Journal of Hunan University

(Natural Sciences)

Vol. 52 No. 1

January 2025

Available online at

https://jonuns.com

ELSEVIER
Scopus

Clarivate
WEB OF SCIENCE

Original research article

https://doi.org/10.55463/issn.1674-2974.52.1.3

Modeling and Simulation with System Dynamics: A Literature Review on Validation Aspects

Andrés Rey Piedrahita*[1], Laura Milena Cárdenas Ardila[2], Jorge Andrick Parra Valencia[3], Ángel José Lozada Das Dores[1] and José Gabriel Pérez Canencio[1]

[1] Faculty of Engineering, Unidad Central del Valle del Cauca, Colombia

[2]Universidad de Antioquia, Colombia

[3]Universidad Autónoma de Bucaramanga, Colombia

* Corresponding author: arey@uceva.edu.co

Article History:

Received: November 18, 2024
Reviewed: December 28, 2024
Revised: January 17, 2025
Accepted: January 25, 2025
Published: February 20, 2025

Abstract: This study **aimed** to address the critical issue of model validation in the field of system dynamics (SD) modeling. **Methods:** To achieve this objective, a comprehensive literature review was conducted to examine the existing validation approaches and practices. The analysis investigated the historical development of the validation methods, identified key trends and challenges, and explored the diverse range of qualitative, quantitative, and mixed methods employed in SD model validation. **Findings:** The findings revealed a significant gap between the theoretical importance of validation and its consistent practical application. The findings of this study led to the formulation of a novel conceptual framework for SD model validation. **Novelty:** This framework provides a structured guide for researchers to conduct more rigorous and comprehensive validation analyses that incorporate a broader range of techniques and considerations. The aim of the proposed framework is to enhance the credibility and reliability of SD models, thereby improving confidence in their use for decision-making and policy analysis.

Keywords: Complex systems, modeling and simulation, model validation, system dynamics, emergent behavior, validation methods, simulation analysis.

系统动力学建模与仿真：验证方面的文献综述

摘要：目的：本研究旨在解决系统动力学 (SD) 建模领域中模型验证的关键问题。方法：为了实现这一目标，进行了全面的文献综述，以研究现有的验证方法和实践。分析调查了验证方法的历史发展，确定了主要趋势和挑战，并探索了 SD 模型验证中采用的各种定性、定量和混合方法技术。结果：研究结果表明，验证的理论重要性与其在实践中的一致应用之间存在巨大差距。研究结果促成了 SD 模型验证的新型概念框架的形成。新颖性：该框架为研究人员提供了一个结构化的指南，以进行更严格和全面的验证分析，并结合了更广泛的技术和考虑因素。所提出的框架旨在提高 SD 模型的可信度和可靠性，从而提高其用于决策和政策分析的信心。

关键词：复杂系统、建模与仿真、模型验证、系统动力学、突发行为、验证方法、模拟分析

1. Introduction

A complex system is an entity composed of interconnected elements, parts, or components that interact dynamically and continuously [1]. These systems have emergent properties that arise from the structure and interrelationships of their components, and cannot be fully understood by analyzing individual parts. The emergent properties represent behaviors or characteristics that arise from the interaction of the system as a whole and are not found in isolated elements. Moreover, when these systems are subjected to external perturbations, they can respond with new properties through self-organizing processes that manifest in nonlinear and often unpredictable ways [2-3].

System Dynamics (SD) is an ideal tool for the in-depth study and analysis of complex systems. This continuous modeling methodology allows the simulation and understanding of how different elements and their interactions contribute to the overall behavior of the system. It has proven to be highly effective in analyzing complex phenomena and solving problems in various fields, including resource management, policy planning, and economics ways [4-5].

Complex systems are ubiquitous in our reality and include notable examples such as transport and communication systems, biological organisms and the human brain, critical infrastructures such as power grids and electronic systems, Earth's climate behavior, and the social and economic structures that shape our societies [6-10]. These examples illustrate the diversity and relevance of complex systems for understanding and managing the inherent complexity of today’s environment.

The use of models and simulations has played a fundamental role in science and other fields, driven by significant advances in computer science and artificial intelligence, which have aroused growing interest among academics and professionals [11]. A model representing a real-world phenomenon is defined as a set of formal representations, including the specification of the elements of the system, relationships between these elements, parameters that influence the behavior of the system, and links to other relevant real-world systems. It is essential that the knowledge base used to construct the model is accurate and comes from reliable sources relevant to the phenomenon under study [12].

One of the major advantages of computer simulation applied to models is its ability to analyze and explore real-world phenomena that are extremely difficult or impossible to observe in controlled laboratory environments. In addition, simulations make it possible to solve highly complex mathematical problems whose computational complexity makes a manual analytical solution impractical [13]. The ability to address these challenges using computational simulations has expanded the frontiers of research and development, providing powerful tools for in-depth understanding and innovation in various disciplines.

The use of models facilitates the study of real-world performance by allowing the inherent complexity and uncertainty of the systems to be managed [14]. This approach begins with a rigorous effort to understand the dynamics of real systems and their associated problems in depth, providing a solid foundation for optimizing the efficient control and management of such systems and their challenges.

One of the most debated aspects of scientific research is the validity of the results. In model-based methodologies, validation plays a central role because the robustness and credibility of conclusions depend directly on the integrity and fidelity of the model used. However, in the field of system dynamics, the literature for decades has highlighted the worrying lack of attention to consistent and rigorous validation efforts [15-16]. This imbalance underscores the need to redouble efforts to apply validation techniques that will increase confidence in models and their predictions and thus lead to more significant advances in research and practice.

This study focused on the critical issue of model validation within the domain of system dynamics (SD) modeling. Acknowledging the pivotal role of model credibility in determining the reliability of simulation outcomes and subsequent decision-making processes, this study provides a comprehensive analysis of the pivotal aspects advocated for the validation of SD models. To achieve this objective, a comprehensive review of the extant literature was conducted. This encompassed a wide range of studies that examined the historical development of validation practices, identified key trends and challenges, and explored the diverse range of qualitative, quantitative, and mixed methods employed in the SD model validation. This comprehensive analysis culminates in the formulation of a novel conceptual framework. This framework not only summarizes and highlights the most relevant elements identified in the literature to enable future research to carry out more rigorous validation analyses but also provides a roadmap for researchers to apply these findings to enhance the credibility and reliability of their SD models across various domains, including but not limited to environmental management, social systems, and economic policy.

2. Fundamentals

2.1 Modeling and simulation life cycle

In practice, model validation is implemented through a series of tests that must be integrated as an essential part of the modeling lifecycle. The selection of tests depends on the type of model in question and is carried out according to a logical sequence of formal steps designed for the validation process, the end result of which may be the acceptance of the model or identification of errors requiring correction. The concept of the modeling and simulation lifecycle refers to a structured set of stages, each of which involves specific tasks and processes that contribute to the progressive construction of the model. These stages not only ensure the consistency and robustness of the model but also allow for a critical evaluation of its behavior and effectiveness. Figure 1 schematically illustrates the typical composition of this cycle, while Table 1 details the different stages and key activities associated with each cycle.

