Model-independent comparison of simulation output

Nuno Fachada\textsuperscript{a,}\textsuperscript{*}, Vitor V. Lopes\textsuperscript{b,c}, Rui C. Martins\textsuperscript{d}, Agostinho C. Rosa\textsuperscript{a}

\textsuperscript{a} Institute for Systems and Robotics (ISR/IST), LARSyS, Instituto Superior Técnico, Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal
\textsuperscript{b} UTEC – Universidad de Ingeniería & Tecnología, Lima, Jr: Medrano Silva 165, Barranco, Lima, Perú
\textsuperscript{c} CMAF-CIO, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
\textsuperscript{d} INESC TEC, Campus da FEUP, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

1. Introduction

Complex systems are usually described as consisting of mutually interacting objects, often exhibiting complex global behavior resulting from the interactions between these objects. This behavior is typically characterized as “emergent” or “self-organizing” as the system’s constituting parts do not usually obey a central controller \cite{1}. Analytic treatment generally does not yield the complete theory of a complex system. Therefore, modeling and simulation techniques play a major role in our understanding of how these systems work \cite{2}. Methodologies such as agent-based modeling (ABM), system dynamics, discrete event simulation, among others, are frequently employed for this purpose \cite{3}. Of these, ABM provides an instinctive approach for describing many complex systems, as agents are regularly a suitable match to the individual and heterogeneous objects composing these systems. The local interactions of these objects, as well as their individual and adaptive behavior, are often critical for understanding global system response \cite{4,5}. ABMs are commonly implemented as a stochastic process,
and thus require multiple runs (observations) with distinct pseudo-random number generator (PRNG) seeds in order to have appropriate sample sizes for testing hypotheses and differentiating multiple scenarios under distinct parameterizations [6].

Computational models of complex systems in general, and ABMs in particular, are usually very sensitive to implementation details, and the influence that seemingly negligible aspects such as data structures, discrete time representation and sequences of events can have on simulation results is notable [7]. Furthermore, most model implementations are considerably elaborate, making them prone to programming errors [8]. This can seriously affect model validation\footnote{Determining if the model implementation adequately represents the system being modeled [9] for its intended purpose [10].} when data from the system being modeled cannot be obtained easily, cheaply or at all. Model verification\footnote{Determining if the model implementation corresponds to a specific conceptual model [9].} can also be compromised, to the point that wrong conclusions may be drawn from simulation results.

A possible answer to this problem is the independent replication of such models [8]. Replication consists in the reimplementation of an existing model and the replication of its results [11]. Replicating a model in a new context will sidestep the biases associated with the language or toolkit used to develop the original model, bringing to light dissimilarities between the conceptual and implemented models, as well as inconsistencies in the conceptual model specification [9,12]. Additionally, replication promotes model verification, model validation [9], and model credibility [11]. More specifically, model verification is promoted because if two or more distinct implementations of a conceptual model yield statistically equivalent results, it is more likely that the implemented models correctly describe the conceptual model [9]. Thus, it is reasonable to assume that a computational model is untrustworthy until it has been successfully replicated [12,13].

Model parallelization is an an illustrative example of the importance of replication. Parallelization is often required for simulating large models in practical time frames, as in the case of ABMs reflecting systems with large number of individual entities [14]. By definition, model parallelization implies a number of changes, or even full reimplementation, of the original model. Extra care should be taken in order to make sure a parallelized model faithfully reproduces the behavior of the original serial model. There are inclusively reports of failure in converting a serial model into a parallel one [15].

Although replication is considered the scientific gold standard against which scientific claims are evaluated [16], most conceptual models have only been implemented by the original developer, and thus, have never been replicated [8,9,11,17]. Several reasons for this problem have been identified, namely: a) lack of incentive [8,11]; b) below par communication of original models [16,18]; c) insufficient knowledge and uncertainty of how to validate results of a reimplemented model [9]; and, d) the inherent difficulty in reimplementing a model [8,9,12]. This work targets c), with positive influence on d). Replication is evaluated by comparing the output of the reimplementation against the output of the original model [11], and this process, as will be discussed, is empirically driven and model-dependent (or even parameter-dependent). Furthermore, it is sometimes unclear as to what output features best describe model behavior. A robust and ready to use output comparison method would thus reduce or eliminate uncertainty of how to validate reimplementation results (reason c), eliminating this obstacle in the overall process of model replication (reason d).

We present a model comparison technique, which uses principal component analysis (PCA) [19] to convert simulation output into a set of linearly uncorrelated statistical measures, analyzable in a consistent, model-independent fashion. It is appropriate for ascertaining statistical equivalence of a model replication with its original implementation. Besides model-independence, this technique has three additional desirable features: a) it automatically selects output features that best explain implementation differences; b) it does not depend on the distributional properties of simulation output; and, c) it simplifies the modelers’ work, as it can be used directly on model output, avoiding manual selection of specific points or summary statistics. The proposed method is presented within the broader context of comparing and estimating the statistical equivalence of two or more model replications using hypothesis tests. The technique is evaluated against the empirical selection of output summary statistics, using the PPHPC ABM [20] as a test case. This model is thoroughly studied in terms of simulation output for a number of parameter configurations, providing a solid base for this discussion.

The paper is organized as follows. First, in Section 2, we review commonly used methods for comparing the output of simulation models, as well as previous work on model replication using these methods. An overview of hypothesis testing within the scope of model output comparison is conducted in Section 3. The main steps in designing and performing a model comparison experiment using hypothesis tests are presented in Section 4. The proposed model-independent comparison methodology is described in Section 5. Section 6 introduces PPHPC, the simulation model used as a test case for the proposed model comparison approach. Section 7 delineates the experimental setup for assessing this methodology against the manual selection of output summary measures. In Section 8, results of the empirical and model-independent comparison approaches are presented. A discussion and an evaluation of these results is performed in Section 9. Recommendations on using the proposed method are given in Section 10. Section 11 summarizes what was accomplished in this paper.

2. Background

Axtell et al. [17] defined three replication standards (RSs) for the level of similarity between model outputs (Carley [21] calls the RS the emphasis of demonstration): numerical identity, distributional equivalence and relational alignment. The first, numerical identity, implies exact numerical output, but it is difficult to demonstrate for stochastic models and not critical for showing that two such models have the same dynamic behavior. To achieve this goal, distributional equivalence
دریافت فوری متن کامل مقاله
امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات