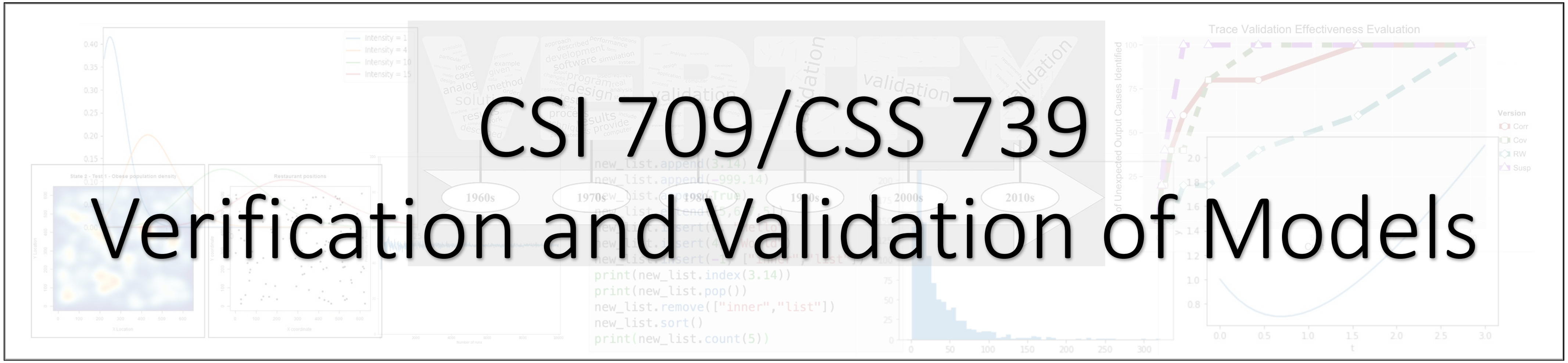


CSI 709/CSS 739
Verification and Validation of Models

Simulation Runtime Verification

Dr. Hamdi Kavak
Computational and Data Sciences Department

http://www.hamdikavak.com
hkavak@gmu.edu

GEORGE MASON UNIVERSITY

Center for Social Complexity

Short Writing Assignment

- Let’s go over the course website and BlackBoard to make we understand which papers we should consider for SWA submissions.

- SWA1 feedback
  - Good first try.
  - Try to write a coherent text rather than discrete sentences.
  - Get closer to 500 words.
  - Don’t hesitate to criticize the paper. **Not** all papers that are assigned to you are great!
  - E.g.: Was it an easy read? Did authors sufficiently substantiate their claims? Was there a methodological contribution or to a specific application area?

Final project proposal

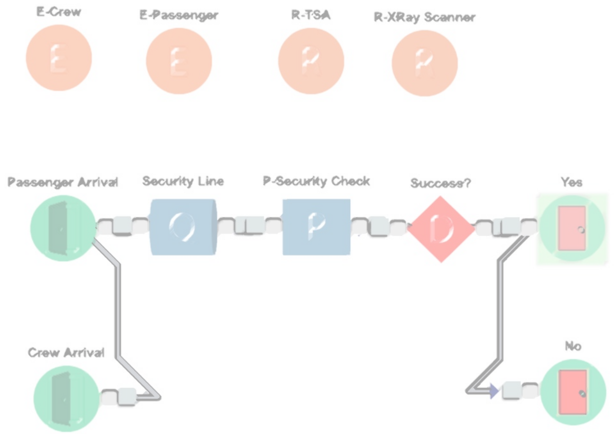
- Submission deadline tonight 11:59 pm

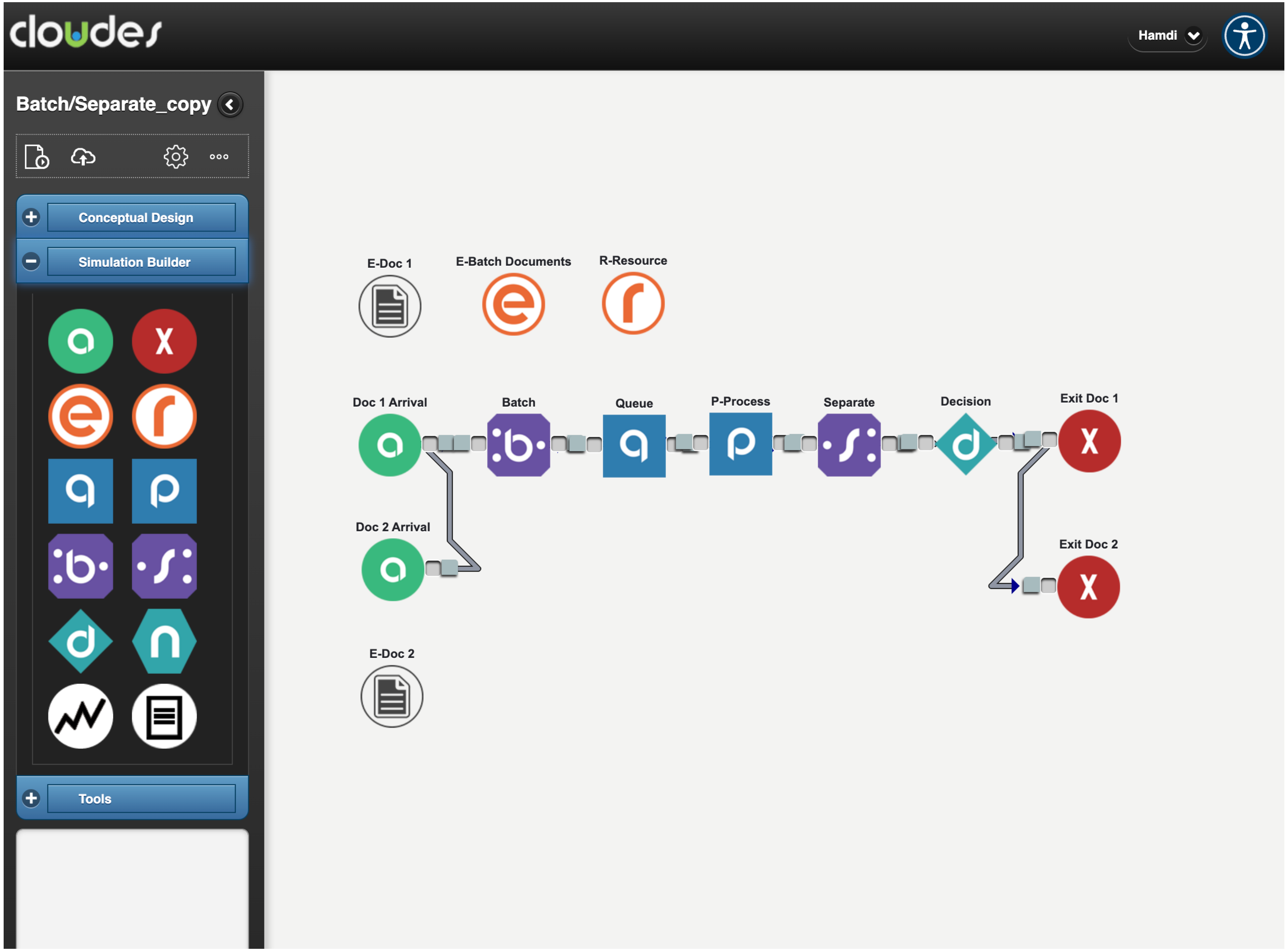
- Any questions?