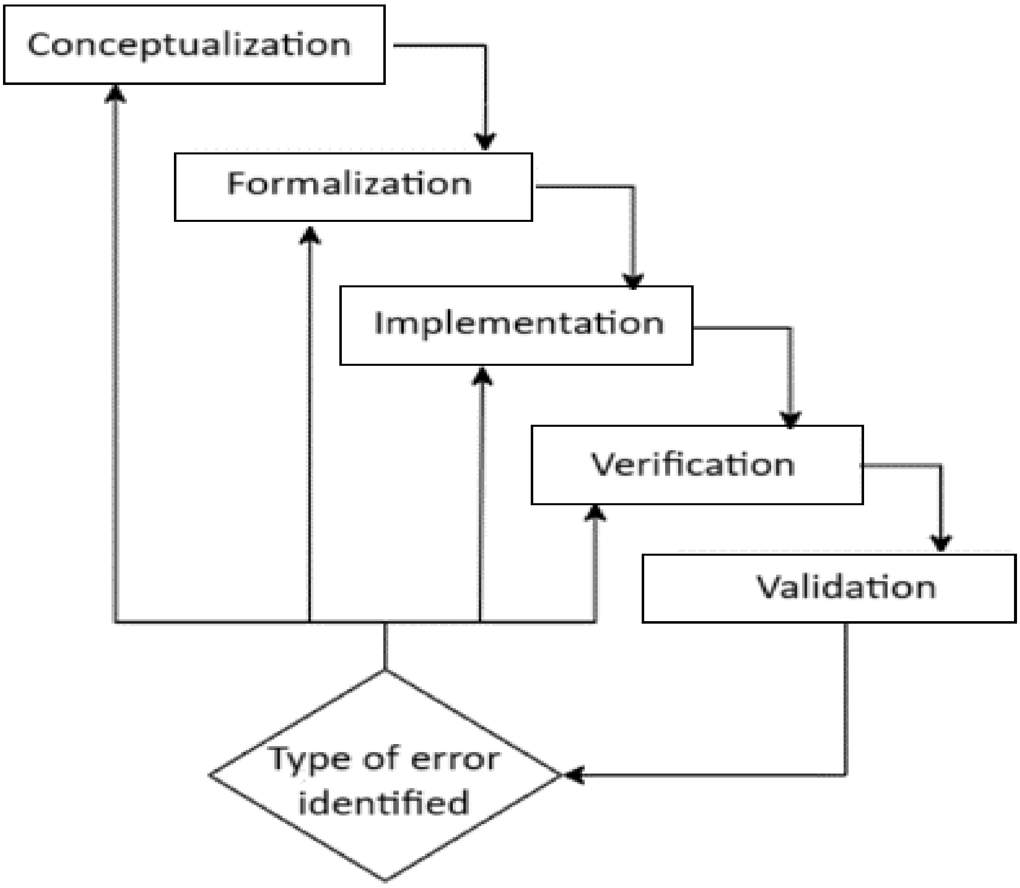

Figure 1. The cycle of the modeling and simulation. Source: Adopted from [12].

Table 1. Stages in the modeling and simulation cycle

Stage	Description
Conceptualization	Here, we first analyze the problem, which must be clearly and precisely formulated, and then proceed to the general analysis of the real system under study, with the aim of identifying the elements of the system and the relevant relationships between them, in order to elaborate a kind of narrative model of the system or similar.
Formalization	Basically, at this stage, the key variables of the model are formally defined (using the results of the previous stage) and the corresponding relationships between the variables are also formally defined, as well as the units and time scales required to solve the proposed problem. The product is a model that can be called a schematic.
Implementation	In the implementation stage, the schematic model is translated into a computer model, which is a codified representation elaborated in computer language, which is required to reorganize equations and adequately represent the different operations.
Functional verification	Functional verification is carried out with tests to assess, for example, whether the model produces reasonable results based on experience, how it behaves when variables are changed, etc.
Validation	Validation consists of the application of various tests, essentially aimed at comparing the model and its behavior with equivalent models developed in previous third-party studies.

Source: [12, 13, 17].

Figure 1 shows one of the most widely accepted models for modeling and simulation cycles. Although there are several versions of the literature that may differ in certain aspects, they all share a similar structure in terms of basic phases and activities. The first three phases of the cycle are well defined and rarely cause ambiguity. However, the functional verification and validation phases are often a source of confusion and debate. Functional verification focuses on ensuring that the results generated by the computational simulation are consistent with empirical observations of the real system.

However, these results may be inconsistent or invalid if the model is inconsistent with existing knowledge and the fundamental laws governing the real system, an aspect that is specifically addressed in the validation stage.

Several authors argue that the objective of verification is to assess the integrity and quality of the various representations and transformations that the model undergoes during the modeling process, as well as the algorithm used for the simulation and the computational tools used. By contrast, validation seeks to determine whether the simulation model is an accurate and faithful representation of the real system under study, focusing on the validity of the knowledge and reliability of the results provided by the model [13, 17].

2.2. Models and Validation Types

To properly select the validation tests to be applied to a model, it is essential to first understand the nature and type of the model in question. This involves determining whether it is a correlational or causal-descriptive model, commonly known as a “black box” model [16, 18], which is constructed strictly from empirical data and does not reveal the internal structure of the system. In contrast, there are “white boxes or theoretical models that provide a transparent representation of the system, explicitly showing the underlying relationships and mechanisms [19, 20].

Note that the classification of models is not limited to these categories; other classifications are based on criteria such as physical versus symbolic, static versus dynamic, and deterministic versus stochastic.

However, for the purposes of this study, the distinction between black- and white-box models is considered particularly useful in selecting the most appropriate validation tests. The distinguishing characteristics of these two types of models are detailed in Table 2, which provides a comparative framework to facilitate the selection of appropriate validation methods.

It is important to note that validation strategies vary considerably depending on the type of model. To explore in detail the most representative validation tests applicable to black box models, relevant sources can be consulted [19-21]. These tests primarily focus on validating a model by comparing its output with the observed output of the real system within a given margin of accuracy. However, as mentioned above, these types of tests do not assess the validity of the internal relationships of the model because they are limited to aspects of accuracy that can be verified by conventional statistical tests.

The following section presents the methodology used to validate the white-box models developed using the SD. This modeling technique is particularly valuable because it allows the investigation of the underlying causes of the observed behavior when describing the internal structure of complex systems.

Such systems, which are characterized by non-linear relationships between their main variables, the presence of delays, and links formed by feedback loops, benefit from the ability of SD to model and analyze their behavior [5, 22].

