

A SURVEY OF VALIDATION IN HEALTH CARE SIMULATION STUDIES

Mohammad Raunak
Megan Olsen

Department of Computer Science
Loyola University Maryland
Baltimore, MD 21210, USA

ABSTRACT

The importance of proper verification and validation (V&V) of simulation models and experiments is well accepted in the modeling and simulation community. There is currently a push to improve the scientific aspects of modeling and simulation, including validation. To determine the amount of validation performed in practice, we analyze 110 health care simulation papers from the last eight years of the Winter Simulation Conference for the level of verification and validation reported and the types of validation techniques applied. Our results show that validation is not discussed sufficiently in most published papers. Although many authors are performing some level of validation on their simulation, close to a fourth of the papers do not mention V&V, while more than half do not provide detail of the type and amount of V&V performed.

1 Introduction

Modeling and Simulation has seen a steady increase in popularity as a useful way to perform “what-if” analysis of difficult-to-analyze systems. It is well accepted that validation of simulation models is difficult to ascertain, especially in published papers where validation details are often not included. Moreover, validity is usually not independently verified as studies are rarely replicated (Taylor et al. 2013, Tolk et al. 2013). Thus, despite the level of research conducted on verification and validation and the number of techniques developed for simulation validation, it is unclear how much verification and validation is actually performed. In this work, we survey health care related simulation papers to investigate this question.

Previous surveys on simulation papers in the health care application domain noted that health care modelers almost never reuse models produced by others, and the study results or recommendations are rarely implemented in the real world (Thorwarth and Arisha 2009, Gunal and Pidd 2010). We argue that the primary reason behind both the lack of reuse and implementation is due to the inability to rely on the published models being sufficiently validated. It is worth noting that none of the simulation survey papers in the existing literature analyzed simulation model validation reporting in papers.

2 The Survey

Our survey focuses on health care simulations as it has been one of the most studied areas in recent years. From the 192 papers published in the ‘Health Care Application’ segment of WSC from 2006 to 2013, we examine 110 papers that include development and experimentation with a concrete executable simulation model. We analyze these papers on the type of validation activities reported, the tools used to develop and analyze the model, use of statistical analysis in validation, and whether the papers include a validation section. Based on the level of validation discussion, each surveyed paper was rated as “Yes” (describes approach(es), provides data), “Minimal” (stated that validation was performed, no details nor data), or “No” (no validation mentioned). We also look at the validation techniques used in papers that discuss validation.

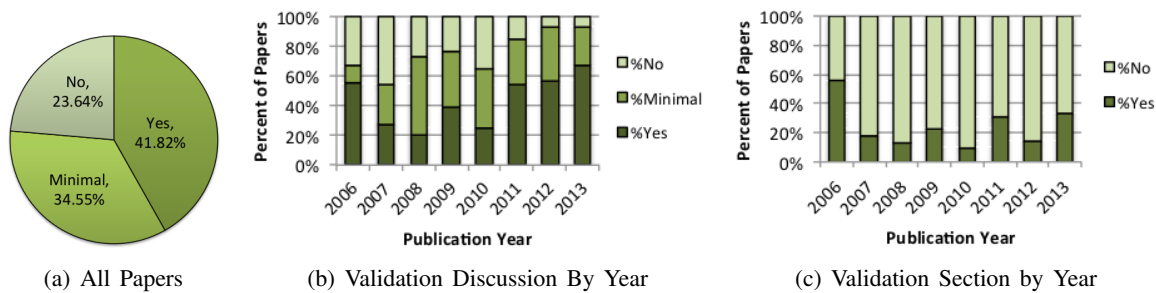


Figure 1: (a) Percentage of papers with each of the rankings on their validation discussion: Yes, Minimal, No. (b) Percentages by year. (c) Percentage of papers with a validation section by year.

In general, we observe a lack of standard terminology in the surveyed papers for describing their validation process and techniques. For example, the term *verification* is sometimes incorrectly exchanged for *validation*, and some papers do not use standard validation technique names. For our analysis, we identify validation techniques based on the description of validation process or data presented in the papers.

We find that 35% of papers have 'Minimal' discussion of validation, while 42% discuss validation in detail ('Yes'). Close to a fourth (26 papers) do not discuss validation ('No'), but 11.6% of those papers mention validation as future work. Although the trends over time are clearly not linear, the percentage of papers not discussing validation ('No') decreased from 2006 (33.3%) to 2013 (6.7%). On average, 23% of all papers include a separate validation section. Of all papers surveyed, 12% use statistical testing to accept or reject their model, and on average 1.08 (standard error 0.11) validation techniques are used. Papers that mention at least one validation technique discuss 1.8 techniques on average (standard error 0.13).

The number of techniques utilized, however, is an insufficient measure of validation activity as some techniques are more powerful than others. The most common validation technique used is "result validation" (30% of all papers) where results of the simulation output is matched with the data from the real world. The second most common validation technique is "face validation" (15% of all papers). "Input validation" is used in 12.5% papers. When only one validation technique is discussed in a paper, "Result validation" is the most common technique, while the others are used rarely by themselves. Animation, which can only give a nominal amount of confidence in the model, has been found to be used only in conjunction with other techniques.

These results show that the researchers who do discuss validation well in their papers are either validating appropriately, or on the right track to sufficient validation. However, too many papers are either not discussing validation or are not discussing it with enough depth. There is a strong need in the community to develop a set of standard terms for V&V activities and to ensure that these are used consistently in published papers, as proposed by Tolk et al. (2013).

REFERENCES

- Gunal, M., and M. Pidd. 2010. "Discrete event simulation for performance modelling in health care: a review of the literature". *Journal of Simulation* 4:42–51.
- Taylor, S., S. Brailsford, S. Chick, P. Ecuier, C. Macal, and B. Nelson. 2013. "Modeling and Simulation Grand Challenges: An OR/MS Perspective". In *Proceedings of the 2013 Winter Simulation Conference*, 1269–1282. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Thorwarth, M., and A. Arisha. 2009. "Application of Discrete-Event Simulation in Health Care: a Review". Technical report, Dublin Institute of Technology.
- Tolk, A., B. Heath, M. Ihrig, H. Padilla, E. Page, E. D. Suarez, C. Szabo, P. Weirich, and L. Yilmaz. 2013. "Epistemology of Modeling and Simulation". In *Proceedings of the 2013 Winter Simulation Conference*, 1152–1166. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.