CLOUDES Simulation Tool

- I will be using the CLOUDES simulation tool to highlight certain techniques.
- Visit http://cloudes.me to register and create models online
- Visit http://blog.cloudes.me to learn more about the tool

cloudes.me




clouder
Hamdi
Batch/Separate_copy
Conceptual Design
Simulation Builder
Tools
E-Doc 1
E-Batch Documents
R-Resource
Doc 1 Arrival
Batch
Queue
P-Process
Separate
Decision
Exit Doc 1
Doc 2 Arrival
Exit Doc 2
E-Doc 2

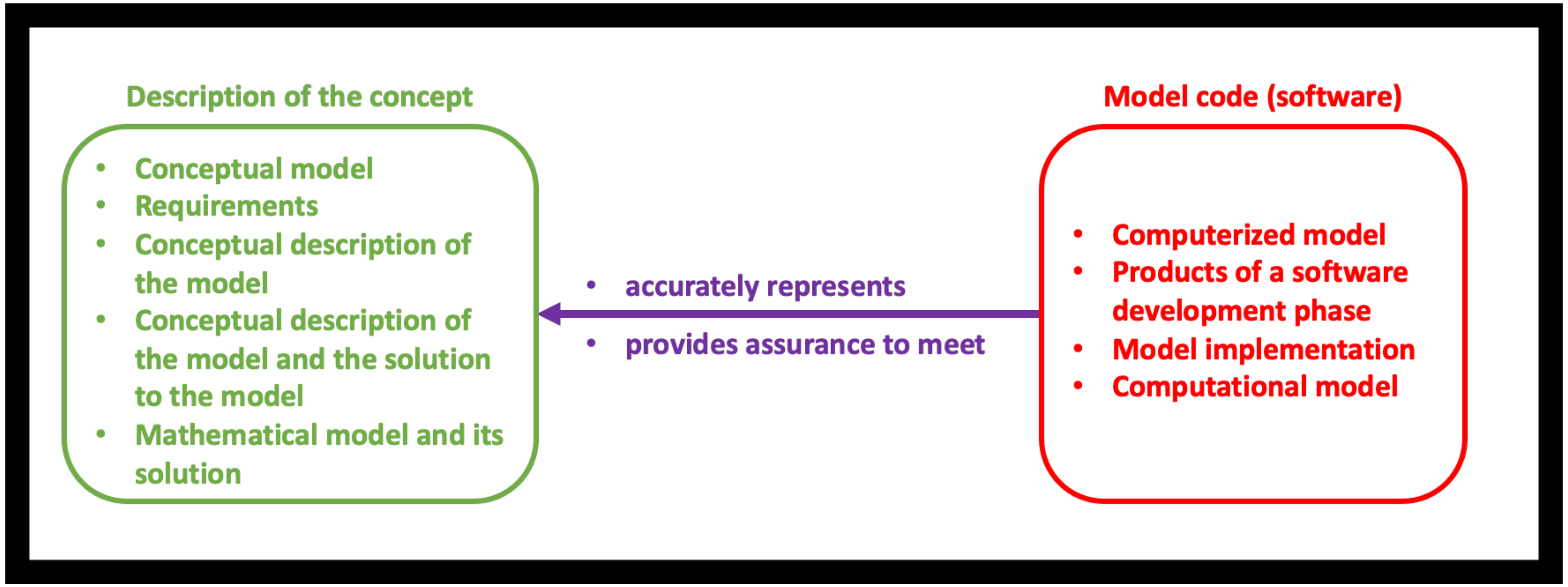
Objective of this lecture

- Go deeper into the simulation verification literature and particularly runtime verification
- Topics include:
  - Terminology
  - Techniques
  - Pros & cons
  - Challenges
  - Open research avenues

Verification

- Concerns with checking how good the model code reflects model specifications (i.e., conceptual model).

recall from week 2



Runtime verification

Definition

- “Structured approaches that rely on human reasoning to evaluate the model during its **execution**”

Goal

- “Utilize **monitors** to reveal behaviors for manual evaluation”

Types of techniques

- “Informal and dynamic V&V”

Source: Lynch (2022)

Runtime verification techniques

- Visualization
- Animation
- Parameter-verification test
- Extreme conditions test
- Object flow testing
- Traces
- Execution tracing
- Message Sequence Charts for Process Interaction Models

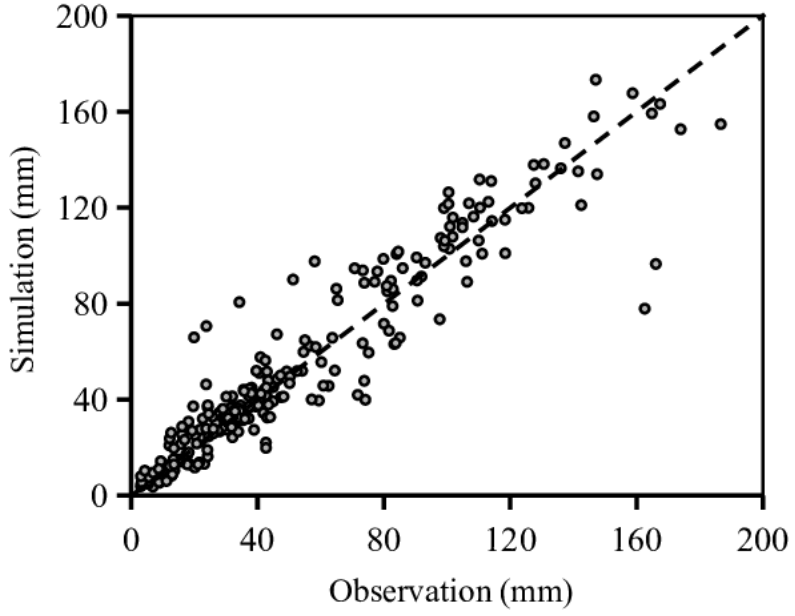
Source: Lynch (2019)

Visualization

- Provides a good starting for investigating model verification.
- Involves collecting and combining data from simulation runs.
- Simulation tools often provide instrumentation for visualization and data collection.
- This lecture will cover highly used visualization techniques

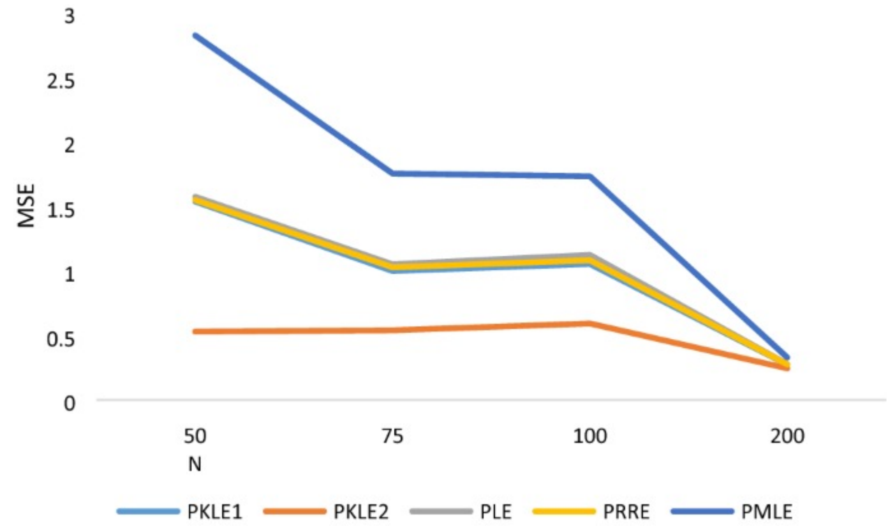
Visualization cont.