Table 2. Description of the white box and black box models that address the validation aspects

Model	Features	Validation issues
Black box	This category includes models built primarily for forecasting (e.g. regression models, time series). The characteristic of these models is that they are strictly correlational, which means that there are no claims about causality in the model structure, as the interest is only in the aggregate output behavior of the model.	The commonly used and accepted tests are statistical significance tests, where a model is considered valid if its output matches the observed output of the real system within a certain range of precision. This type of test does not question the validity of the relationships present in the model, but only the precision aspects.
White box	In this category, we have models that are created with the purpose of providing an understanding of how real systems work and behave in relation to some aspect of interest (examples of models of this type are system dynamics models and all design-oriented models in general). The characteristic of these models is that they provide theories about the real system, not just reproducing or predicting the behavior, but providing an understanding of the causes of the behavior and suggesting ways to change the existing behavior.	The tests generally applied include both the evaluation of the validity of the internal structure of the model and the evaluation of the behavior described by the model output. In this sense, a model is considered valid if it first successfully passes the validity tests of the internal structure of the model, and only if it successfully passes these tests is the behavior evaluated. For this type of model, it is considered that generating accurate output behavior is not sufficient to assess the validity of the model, so statistical significance tests alone are not sufficient.

Source: Adopted from [16, 18].

3. Methods

3.1. Definition of the Research Question

The methodology of this literature review was organized into key steps inspired by robust methodological guidelines, such as those proposed in [23, 24]. The first and most critical step was to develop a clear and precise research question: What are the best practices and methods for validating system dynamics models? This question not only defined the purpose of the study, but also acted as a fundamental axis that guided the entire process of searching and analyzing the existing literature, ensuring a coherent and systematic approach at each stage of the work.

3.2. Literature Search Strategy

Based on this question, a comprehensive search was conducted in high-impact academic databases such as Scopus and Web of Science, as well as in leading publishers such as Wiley, which hosts key publications such as the Journal of System Dynamics Review. Specific terms and an advanced search string were used to ensure the relevance and accuracy of the results: “system dynamics” AND (validation OR (validation AND models) were used in the titles of the papers.

3.3. Inclusion and Exclusion Criteria

3.3.1. Inclusion Criteria

Only articles written in English that explicitly addressed the validation of the system dynamics models were included. There were no restrictions on the publication date or publication type (articles, conferences, and reviews), which allowed for a broad and comprehensive collection of information.

3.3.2. Exclusion Criteria

Studies that did not directly address the validation of system dynamics models or those that dealt with tangential aspects without specific relevance to the objective of the review were excluded.

3.4. Analysis and Synthesis of the Findings

Once the relevant articles were selected, a detailed analysis was conducted to identify recurring patterns, methodologies used, and gaps in the existing literature. This phase involves the extraction and comparison of relevant data, facilitating a comprehensive synthesis of current knowledge and highlighting areas for future research.

3.5. Organization and Discussion of the Information

The information obtained was organized logically and coherently within the article, highlighting the implications of the findings and providing a framework for further study. The importance of continuing to explore and deepen the validation practices applied to system dynamics models was emphasized, thus promoting a more robust and informed academic discussion on the subject.

4. Results and Discussion

This section provides a general overview of publications on the validation of system dynamics models. Next, the main advances in the field are reviewed, highlighting the efforts of the academic and research communities. Finally, key issues and their implications in the validation process are discussed.

4.1 Overview of Publications on System Dynamics Model Validation

Rigorous application of the eligibility criteria and implementation of the search equation in academic databases resulted in the identification of 41 articles (22 in Scopus, 6 in Web of Science, and 13 in Wiley). Of these, 27 were discarded after a thorough analysis of the titles and abstracts as they did not directly address the validation of the system dynamics models. The final selection included 14 articles that met the inclusion criteria. These studies are detailed in Table 3, where they are presented together with relevant information such as quality indicators and characteristics of the journals in which they were published.

The data in Table 3 provides key insights into the disciplinary focus of published articles, shedding light on the dominant fields of knowledge from which these studies originate. There is a notable concentration in the field of *Modeling and Simulation*, which accounts for 35.7% of publications, underlining the central role of this field in the research landscape studied. *Computer Science* was followed by 21.4%, indicating a significant contribution from computational advances and innovations. *Engineering* contributed 14.2%, further emphasizing the applied engineering approaches prevalent in these studies. Meanwhile, *Applied Mathematics, Decision Sciences*, *Medicine* and *Biochemistry, Genetics and Molecular Biology* accounted for 7.1%, underlining their targeted but central role. Taken together, these distributions reveal a landscape characterized by interdisciplinary synergies with a clear bias toward fields focused on simulation, computation, and engineering problem solving, which together shape the research direction and application in the context of this dataset.

Table 3. A comprehensive overview of studies on the validation of system dynamics models (compiled by the authors)

Author	Research Areas	Validation of the system dynamics models	Methodologies Employed	Key findings Contributions
[32]	System Dynamics and Simulation Models	Cross-validation modeling framework based on system dynamics (CVMFPS)	Quantitative, Qualitative	Training and validation sets to test the generalization ability of the model and estimate its real-world performance
[33]	Systems Thinking and Model Analysis	Fit (R2), mean square error (MSE) and Theil statistic	fsQCA (Fuzzy Qualitative Comparative Analysis)	Epidemic prevention and control strategies, specifically for the COVID-19 pandemic
[34]	Policy Modeling and System Design	Causal relationships in the system dynamics models	Comparative Analysis	Increase the validity of the system dynamics model and stakeholder confidence
[18]	System Dynamics and Simulation Models	Testing for phase lags between the model and the real system behavior using the cross-correlation function	fsQCA (Fuzzy Qualitative Comparative Analysis)	“Discrepancy Coefficient” as a final summary measure, but only after the model has passed the previous tests
[16]	System Dynamics and Simulation Models	Formal aspects of validation, with an emphasis on the importance of structural validity	Quantitative, Qualitative	Direct structure tests, structure-oriented behavior tests, and behavior pattern tests
[30]	System Dynamics and Simulation Models	System dynamics modeling to assess the impact of renewable energy systems and energy efficiency	Quantitative, Qualitative	Dependency and unit consistency tests to check the relationships between the parameters and their units
[13]	System Dynamics and Simulation Models	Time-course data with complex error structures	Quantitative, Qualitative	Developing a validation model that accounts for complex error structures in the time-course data, such as heteroscedasticity, correlations between objects, and time dependence
[29]	System Dynamics and Simulation Models	Structural and behavioral validation, focusing on Forrester and Senge	Comparative Analysis	Discussing the simulation results with the clients revealed surprises and led to the incorporation of supply-side factors that were initially overlooked
[31]	Policy Modeling and System Design	Validation of a system dynamics model for cybersecurity	Comparative Analysis	Comparing the generated behavior to the observed behavior of the real system.

