  - Average delay observed, $Z = 4.3$ minutes, consider this to be the true mean value $\mu_0 = 4.3$.
  - When the model is run with generated random variates X_{1n} and X_{2n}, Y should be close to Z.
  - Six statistically independent replications of the model, each of 2-hour duration, are run.

Hypothesis Testing

[Bank Example: Validate I-O Transformation]

- Compare the average delay from the model Y with the actual delay Z (continued):
  - Null hypothesis testing: evaluate whether the simulation and the real system are *the same* (w.r.t. output measures):

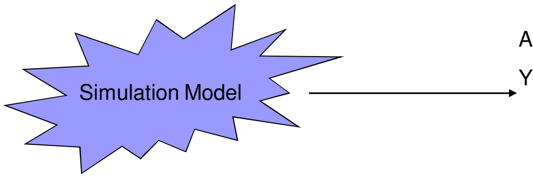
$$H_0 : E(Y) = 4.3 \text{ minutes}$$

$$H_1 : E(Y) \neq 4.3 \text{ minutes}$$

    - If H_0 is not rejected, then, there is no reason to consider the model invalid
    - If H_0 is rejected, the current version of the model is rejected, and the modeler needs to improve the model

Hypothesis Testing

[Bank Example: Validate I-O Transformation]



Average Delay Times

$Y_1, Y_2, \ldots, Y_6$ iid random variables

Replication	Average Delay
1	2.79
2	1.12
3	2.24

Replication	Average Delay
4	3.45
5	3.13
6	2.38

Hypothesis Testing

[Bank Example: Validate I-O Transformation]

- Conduct the *t* test:
  - Choose level of significance ($\alpha = 0.5$) and sample size ($n = 6$).
  - Compute the same mean and sample standard deviation over the n replications:

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = 2.51 \text{ minutes} \qquad S = \left[\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}\right]^{1/2} = 0.81 \text{ minutes}$$

  - Compute test statistics:

Student's t distribution

$$|t_0| = \left|\frac{\bar{Y} - \mu_0}{S/\sqrt{n}}\right| = \left|\frac{2.51 - 4.3}{0.82/\sqrt{6}}\right| = 5.24 \quad > \quad t_{\alpha/2,n-1} = 2.571$$

  - Hence, reject H_0. Conclude that the model is inadequate.
  - Check: the assumptions justifying a *t* test, that the observations (Y_i) are normally and independently distributed.

Summary

- Model validation is essential:
  - Model verification
  - Calibration and validation
  - Conceptual validation
- Best to compare system data to model data, and make comparison using a wide variety of techniques.
- Some techniques that we covered (in increasing cost-to-value ratios):
  - Insure high face validity by consulting knowledgeable persons.
  - Conduct simple statistical tests on assumed distributional forms.
  - Compare model output to system output by statistical tests.