Scatter plot



Source: Wang, X. S., & Zhou, Y. (2016). Shift of annual water balance in the Budyko space for catchments with groundwater-dependent evapotranspiration. *Hydrology and Earth System Sciences*, *20*(9), 3673-3690.

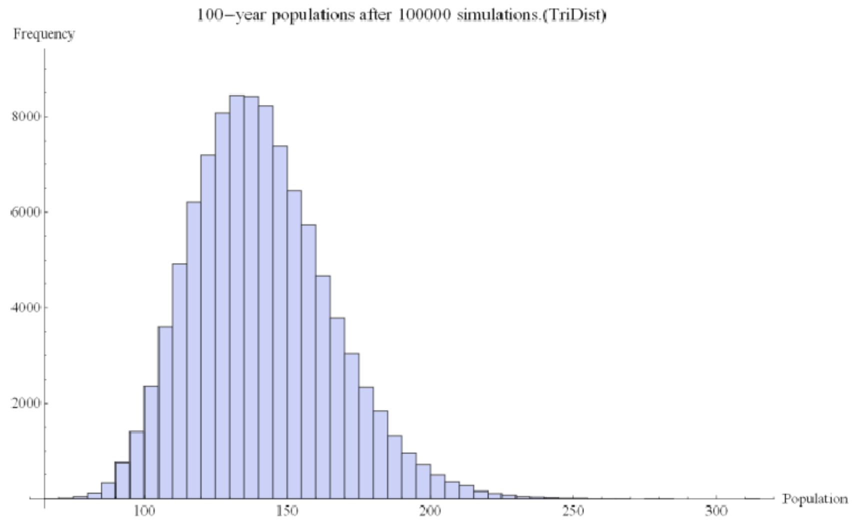
Line plot



Source: Lukman, A. F., Adewuyi, E., Månsson, K., & Kibria, B. G. (2021). A new estimator for the multicollinear Poisson regression model: simulation and application. *Scientific Reports*, *11*(1), 3732.

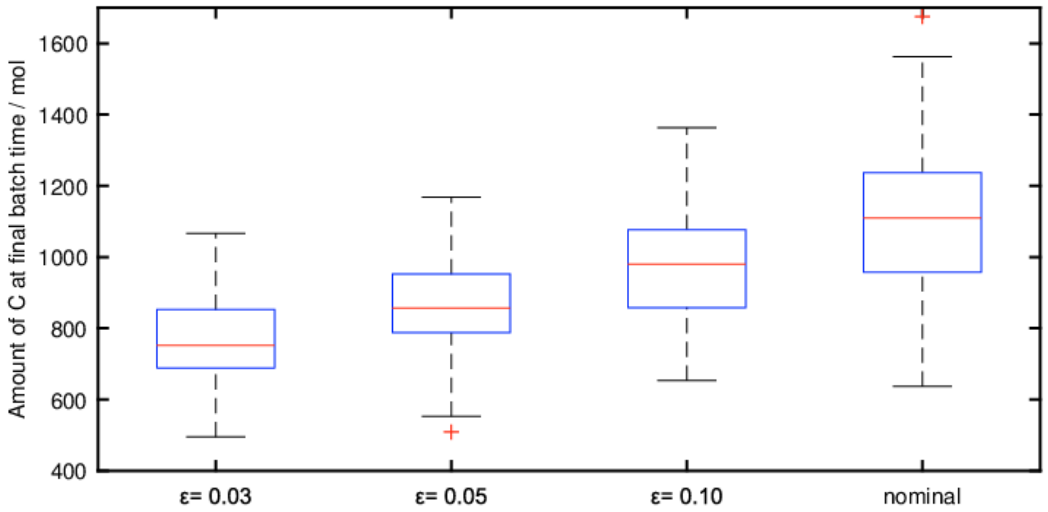
Visualization cont.

Histogram



Source: Hayes, M. A. (2011). *An analysis of fringed myotis (Myotis thysanodes), with a focus on Colorado distribution, maternity roost selection, and preliminary modeling of population dynamics*. University of Northern Colorado.

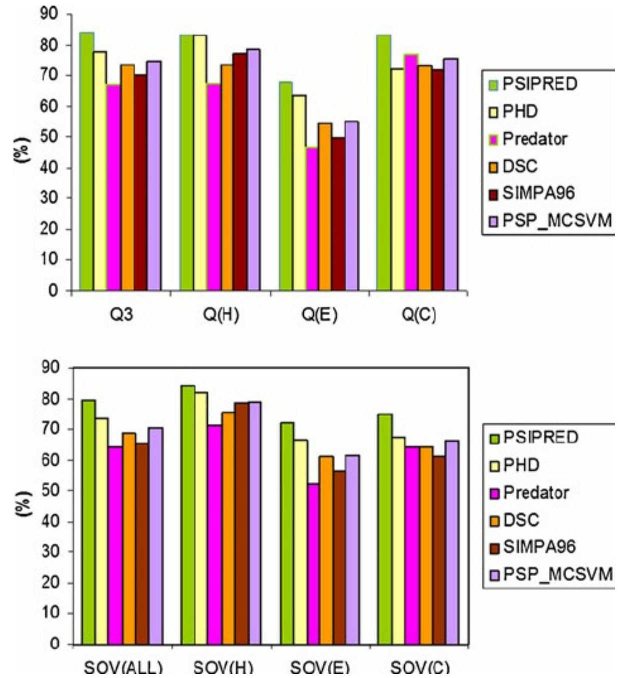
Box plot



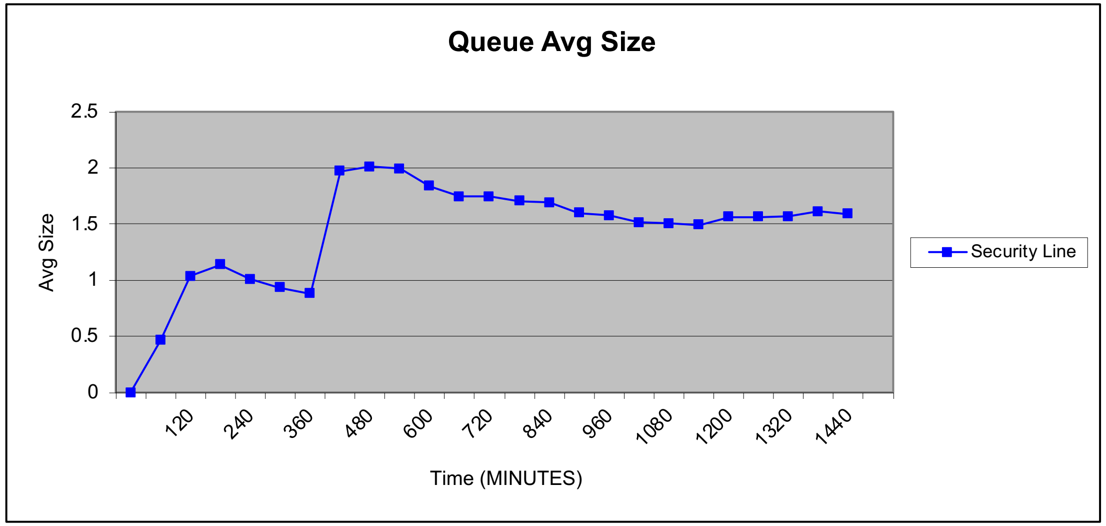
Source: Bradford, E., & Imsland, L. (2018). Economic stochastic model predictive control using the unscented kalman filter. *IFAC-PapersOnLine*, *51*(18), 417-422.