[27]	System Dynamics and Simulation Models	Use of partial-model testing as a validation tool	Quantitative, Qualitative	The model is constructed and parameterized based on detailed information about the system components, rather than just fitting aggregate time-series data.
[35]	System Dynamics and Simulation Models	Novel machine learning approach based on graph neural networks (GNN) for full-field structural dynamics prediction	Quantitative, Qualitative Comparative Analysis	Inherent generalizability and can produce satisfactory prediction even when all inputs are sampled outside the training distribution
[25]	System Dynamics and Simulation Models	System dynamics models constructed for fee-paying clients as opposed to for academic research purposes	fsQCA (Fuzzy Qualitative Comparative Analysis)	Operational reality in a consultancy firm with a set of applicable criteria
[26]	System Dynamics and Simulation Models	Build confidence in the simulation model regardless of how well the model passes the behavior Validity tests	Comparative Analysis	Increased appeal for the simulation models for policy analysis and design
[28]	System Dynamics and Simulation Models	Validity tests such as Boundary adequacy; Structure verification; Dimensional consistency; Parameter verification	Comparative Analysis	Validation ideas and testing procedures can be applied to a wider variety of simulation models, agent-based models like SD models are causal models

4.2 Importance of Model Validation in system dynamics

The data presented in Table 3 reveal several key aspects related to the validation of system dynamics models, particularly within the context of academic research and business consulting.

Model validation is of paramount importance for instilling confidence in the model’s capacity to mirror the actual system and furnish valuable insights. The validation process encompasses an examination of both the structure and behavior of the model [25,27,31].

The validation of SD models is a crucial process that serves to substantiate their credibility and utility in both academic and applied contexts, such as business consulting. Existing literature emphasizes that this process requires the implementation of a range of validation tests at different stages of the modeling and simulation lifecycle. Figure 2 synthesizes the most significant aspects of validation identified in the literature, thereby providing a comprehensive framework that supports the methodological rigor required in this domain.

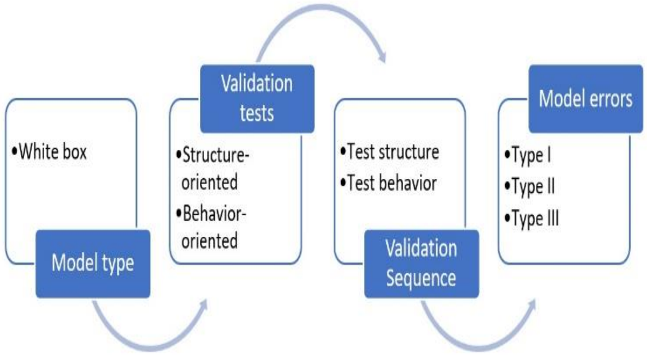

Figure 2. A comprehensive framework of key issues for the validation of system dynamics models (authors’ design)

The test results reported in Table 4 serve not only to guarantee the structural soundness of the model but also its behavioral accuracy. This instills confidence in their capacity to replicate real-world systems and generate actionable insights. This framework provides a foundation for the subsequent discussion, aligning with the findings presented in Table 3, to highlight the integral role of validation in achieving robust and reliable models.

Table 4. Validation tests for the SD simulation models (compiled by the authors)

Test name	Purpose of the test
Boundary adequacy	This test assesses the suitability of the model’s boundaries in relation to the stated objectives. The process starts with the identification of the exogenous variables in the model equations, which enables the current boundary of the system to be established [29]. This boundary is reviewed through a multi-faceted approach that includes interviews with key stakeholders and external experts, a thorough analysis of relevant literature and archival materials, and direct experience with the

	modeled system. This comprehensive approach ensures an accurate allocation of endogenous and exogenous variables, eliminating potential inconsistencies that could compromise the validity of the model and its ability to adequately reflect the real system [36, 37].
Structure verification	The objective of structure verification is to determine whether the components of the model, including causal relationships, boundaries and parameters, accurately represent the system under study [27]. This process entails a direct comparison between the model structure and the available knowledge about the system, whether theoretical or empirical. In the theoretical case, the model structure is evaluated in accordance with the established principles, physical realities, and fundamental laws that govern the system [25]. Empirical validation, on the other hand, entails assessing the model structure against quantitative or qualitative data obtained directly from the observed system. This comprehensive approach guarantees that the model is not only consistent with theoretical foundations but also with empirical evidence, thereby enhancing its capacity to accurately reflect the dynamics of the real system [29].
Dimensional consistency	The dimensional consistency test represents a fundamental step in the validation process, as it serves to guarantee that the units of measurement of each variable are consistent on both sides of all model equations [35]. Although this is a fundamental test, its importance is beyond question, as it provides an indispensable basis for ensuring the mathematical and conceptual integrity of the model. Therefore, it is recommended that this test be carried out at the outset of the validation process, serving as a preliminary filter to identify and correct possible errors before moving on to more complex analyses [37].
Parameter verification	It is of the utmost importance to guarantee the precision of the parameter verification, as this is a fundamental aspect that determines the validity and predictive capacity of the model. This process must be based on reliable empirical data or, alternatively, on expert knowledge. Parameter estimation can be conducted through formal statistical methods, utilizing available numerical data, or through qualitative approaches employing expert judgment [16]. In the latter case, interviews, workshops, analysis of archival materials, direct experience with the system, and other techniques can be employed. These combined approaches ensure that the model parameters accurately reflect the dynamics of the real system, thereby enhancing its robustness and applicability [29].
Extreme condition	The objective of this test is to assign extreme values to a selection of parameters within the model. A comparison is made between the behavior generated by the model and the observed behavior of the system in reality, under the same conditions that are considered extreme. The test is deemed to have been passed when the model demonstrates realistic behavior, irrespective of the degree of extremity of the imposed conditions. For example, inputs may be taken to have extreme values, such as zero or infinity [28].
Integration error	The objective of the integration error test is to determine whether the results produced by the model are susceptible to alterations in the numerical integration method and the time step. This is the initial simulation test that must be applied, as failure at this stage would render all subsequent model results meaningless. The test is deemed to have been passed when the model results are demonstrated to be insensitive to the choice of time step or integration method [38].
Reproduction behavior	The objective of this test is to determine whether the model in question is capable of reproducing the system behavior that is of interest. This is essentially achieved by comparing the behavioral patterns derived from the meticulous observation of the actual system under investigation with the behavior exhibited by the simulation model. Empirical data of the system are required, and the comparison is performed using descriptive statistics to evaluate the point-by-point fit. The objective of the test was to identify deficiencies in the structure or parameters of the model and to assess their significance in relation to the intended purpose. The focus is not on demonstrating that the model also fits the data, but rather on demonstrating that the model accurately reflects the data [29-31].
Behavioral abnormalities	The objective of these tests is to determine the significance of the specific structural elements within the model. This is achieved by examining whether anomalous behavior emerges when a relationship (being examined) is removed or modified. The anomalous behavior generated by the elimination of a relationship provides some evidence of its importance. In the absence of comprehensive data, it is challenging to determine the strength of significant relationships or the formulation of the model through statistical means. This type of test is instrumental in addressing this challenge [26, 28, 30].
Surprising behavior	The objective of the surprise behavior test is to determine whether the model is capable of generating a specific behavior that has been observed in the real system but has not been previously identified by the stakeholders. In such instances, the model is deemed to have passed the test. This test necessitates a comprehensive examination of the model's behavior, encompassing the observation of all variables, not merely the principal ones [32-34].
Sensitivity analysis	The objective of this test is to determine whether the actual system will demonstrate comparable sensitivity to the behavior generated by the model when parameters for which the model is highly sensitive are manipulated. A pass result for the model indicates that the observed behavior in the simulations accurately reflects that of the real system. The test can be conducted if data on the

	behavior of a modified version of the real system is available [13-29].
System improvement	These tests are the most challenging to conduct because they aim to establish the impact of the model, specifically whether the model and the policies it enables to develop enhance the system and/or address the problem. The evaluation of the aforementioned criteria is a complex process that is contingent on the specific circumstances of the case in question. To conduct an effective evaluation, it is essential to design the intervention as an experiment, delineating from the outset the evaluations that will be conducted throughout the process and the methodology for the management of all pertinent information [36-38].
Turing test	In this test, experts are presented with a random collection of behaviors from the real and simulated output. If the experts detect significant differences in behavior, the model fails the test [25].

