講演番号	033

Why Model Credibility Isn’t Enough:

- Rethinking Trust in Simulation Architectures -

Romain Barbedienne [1)] **Adeline Lanugue [1)]** **Rim Kaddah [1)]** **Julien Silande [1,2)]** **Anthony Levillain [3)]**
Cedric Leclerc [4)] **Maxime Hayet [5)]** **Boussaad Soualmi [1)]** **Cristian Maxim [1)]**

1) IRT SystemX, 2 Bd Thomas Gobert, 91120 Palaiseau, France (E-mail: romain.barbedienne@irt-systemx.fr)
2) Keysight Technology, La janais 3 rue Pierre et marie curie, 35131 Chartres-de-Bretagne, France
3) OPmobility Alphatech, ZAC du Bois de Plaisance, 214 Av. de la Mare Gessart, 60280 Venette, France
4)Renault Technocentre, 1 avenue du Golf, 78280 Guyancourt, France
Stellantis Green-Campus, 43 Rue Jean Pierre Timbaud, 78300 Poissy, France

KEY WORDS: Credibility, Model Architecture, Sensitivity Analysis

Computer simulation is widely used validate system requirements, as it is less expensive than real-life testing and allows numerous tests in a short time. However, simulation results must be credible to prevent errors in decision-making. In practice, engineers build simulation architectures composed of several submodels connected through interfaces, each carrying its own credibility level. This paper investigates whether the credibility of a simulation architecture can be assessed based on the credibility of its constituent submodels.

Several frameworks exist for assessing the credibility of individual simulation models. NASA-STD-7009B defines requirements covering data traceability, verification, validation, and technical review. The Predictive Capability Maturity Model (PCMM) evaluates model maturity across dimensions such as geometric fidelity, code verification, and uncertainty quantification. The Prostep ivip and IRT SystemX frameworks offer structured tools facilitating communication between stakeholders in industrial contexts. The European Commission regulation provides guidelines for certifying autonomous vehicles using virtual tools. However, all these frameworks are designed for standalone models. Their extension to simulation architectures raises unresolved challenges: most validation methods require experimental reference data unavailable at the architecture level, and none provides a systematic mechanism for addressing compositional aspects such as uncertainty propagation across interfaces.

To evaluate the impact of each submodel within an architecture, two families of methods are explored; sensitivity analysis methods, and statistical correlation methods. Sensitivity analysis methods, including local approaches (finite differences), global screening (Morris method, OAT, sequential bifurcation), variance-based approaches (Sobol indices, FAST), and metamodel-based methods (Kriging), provide tools for understanding how parameter variations propagate to outputs. However, these methods operate at the parameter level rather than at the submodel level, creating a conceptual mismatch with the problem of assessing submodel influence. Statistical correlation methods, such as Granger causality and Lasso regression, require only a single simulation run but are limited to detecting linear dependencies. Ablation studies from AI explainability target the identification of the most influential components within a system, which is conceptually aligned with the assembly problem, but cannot be directly applied because removing a physical submodel disrupts interface connectivity.

A structured comparison across objectivity, cost, scope, and maturity reveals a clear trade-off between rigor and computational cost. The most rigorous methods (Sobol indices) provide quantitative variance decomposition but at high computational expense, while statistical methods are efficient but restricted to linear models. No single method adequately addresses the assembly credibility problem. Among the reviewed approaches, the conceptual framework of ablation studies appears the most aligned with the problem structure. A viable adaptation could consist in substitution strategies, where a submodel is replaced by a simplified or degraded version rather than removed entirely, thereby preserving interface integrity while enabling assessment of the submodel’s contribution. Ultimately, the credibility of a simulation architecture cannot be reduced to the aggregation of individual model credibilities. It requires a dedicated methodological framework integrating compositional uncertainty propagation, interface characterization, and decision-relevant impact assessment.

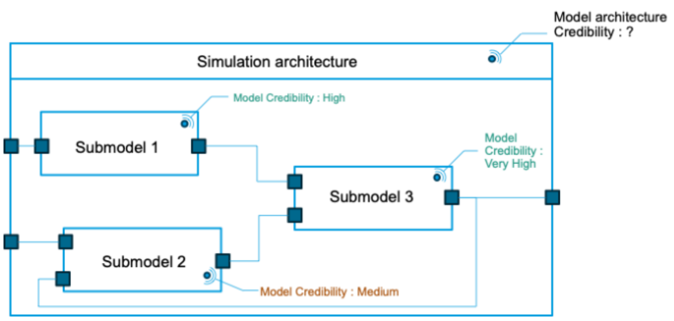

Figure 1 Credibility of simulation architecture