Visualization cont.

Bar chart/graph



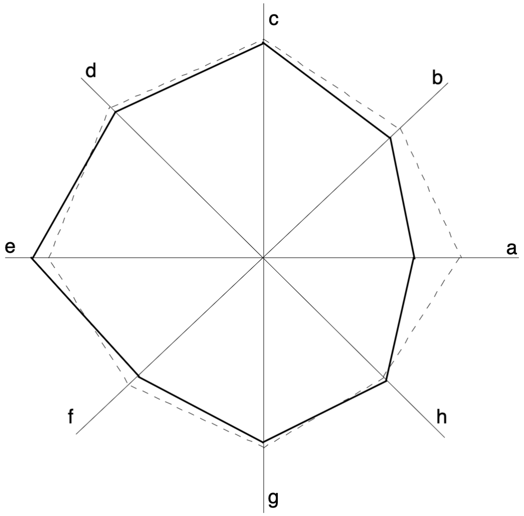
Time series plot



Source: Chatterjee, P., Basu, S., Kundu, M., Nasipuri, M., & Plewczynski, D. (2011). PSP_MCSVM: brainstorming consensus prediction of protein secondary structures using two-stage multiclass support vector machines. *Journal of molecular modeling*, *17*, 2191-2201.

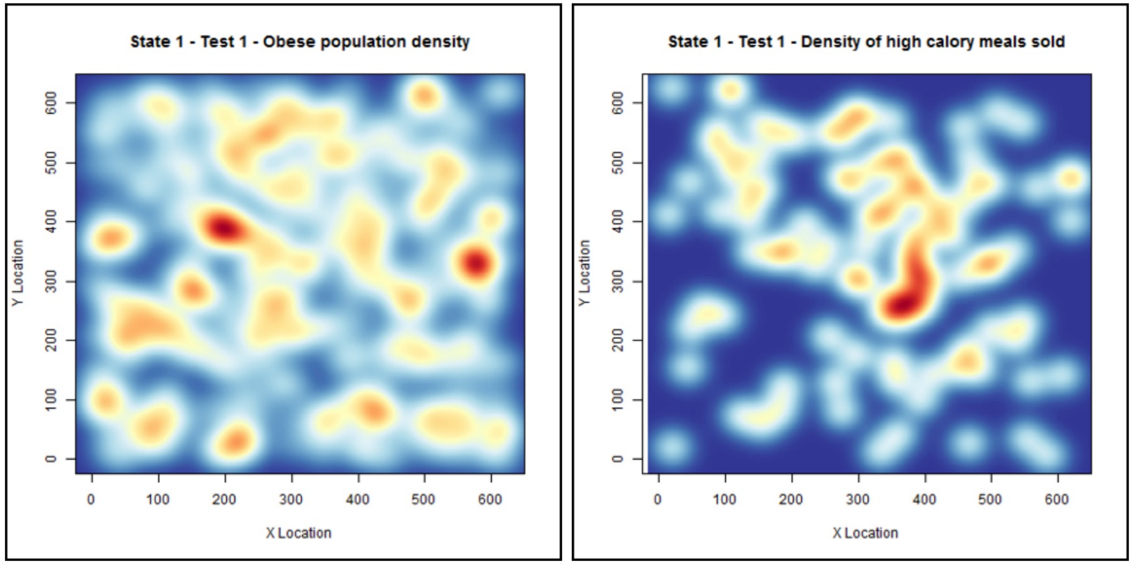
Visualization cont.

Polar diagram

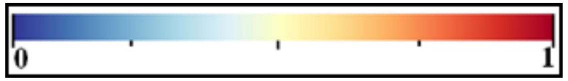


Source: M. I. Smith, D. J. Murray-Smith, and D. Hickman, "Verification and Validation Issues in a Generic Model of Electro-Optic Sensor Systems," The Journal of Defense Modeling and Simulation, vol. 4, no. 1, pp. 17-27, 2007.

Heat map



(a) Density of people classified as obese. (b) Density of high calorie meals sold.



(c) Color scale.

Source: Lynch, C. J., Kavak, H., Gore, R., & Vernon-Bido, D. (2019). Identifying unexpected behaviors of agent-based models through spatial plots and heat maps. *Complex Adaptive Systems: Views from the Physical, Natural, and Social Sciences*, 129-142.

Visualization cont.

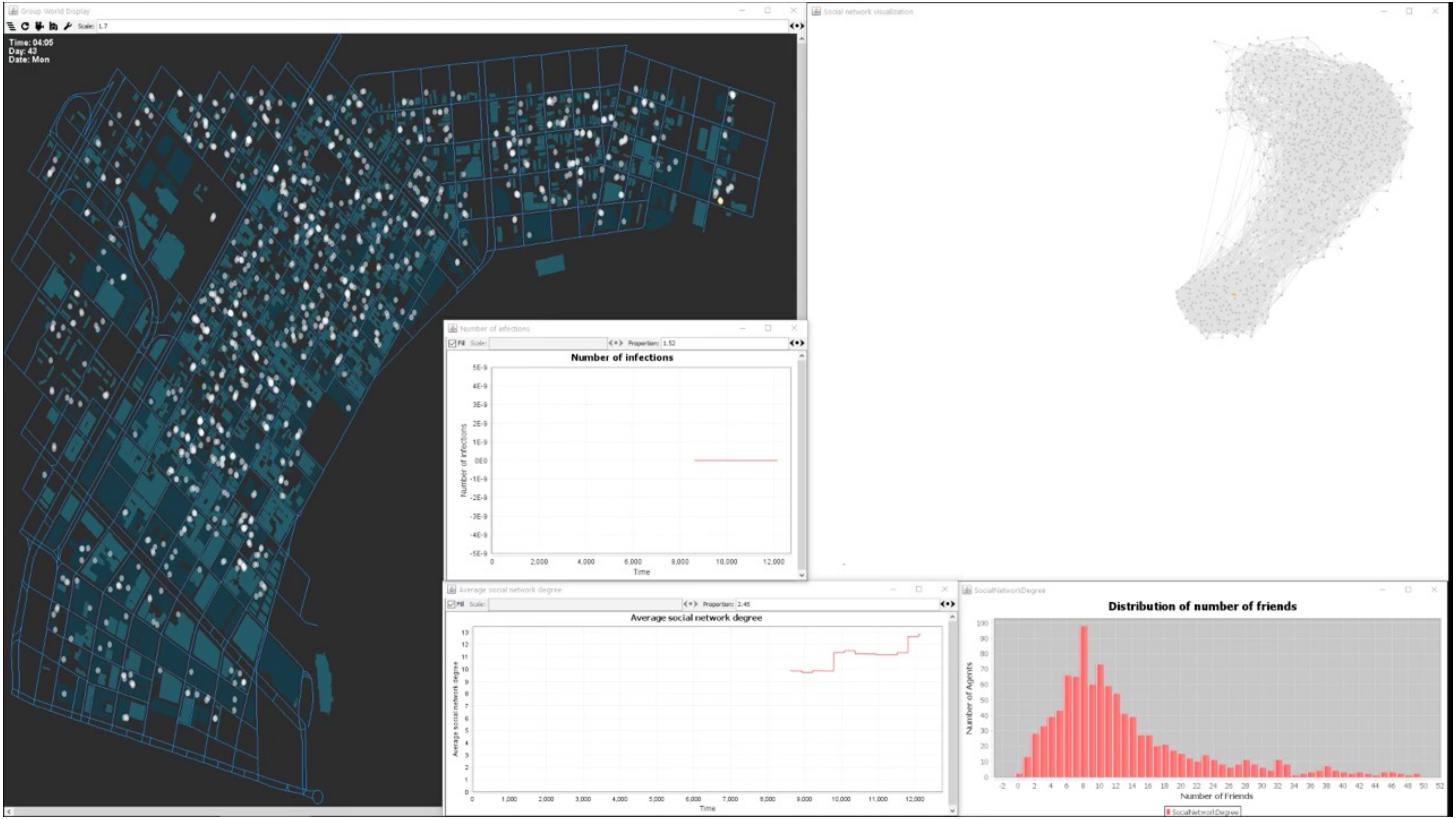
- Many of the techniques can capture uncertainty or error.
- These are just a fraction of techniques to visualize runtime simulation data. Others include.
  - Violin plots, geospatial plots, 3D plots, pie charts, radar charts, contours…
- Need to consider color selection for accessibility.