4.3 Recommended Sequence for the Validation Tests in the SD Models

Table 4 presents a range of recommended validation tests for evaluating the simulation models developed using System Dynamics. The application of these tests follows the generally recommended logical sequence. The process begins with an assessment of the validity of the model’s structure, which should be conducted before any evaluation of its behavioral accuracy. Behavioral testing should only be conducted if the model successfully passes the structural tests. This approach ensures that structural errors do not lead to models that can replicate system behaviors with high precision, despite the underlying flaws [15, 16, 18]. Structural tests can be further categorized into two types.

4.3.1 Direct Structure Tests

Direct structural tests are employed to evaluate the structural validity of the model by directly comparing it to the known structure of a real-world system. These tests are conceptually rich, qualitative, and often involve a less formal approach.

4.3.2 Structure-Oriented Behavior Tests

These tests employed rigorous numerical experimentation (simulation) to indirectly assess the structure of the model. The results yield quantitative insights, thereby facilitating a more formal evaluation and enabling precise measurement of the findings.

4.3.3 Logical Sequence for Applying the Validation Tests

Figure 3 shows a more detailed logical sequence for the implementation of the validation tests. It is of utmost importance to highlight that modeling is an iterative feedback process rather than a linear progression. These models are subject to continuous refinement, questioning, validation, and adjustment. As illustrated in Figure 3, the modeling process encompasses multiple feedback loops triggered by the outcomes of the model reviews. Progress is achieved when a review provides sufficient confidence in the model. For a comprehensive explanation of the application of each test, refer to the bibliographic sources used in this section.

4.4 Key Errors in the SD Models

The creation of a simulation model may result in three principal categories of error:

4.4.1 Type I Error

This error was defined as the rejection of a credible model based on biased or inaccurate observations during its evaluation. This occurs when a model with sufficient credibility is incorrectly rejected, often because of biased or inaccurate observations during the evaluation process.

4.4.2 Type II Error

The converse situation, in which an untrustworthy model is erroneously accepted as credible, frequently arises from flawed or biased observations made during the model-development process.

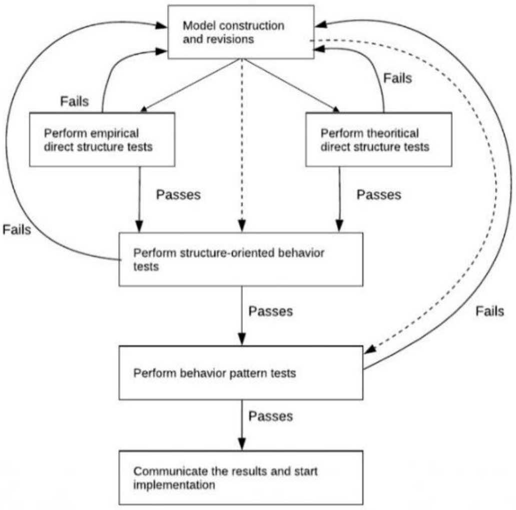

Figure 3. Logical sequence of the formal validation steps for the SD models. Source: Adopted from [18].

4.4.3 Type III error

This occurs when the model addresses an entirely inappropriate problem [19, 20, 39]. As illustrated in Figure 4, these errors can occur at various stages of the modeling process, with each stage carrying distinct costs and consequences.

To illustrate, Type I errors result in increased model development costs due to delayed acceptance.

The impact of Type II and III errors is contingent on the manner in which the flawed model is utilized and the decisions it provides. In particular, Type III errors render the model completely irrelevant [19, 20].

The successful conclusion of the modeling process is contingent on the model, and its results are deemed credible by relevant stakeholders [11]. This outcome is contingent on the establishment of trust in the model, accompanied by an acknowledgment of its specific utility in accordance with its intended purpose [15, 18]. Trust is not a statistical metric; rather, it is the confidence that stakeholders have in the model and its outcomes [40, 41].

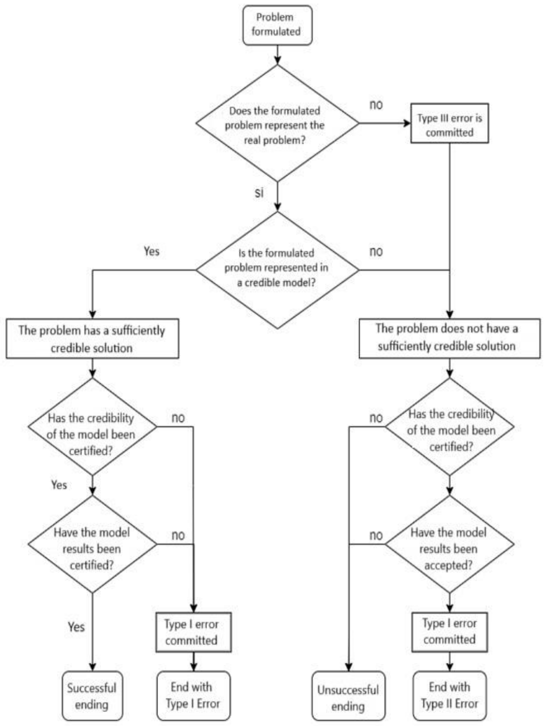

Figure 4. Errors in the modeling and simulation process. Adopted from [16, 20].

It is important to recognize that the development of a simulation model carries the risk of making Type I, II and III errors, which can affect its credibility and usefulness. It is imperative that stakeholders accept the model because their trust is contingent on their ability to address the problem at hand. Therefore, it is essential to subject the data and results to a rigorous evaluation to guarantee the credibility of the model and its relevance to the stated objectives. It is of paramount importance that the field continues to adopt and adapt modern data science and AI techniques throughout the modeling process, as these could transform the way complex problems are approached and models developed [24].

4.5 The Limitations of Model Validation in SD

The process of model validation in SD, particularly when applied to white-box or descriptive-causal models, is inherently constrained by several limitations that must be acknowledged and addressed [16, 23].

4.5.1 Philosophical Limitations

The subjective nature of the validation is a key consideration in this context. The validity of a model is inextricably linked to its purpose, which, in turn, is subject to individual interpretations and perspectives. The relativistic/holistic philosophical perspective posits that the validation of the internal structure of a model cannot be entirely objective, formal, or quantitative [16, 25, 26]. This is because the confirmation of a scientific theory inherently involves both informal and subjective aspects.

Validating the internal structure of a model presents a significant challenge. The question of whether the internal structure of a model accurately represents reality presents a significant challenge from both the philosophical and technical perspectives. It is not possible to definitively determine how close the structure of a model is to the “real” structure using established formal tests, such as statistical hypothesis tests [16, 18].

4.5.2 Technical Limitations

The complexity of real systems is a significant challenge in the field of system engineering. It is inherent to the nature of these models that they are simplifications of reality. Consequently, certain details and complexities are inevitably omitted. The selection of the model boundary, that is, the determination of what is included and what is excluded, is of great consequence and has the potential to exert a significant influence on the validation results [25-27].

It is challenging to capture qualitative relationships. Tests that directly compare the structure of models with real-world systems often rely on highly qualitative information, which is difficult to quantify.

The limitations of statistical tests must also be considered. It is questionable whether standard statistical tests, such as significance tests, are appropriate for validating the behavior of system dynamics models [16, 25]. This is because of problems such as autocorrelation and multicollinearity. The stochastic nature of real-world data and the difficulty of accurately estimating noise terms in system dynamics models also present significant challenges for behavioral validation [30, 31].

4.5.3 Practical Limitations

The issue of data availability is significant in this context. The validation of models, particularly those based on behavioral data, frequently necessitates the availability of extensive and reliable historical data. However, this may not always be forthcoming, particularly in the case of complex social and economic systems [16, 18].

Time and resource constraints represent significant challenges in this regard. The implementation of a comprehensive set of validation tests is often financially and temporally prohibitive, which can restrict the extent of testing that can be conducted in practice [27-29].

4.5.4 Specific limitations of the domain in question

The process of validating models used to inform policy decision-making presents several distinctive challenges, particularly within a specific domain. Often, the controversial nature of policy issues, multiplicity of stakeholder perspectives, and difficulty of conducting controlled experiments in the real world combine to render the process of validation particularly complex [16, 23].

4.5.5 Mitigation of the Limitations

Although the limitations of model validation are intrinsic, there are methods that can be employed to mitigate these limitations and enhance confidence in the SD models.

It is essential to prioritize structural validity. The prioritization of structural validity tests, such as structure checking and parameter confirmation, can facilitate assurance that the model is firmly grounded in domain knowledge and theory [16, 18].

The triangulation of methods is another means of ensuring the rigor of the process. The use of a combination of quantitative and qualitative tests, including expert reviews, sensitivity analysis, and testing under extreme conditions, can facilitate a more comprehensive assessment of model validity.

Transparent communication is essential for ensuring that all parties involved in the process are kept fully informed and up-to-date with the latest developments. It is beneficial to communicate the assumptions, limitations, and results of the validation process to relevant stakeholders in a clear and transparent manner, as this can foster trust and enhance the overall transparency of the process [31-33].

Stakeholder involvement is a crucial aspect of the validation process. Involving stakeholders in the validation process can facilitate the alignment of the model with their needs and representation of their understanding of the system [16, 18].

4.6 A Proposed Research Agenda on the Validation of Models in the Context of SD

4.6.1 Development of Formal Methods for the Testing of Behaviorally Oriented Structures

Testing, in which a model is subjected to extreme or atypical conditions to assess whether its behavior is consistent with domain knowledge and remains plausible, is a powerful tool for uncovering structural weaknesses. To advance this approach, further research is required to formalize and quantify such testing procedures. This could include the development of metrics to assess model plausibility under extreme scenarios, or the development of algorithms to automate the generation of test sets for such conditions. Embedding these behavior-focused structural testing methods in simulation software can improve their accessibility and usability, thus streamlining their application for model developers [16, 18].

4.6.2 Research into the Validation of Models for Specific Purposes

Models designed to support organizational learning often require different validation approaches to those developed for forecasting or decision making, reflecting their unique objectives and application contexts. Similarly, consultants using system dynamics models in commercial settings face specific challenges, including the need for rigorous formal testing and the need to align model development with client expectations and practical limitations [19, 20].

4.6.3 Enhanced Software Tools for Validation

The current generation of simulation software offers only basic capabilities for model validation, leaving a considerable scope for enhancement. The creation of integrated validation environments that provide systematic guidance to users throughout the validation process, including structural testing, behavioral evaluation, and sensitivity analysis, would represent a significant advancement in this field. Furthermore, the development of composite validation summary measures that would distill the outcomes of multiple tests into clear, informative, and user-friendly formats would significantly enhance the accessibility and interpretability of the validation results [21, 22].

4.6.4 Empirical Research on the Effectiveness of the Validation Methods

Significant progress has been made in controlled studies in this domain. However, there remains a pressing need for rigorous experimental research to evaluate the effectiveness of diverse validation techniques for uncovering model deficiencies and improving model robustness. It is imperative that such studies be conducted to establish an empirical foundation for best practices, offering actionable insights into the comparative strengths and limitations of various approaches [16, 18]. In addition, well-documented case studies that exemplify optimal validation practices can serve as crucial resources for practitioners. By exemplifying the successful application of effective validation in real-world scenarios, these case studies can bridge the gap between theoretical advancements and practical implementation, fostering a deeper understanding of how to ensure model reliability across diverse applications [24, 25].

5. Conclusion

5.1. Research Summary

The study of complex systems emphasizes the critical importance of comprehending the intricate interplay between numerous interacting elements. These interactions frequently result in emergent behaviors that elude immediate comprehension. In this context, validation of the models assumes a pivotal role in determining the credibility and utility of the results. This study addresses the observed dearth of attention given to validation within the field of System Dynamics (SD) modeling. A comprehensive review of the literature was conducted to investigate the evolution of validation practices in SD and identify key trends, challenges, and best practices.

A significant contribution of this study was the development of a novel conceptual framework for SD model validation. This framework, synthesized from the existing literature, provides researchers with a structured guide for conducting more rigorous and comprehensive validation analyses, incorporating a broader range of techniques and considerations. This framework represents a substantial advancement in the field, offering a valuable resource for enhancing the credibility and reliability of SD models and, ultimately, improving confidence in their application for decision-making and policy analysis.

5.2. Recommendations

Considering the findings of this study, the following key recommendations can be formulated: i) Validation should be prioritized. Researchers should prioritize model validation at every stage of the SD modeling process, from the initial conceptualization of the model to its final implementation and analysis. ii) A multi-method approach should be embraced: a combination of qualitative, quantitative, and mixed-methods techniques should be employed to ensure a comprehensive and robust validation process. iii) Collaboration should be fostered: collaboration between modelers, domain experts, and stakeholders is crucial for effective model validation. iv) The utilization of available tools and resources is also paramount. Researchers should leverage existing tools and resources such as software packages and online communities to support their validation efforts.