Animation

- Created by combining a set visualizations taken at specific times during simulation runs.

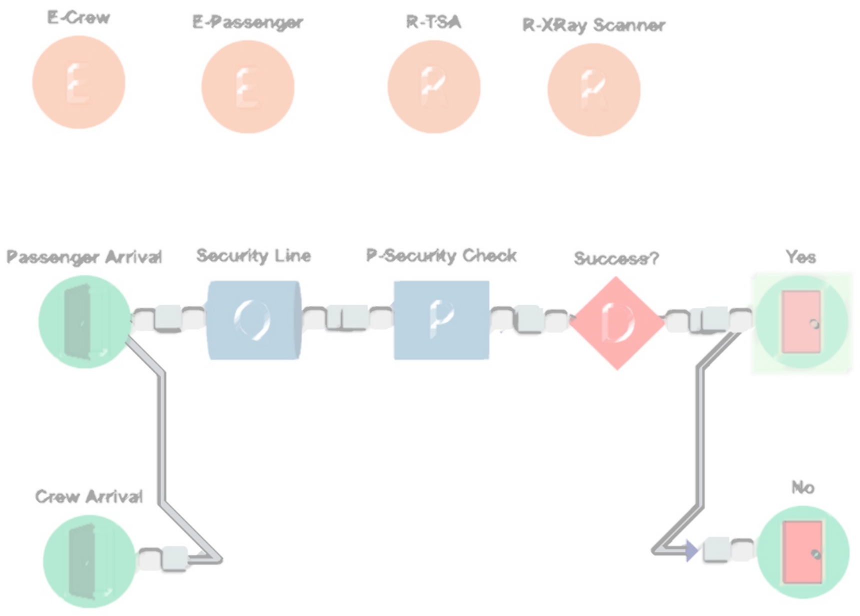
- Captures dynamic representation of simulation model behavior.

- Helps identify errors in the implementation by visual inspection.


Group World Display
Time: 04:05
Day: 43
Date: Mon
Number of infections
Number of infections
Time
Average social network degree
Average social network degree
Time
Distribution of number of friends
Number of Agents
Number of Friends

CLOUDES animation feature

- http://cloudes.me



Face validation

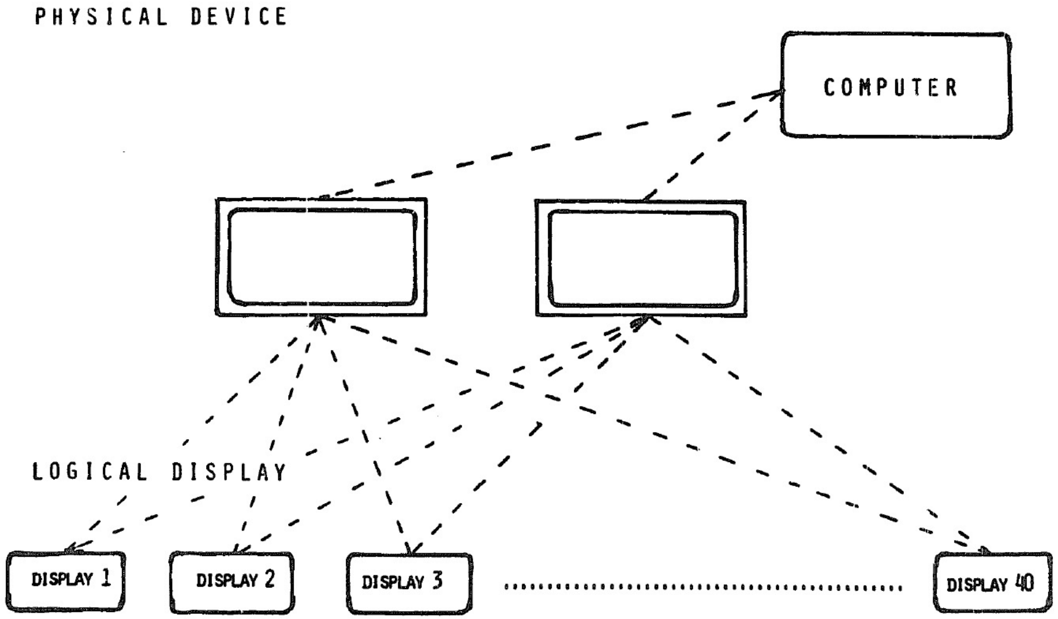
- Is a term used to describe the process of subject matter experts visually checking model runtime output.

- The goal is to identify verification and validation-related issues.

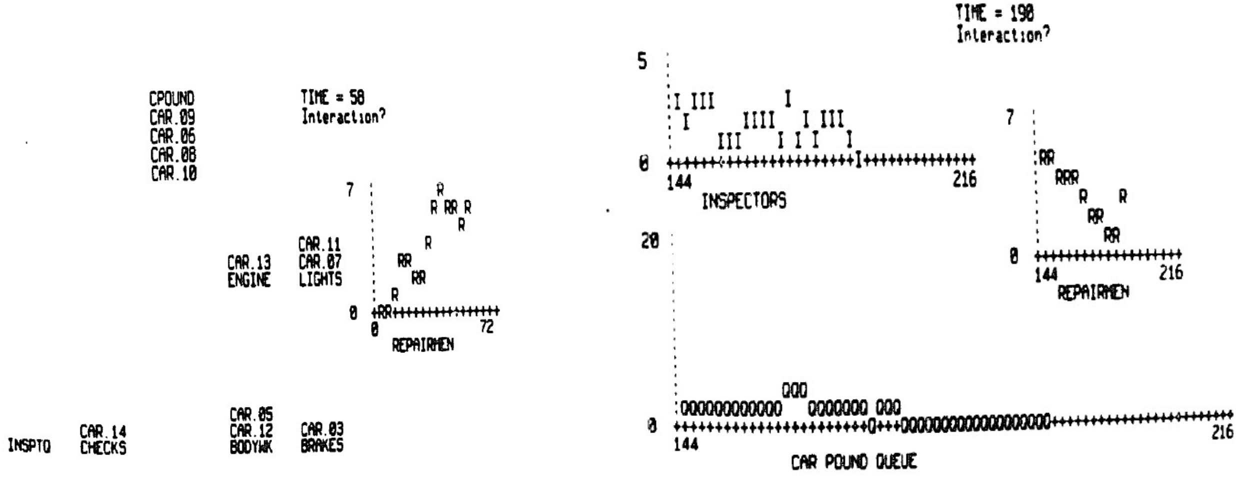
- Visualization and animation techniques are used to convey model processes, variables, decision factor, etc.