5.3. Research Perspectives

This study suggests several avenues for future research:

- The development and refinement of advanced validation techniques are imperative. Further research is required to develop and refine novel validation techniques such as those incorporating artificial intelligence and machine learning.
- A further research objective is to investigate the impact of different validation approaches on the model performance. Empirical studies are required to investigate the impact of different validation approaches on the model accuracy, robustness, and decision-making effectiveness.
- The development of standardized validation guidelines is also recommended. The development of such guidelines for SD model validation can enhance the consistency and rigor across diverse studies.
- Finally, there is a need to explore the ethical implications of model validation. Research is needed to explore the ethical considerations associated with model validation, such as the potential for bias and misuse of the validated models.

Declarations

Author Contributions

Conceptualization, A.R.P., L.M.C.A. and J.A.P.V.; methodology, A.R.P., L.M.C.A. and J.A.P.V.; software, A.R.P., L.M.C.A. and J.A.P.V.; validation, A.R.P., L.M.C.A. and J.A.P.V.; formal analysis, A.R.P., L.M.C.A.; J.A.P.V. A.J.L.D. and J.G.P.C.; investigation, A.R.P., L.M.C.A.; J.A.P.V. A.J.L.D. and J.G.P.C.; resources, A.R.P.; data curation, A.R.P.; writing—original draft preparation, A.R.P., L.M.C.A. and J.A.P.V.; writing—review and editing, A.R.P., L.M.C.A.; J.A.P.V. A.J.L.D. and J.G.P.C.; visualization, A.R.P.; supervision, A.R.P.; project administration, A.R.P.; funding acquisition, A.R.P. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

Funding information is not available.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this manuscript. In addition, ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies, have been completely observed by the authors.

References

[1] YOON S, ANDERSON E, KLOPFER E, KOEHLER-YOM J, SHELDON J, SCHOENFELD I, … & GOH SE. Designing Computer-Supported Complex Systems Curricula for the Next Generation Science Standards in High School Science Classrooms. *Systems*, 2016, 4: 38. https://doi.org/10.3390/systems4040038

[2] BARELLI E, & LEVRINI O. Computational simulations

at the interface of physics and society: A teaching-learning module for high-school students. *IL Nuovo Cimento. Communications: SIF Congress*, 2022, 45(6): 1–4. http://dx.doi.org/10.1393/ncc/i2022-22213-6

[3] YOON SA. An Evolutionary Approach to Harnessing Complex Systems Thinking in the Science and Technology Classroom. *International Journal of Science Education*, 2008, 30: 1–32. https://doi.org/10.1080/09500690601101672

[4] HELBIG N, CRESSWELL AN, BURKE GB, & LUNA-REYES L. *The dynamics of opening government data*. Albany, New York: 2012.

[5] NAJAFABADI MM, & LUNA-REYES LF. Open Government Data Ecosystems: A Closed-Loop Perspective. *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017, pp. 2711-2720.

[6] ARROYO MENÉNDEZ M, & HASSAN COLLADO S. Simulación de procesos sociales basada en agentes software. Empiria *Revista de Metodología de Ciencias Sociales*, 2007. 14: 139-161. https://doi.org/10.5944/empiria.14.2007.1175

[7] CASABOZA J, & CÁRDENAS DE. Analysis and Modeling of Dynamic Behavior of the COVID-19 Outbreak: Study Case of Panama. *IEEE Latin America Transactions*, 2021, 19: 893–900. https://doi.org/10.1109/TLA.2021.9451233

[8] FENNER G, LIMA A, DE SOUZA J, MOURA JAB, & BEZERRA TR. Supporting Infrastructure as a Service Capacity Management through Business Scenarios Simulation. *IEEE Latin America Transactions*, 2020, 18: 473–486. https://latamt.ieeer9.org/index.php/transactions/article/view/210

[9] PARHIZKAR T, & MOSLEH A. Guided Probabilistic Simulation of Complex Systems Toward Rare and Extreme Events. *2022 Annual Reliability and Maintainability Symposium (RAMS), IEEE*; 2022, pp. 1–7. https://doi.org/10.1109/RAMS51457.2022.9893976

[10] NCUBE C, & LIM SL. On Systems of Systems Engineering: A Requirements Engineering Perspective and Research Agenda. *Proceedings of the 2018 IEEE 26th International Requirements Engineering Conference (RE), IEEE;* 2018, pp. 112–123. https://doi.org/10.1109/RE.2018.00021

[11] CHATTOPADHYAY S, ROY T, SENGUPTA S, & BERGER-VACHON C. *Modelling and Simulation in Science, Technology and Engineering Mathematics*. vol. 749. 1st ed. Cham: Springer International Publishing, 2019. https://doi.org/10.1007/978-3-319-74808-5

[12] MALDONADO GRANADOS LF. *El modelamiento matemático en la formación del ingeniero*. 1st ed. Bogotá DC., Colombia: Universidad Central, 2013. https://www.ucentral.edu.co/sites/default/files/inline-files/2015_modelamiento_matematico_001.pdf.

[13] KIM Y, & KIM J. Model validation in dynamic systems for time-course data with complex error structures. *Journal of Chemometrics*, 2019, 33(5): e3108. https://doi.org/10.1002/cem.3108

[14] GARCÍA-VALDECASAS MEDINA JI. Explicación, mecanismo y simulación: otra manera de hacer sociología. *Empiria Revista de Metodología de Ciencias Sociales*, 2014, 28: 35. https://doi.org/10.5944/empiria.28.2014.12120

[15] BARLAS Y, & CARPENTER S. Philosophical roots of model validation: Two paradigms. *System Dynamics Review*, 1990, 6(2): 148–166. https://doi.org/10.1002/sdr.4260060203

[16] BARLAS Y. Formal aspects of model validity and validation in system dynamics. *System Dynamics Review*, 1996, 12: 183–210. https://doi.org/10.1002/(SICI)1099-1727(199623)12:3%3C183::AID-SDR103%3E3.0.CO;2-4

[17] SCHOENBERG W, & SWARTZ J. System Dynamics, Machine Learning and Structural Validation. *LIFE: A Transdisciplinary Inquiry*, Intellect Books/University of Chicago Press, 2024, pp. 289–308.

[18] BARLAS Y. Multiple tests for validation of system dynamics type of simulation models. *European Journal of Operational Research*, 1989, 42: 59–87. https://doi.org/10.1016/0377-2217(89)90059-3

[19] BALCI O. Principles and techniques of simulation validation, verification, and testing. *1995 Winter Simulation Conference Proceedings, IEEE*, 1995, pp. 147–154. https://doi.org/10.1109/WSC.1995.478717

[20] BALCI O. Validation, verification, and testing techniques throughout the life cycle of a simulation study. *Annals of Operational Research*, 1994, 53: 121–173. https://doi.org/10.1007/BF02136828

[21] XIANG X, KENNEDY RC, MADEY GR, & CABANISS SE. Verification and Validation of Agent-based Scientific Simulation Models. *Proceedings of the Agent-Directed Simulation Conference*, 2005, pp. 47-55. https://www3.nd.edu/~nom/Papers/ADS019_Xiang.pdf.