Historical example: Visual Interactive Simulation

<u>VISION dynamic display mechanism architecture</u>



<u>Example outputs</u>



Source: Hurrion, R. D. (1980). An interactive visual simulation system for industrial management. *European Journal of Operational Research*, 5(2), 86-93

Parameter verification tests

- Verifies that parameter values appear consistent with respect to the relevant descriptive and numerical knowledge provided of the system.
- Parameter values of interest need to be observable during run time.
- Animation or visualization techniques can be used to see parameter values and their change.

Source: Lynch (2019)

Extreme conditions tests

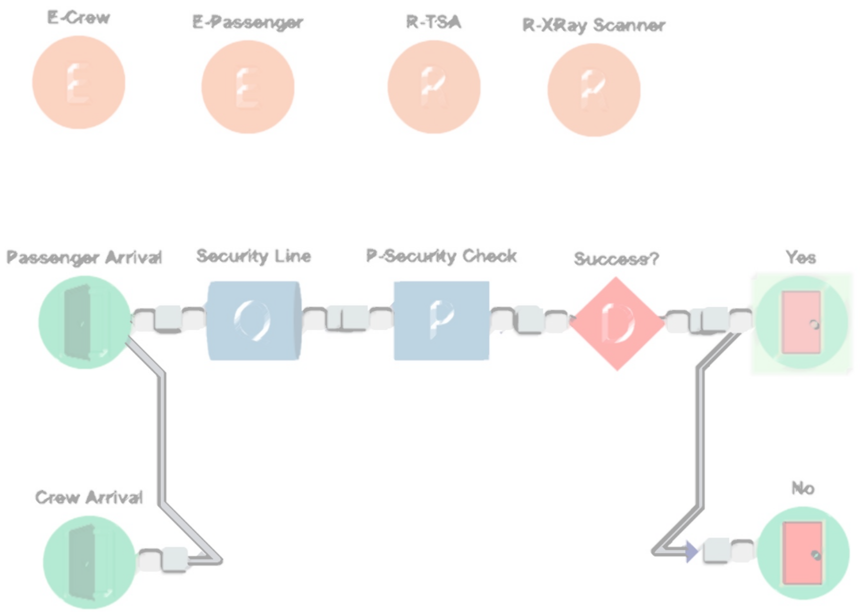
- Checks if plausible results occur when using extreme and unlikely combinations of parameter values.

- Example questions to check
  - Does the model structure allow for potential attempted calculations of dividing by zero?
  - How do model components behave as their denominator values approach zero?
  - For model components that are strictly greater than or less than a given value, does this bound hold throughout execution?
  - For model components that are greater than or less than a given value y, does y actually occur during execution?

Source: Lynch (2019)

Generating extreme condition on CLOUDES

- http://cloudes.me



Object flow testing

- Examines dynamic object flows throughout runtime to capture activities, processes, and interactions within the environment to identify errors.
  - E.g.: state transitions
  - Highly useful in simulations developed using object oriented programming.

- Visual inspection of object flows within the simulation environment can assist in verifying simulation component’s logic and behaviors by watching event-by-event execution, setting up conditions to force certain events, and tracing the progress of simulated entities

Source: Lynch (2019)

Traces

- Automatically track, reveal, and store data throughout the execution of a simulation over time.

- This trace data can be manually, visually inspected and compared with the expected occurrences to try and identify potential errors.

- Parameter verification and extreme conditions tests can be applied on trace data.

Source: Lynch (2019)

Traces cont.

- Especially useful in complex simulations like agent-based models.
- Instrumentation is required in order to develop the data structures for collecting or storing the trace values.
- Can become a burden to collect necessary traces.

Source: Lynch (2019)

Execution monitoring

- Execution monitoring aims to reveal the presence of errors by collecting and analyzing low-level information about activities and events that take place during execution.
- Occurrences are revealed to users with identifying details such as the context of their location and corresponding time.
- Instrumentation is required to calculate, collect, and/or provide the statistics necessary for uncovering potential errors.

Source: Lynch (2019)

Message sequence charts for process interaction models

- A combination of a trace-based and execution monitoring technique

- Used to verify that a model's implemented logic appears to be correct by aiding in the identification of areas suffering from starvation and deadlock.

- Instrumentation is generally required to obtain the needed simulation data
  - E.g.: message passing between simulation objects, resource utilization requests

- Feed information into the message sequence charts setup so that the process interaction visualizations can be generated for inspection.

Source: Lynch (2019)

The Gaemas Virtual Laboratory (GVL) tool

- Also, a combination of a trace-based and execution monitoring technique.

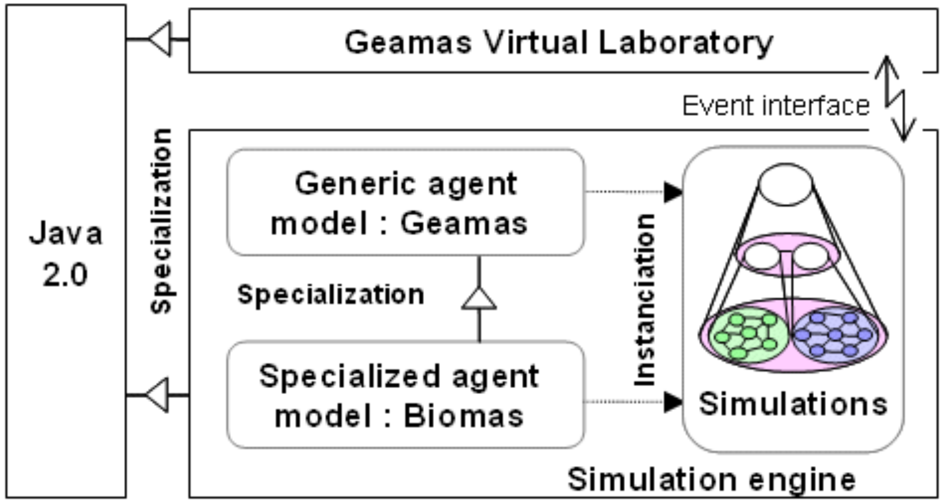
- The GVL traces and stores information of ABM components to examine the individual-level behaviors within the simulation at runtime to identify suspicious outcomes.

- Traces: sets of messages exchanged between agents (individual level), or a historical accounting of simulation execution per agent or group of agents (aggregate level)

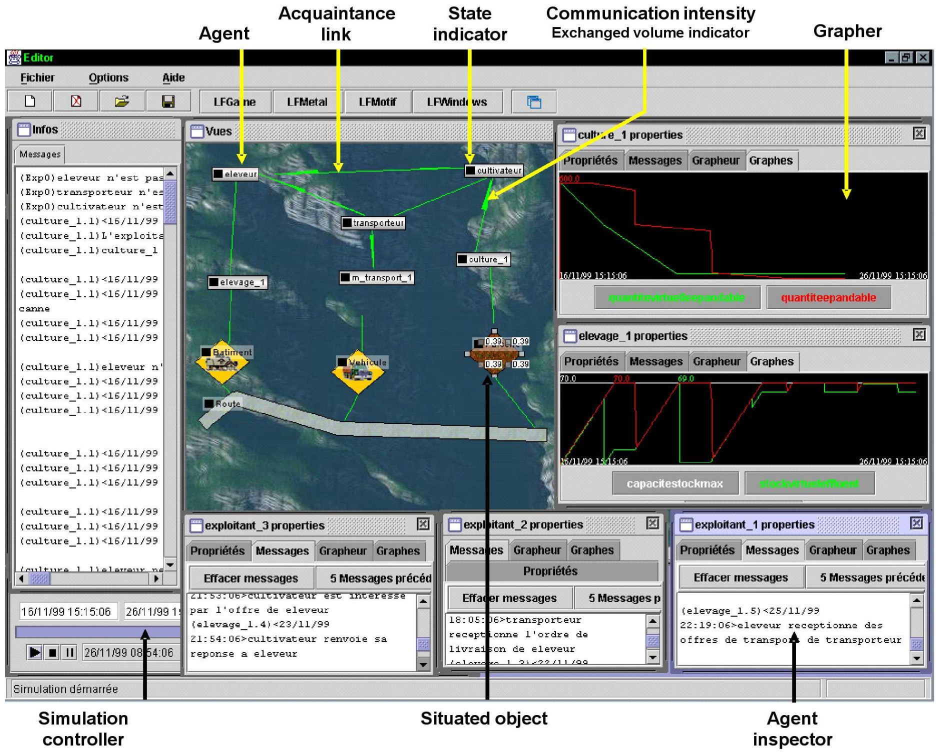
Source: Lynch (2019)

The Gaemas Virtual Laboratory (GVL) tool

The Geamas platform: relations between software modules



User interface



Source: Courdier, R., Guerrin, F., Andriamasinoro, F. H., & Paillat, J. M. (2002). Agent-based simulation of complex systems: application to collective management of animal wastes. *Journal of Artificial Societies and Social Simulation*, *5*(3).

Runtime verification techniques: recap

- Visualization
- Animation
- Parameter-verification test
- Extreme conditions test
- Object flow testing
- Traces
- Execution tracing
- Message Sequence Charts for Process Interaction Models

Source: Lynch (2019)

Pros and cons

Pros

- Displays simulation behavior within the context of a run
- Intuitive and less complex (Whitner and Balci 1989; Ng et al. 2004; Eek et al. 2015)
- Feedback supports interpretation (Bell and O'Keefe 1994; Whitner and Balci 1989)

Cons

- Time and resource intensive (Whitner and Balci 1989; Eek et al. 2015)
- Use and effectiveness are heavily dependent upon simulation platform (Glasow and Pace 1999)
- Monitoring is attention demanding (McCormick 1957; Crossan et al. 2000)

Source: Lynch (2022)

Lightweight, Feedback-driven Runtime Verification (LFV)

Goal

- Providing consistent application and interpretation to runtime verification by incorporating formal specifications.

Formal specifications

- Statements connecting components of the simulation to explicit requirements
- Specifications connect simulation observations to model requirements

Feedback-driven

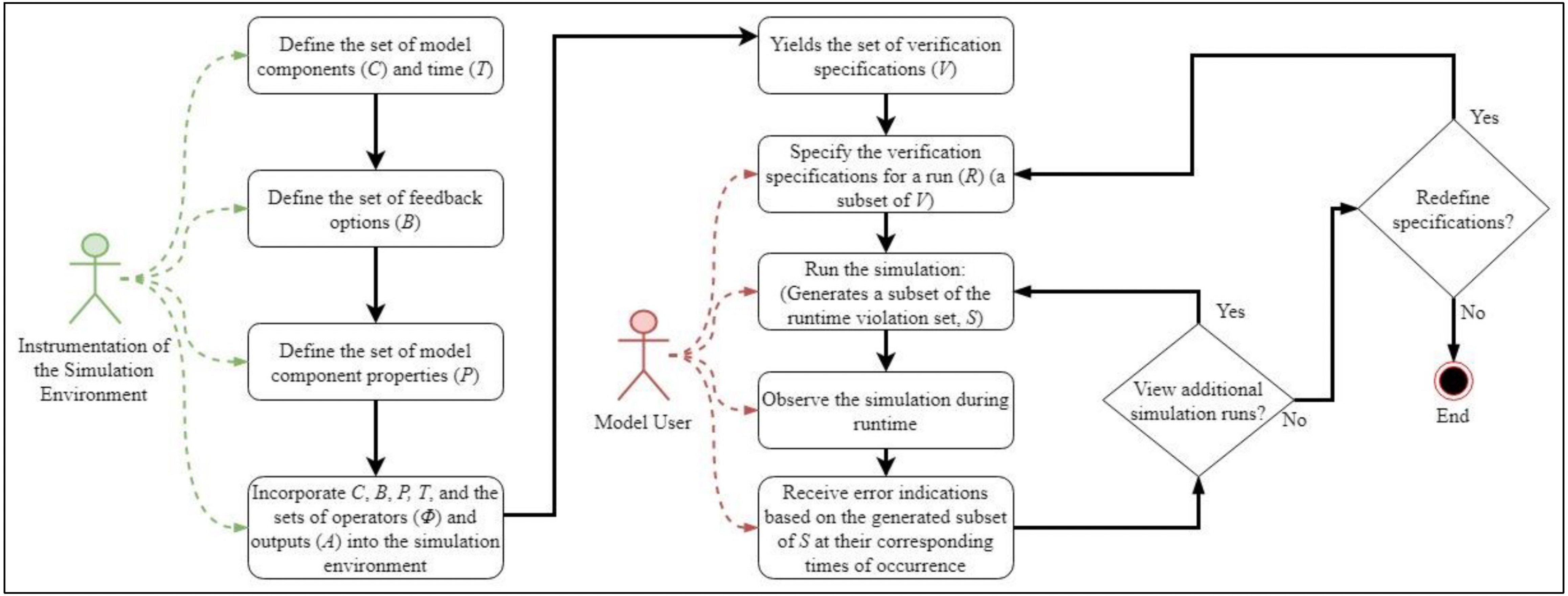
- Feedback indicates violations

Differentiates role of instrumentation and observation

- Structured application independent of paradigm
- Dynamic configuration

Source: Lynch (2022)

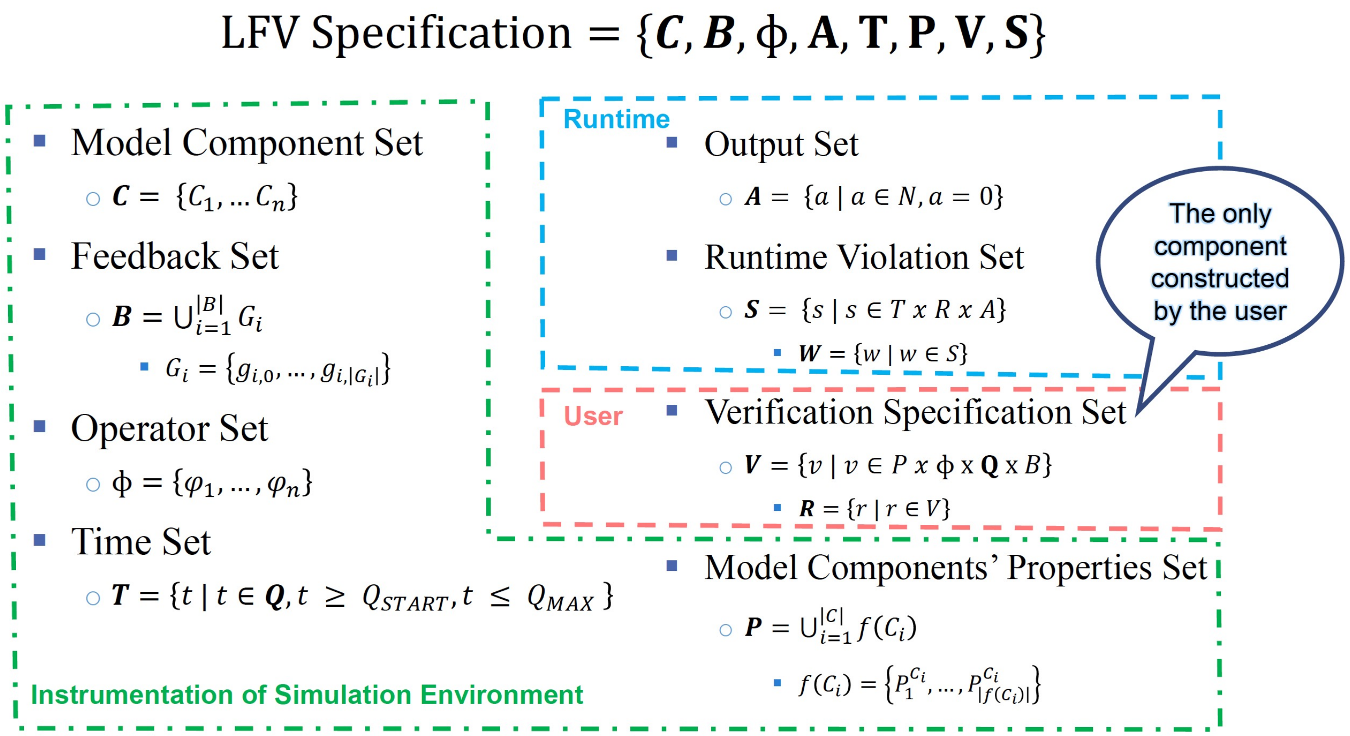
LFV - methodology



Source: Lynch (2022)

LFV - Recap

- Implemented in a separate CLOUDES instance (not available in the main platform).
- Easy to incorporate
- Various feedback options including sound



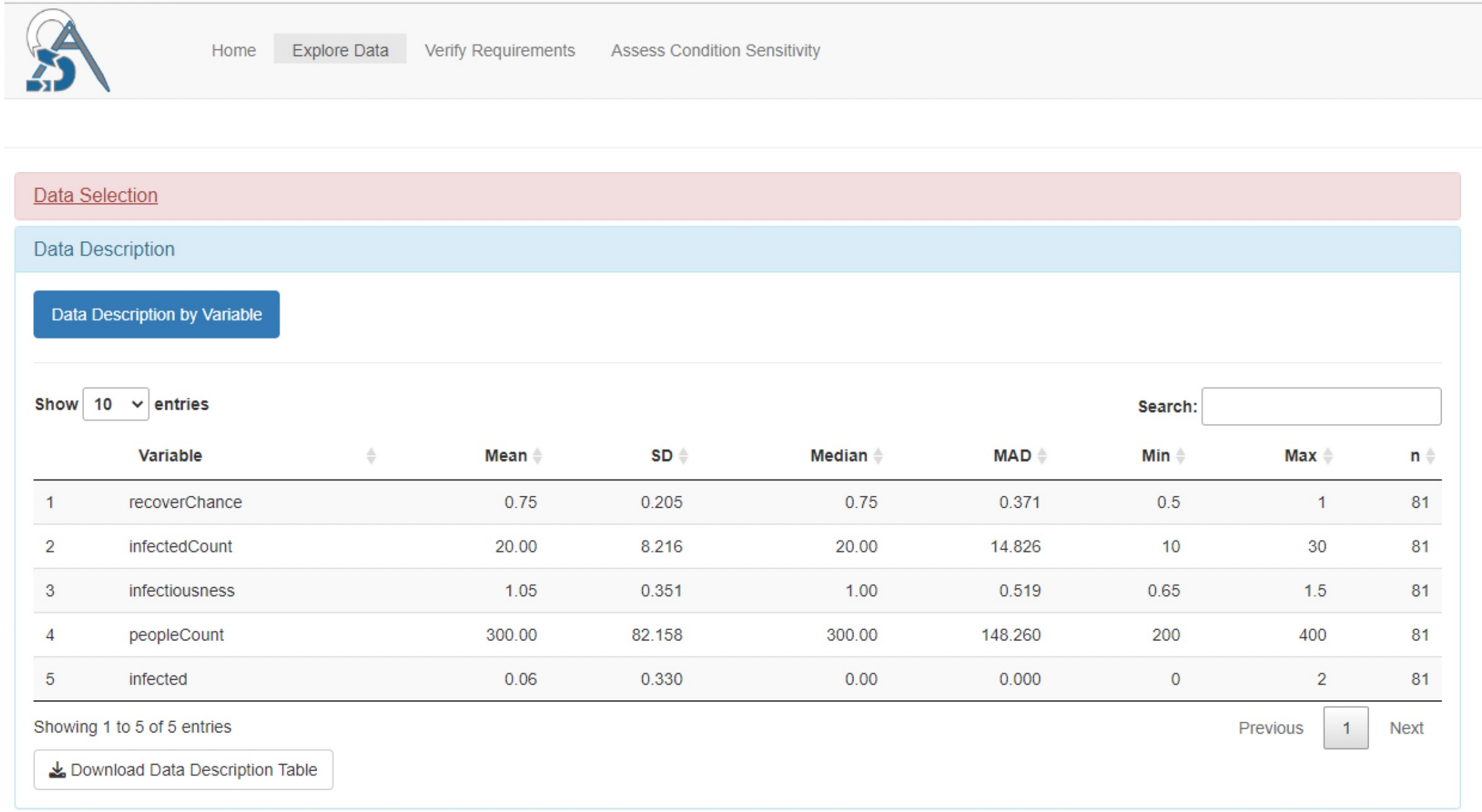
Source: Lynch (2022)

Sensitivity assessor

The Sensitivity Assessor (SA) assists in building evidence to support the interpretation and explanation of simulation model outcomes.

A consistent set of capabilities for exploring data yields assessments that are **reproducible**, **transparent**, and **easily shareable**.

- Exploration and sanity checking of the variables
- Verifying requirements that should always hold
- Quantifying the extent to which certain conditions contribute to, or fail to contribute to, a variable of interest.



	Variable	Mean	SD	Median	MAD	Min	Max	n
1	recoverChance	0.75	0.205	0.75	0.371	0.5	1	81
2	infectedCount	20.00	8.216	20.00	14.826	10	30	81
3	infectiousness	1.05	0.351	1.00	0.519	0.65	1.5	81
4	peopleCount	300.00	82.158	300.00	148.260	200	400	81
5	infected	0.06	0.330	0.00	0.000	0	2	81

The Sensitivity Assessor is publicly available and accessible at: https://vmasc.shinyapps.io/SensitivityAssessor/

Open challenges

Transparency

- Specifications explicitly connect simulation observations to the model

Reproducibility

- Specifications explicitly connect testing to simulation/solution space

Reuse

- Specifications explicitly connect users to model specifications and DoE

Documentation

- Specifications explicitly reflect model specifications, testing setup, and outcomes

Source: Lynch (2022)

Sources

- Lynch, C. J. (2022. V&V of Models. Invited Lecture Slides. George Mason University.
- Lynch, C. J. (2019). A Lightweight, Feedback-Driven Runtime Verification Methodology (Doctoral dissertation, Old Dominion University).