[22] SCHOLL HJ (JOCHEN), & LUNA-REYES LF. Transparency and openness in government: a system dynamics perspective. *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance*, New York, NY, USA: Association for Computing Machinery; 2011, pp. 107–114. https://doi.org/10.1145/2072069.2072088

[23] LUTZ CS, CARR W, COHN A, & RODRIGUEZ L. Understanding barriers and predictors of maternal immunization: Identifying gaps through an exploratory literature review. *Vaccine*, 2018, 36: 7445–7455. https://doi.org/10.1016/j.vaccine.2018.10.046

[24] NAUGLE A, LANGARUDI S, & CLANCY T. What is (quantitative) system dynamics modeling? Defining characteristics and the opportunities they create. *System Dynamics Review*, 2024, 40(2): e1762. https://doi.org/10.1002/sdr.1762

[25] COYLE G, & EXELBY D. The validation of commercial system dynamics models. *System Dynamics Review*, 2000, 16: 27–41. https://doi.org/10.1002/(SICI)1099-1727(200021)16:1<27::AID-SDR182>3.0.CO;2-1

[26] QUDRAT-ULLAH H. Structural validation of system dynamics and agent-based simulation models. *Simulation in Wider Europe - 19th European Conference on Modelling and Simulation*, ECMS, 2005. https://scs-europe.net/services/ecms2005/pdf/abs-07.pdf.

[27] HOMER JB. Partial-model testing as a validation tool for system dynamics (1983). System Dynamics Review, 2012, 28: 281–294. https://doi.org/10.1002/sdr.1478

[28] QUDRAT-ULLAH H. On the validation of system dynamics type simulation models. *Telecommunication Systems*, 2012, 51: 159–166. https://doi.org/10.1007/s11235-011-9425-4

[29] VAN NISTELROOIJ LPJ, ROUWETTE EAJA,

VERSTIJNEN IM, & VENNIX JAM. The Eye of the Beholder: A Case Example of Changing Clients' Perspectives Through Involvement in the Model Validation Process. *Systems Research and Behavioral Science*, 2015, 32: 437–449. https://doi.org/10.1002/sres.2336

[30] LAIMON M, MAI T, GOH S, & YUSAF T. System dynamics modelling to assess the impact of renewable energy systems and energy efficiency on the performance of the energy sector. *Renewable Energy*, 2022, 193: 1041–1048. https://doi.org/10.1016/j.renene.2022.05.041

[31] KANNAN U, & SWAMIDURAI R. Empirical Validation of System Dynamics Cyber Security Models. *Proceedings of the 2019 SoutheastCon, IEEE*, 2019, pp. 1–6. https://doi.org/10.1109/SoutheastCon42311.2019.9020607

[32] HUANG H, CHEN W, XIE T, WEI Y, FENG Z, & WU W. The Impact of Individual Behaviors and Governmental Guidance Measures on Pandemic-Triggered Public Sentiment Based on System Dynamics and Cross-Validation. *International Journal of Environmental Research and Public Health*, 2021, 18: 4245. https://doi.org/10.3390/ijerph18084245

[33] JIA S., LI Y. & FANG T. System dynamics analysis of COVID-19 prevention and control strategies. *Environmental Science and Pollution Research*, 2022, 29: 3944–3957. https://doi.org/10.1007/s11356-021-15902-2

[34] SHALAHUDDIN M, SUNINDYO WD, EFFENDI MR, & SURENDRO K. Fuzzy-set qualitative comparative analysis for validating causal relationships in system dynamics models. *Engineering Reports*, 2024, 6: e12855. https://doi.org/10.1002/eng2.12855

[35] LI Q, WANG Z, LI L, HAO H, CHEN W, & SHAO Y. Machine learning prediction of structural dynamic responses using graph neural networks. *Computers and Structures*, 2023, 289: 107188. https://doi.org/10.1016/j.compstruc.2023.107188

[36] FORRESTER JW. System dynamics, systems thinking, and soft OR. *System Dynamics Review*, 1994, 10: 245–256. https://doi.org/10.1002/sdr.4260100211

[37] STERMAN J. *System Dynamics: Systems Thinking and Modeling for a Complex World*. Cambridge MA: 2002. https://archive.org/details/businessdynamics0000ster

[38] BELETE GF, VOINOV A, & LANIAK GF. An overview of the model integration process: From pre-integration assessment to testing. *Environmental Modelling & Software*, 2017, 87: 49–63. https://doi.org/10.1016/j.envsoft.2016.10.013

[39] SCHRUBEN LW. Establishing the credibility of simulations. *Simulation*, 1980, 34: 101–105. https://doi.org/10.1177/003754978003400310

[40] TEDESCHI LO. Assessment of the adequacy of mathematical models. *Agricultural Systems*, 2006, 89(1-2): 225–247. https://doi.org/10.1016/j.agsy.2005.11.004

[41] ROBINSON S, & BROOKS RJ. Independent Verification and Validation of an Industrial Simulation Model. *Simulation*, 2010, 86: 405–416. https://doi.org/10.1177/0037549709341582

[42] PACE DK. Verification, validation, and accreditation of simulation models. In: OBAIDAT, M.S., & PAPADIMITRIOU, G.I. (eds) Applied System Simulation. Applied System Simulation: Methodologies and Applications, Springer, Boston, MA.: Kluwer Academic Publishers; 2003, chapter 21: 487–506.

参考文:

[1] YOON S、ANDERSON E、KLOPFER E、KOEHLER-YOM J、SHELDON J、SCHOENFELD I、… 和 GOH SE。为高中科学课堂中的下一代科学标准设计计算机支持的复杂系统课程。 系统， 2016， 4： 38。https://doi.org/10.3390/systems4040038

[2] BARELLI E 和 LEVRINI O。物理与社会交界处的计算模拟：面向高中生的教学模块。IL Nuovo Cimento。通讯：SIF 大会，2022 ，45(6)：1-4。http://dx.doi.org/10.1393/ncc/i2022-22213-6

[3] YOON SA。在科学技术课堂中利用复杂系统思维的进化方法。国际科学教育杂志，2008，30：1-32。https://doi.org/10.1080/09500690601101672

[4] HELBIG N、CRESSWELL AN、BURKE GB 和 LUNA-REYES L。开放政府数据的动态。奥尔巴尼，纽约：2012。

[5] NAJAFABADI MM 和 LUNA-REYES LF。开放政府数据生态系统：闭环视角。第 50 届夏威夷国际系统科学会议论文集，2017，第 2711-2720 页。

[6] ARROYO MENÉNDEZ M 和 HASSAN COLLADO S. 社会过程模拟和代理软件。社会科学方法论的经验，200。14：139-161。https://doi.org/10.5944/empiria.14.2007.1175

[7] CASABOZA J 和 CÁRDENAS DE。 COVID-19 爆发的动态行为分析和建模：巴拿马研究案例。 IEEE 拉丁美洲学报，2021，19：893-900。https://doi.org/10.1109/TLA.2021.9451233

[8] FENNER G、LIMA A、DE SOUZA J、MOURA JAB 和 BEZERRA TR。通过业务场景模拟支持基础设施即服务容量管理。IEEE 拉丁美洲交易，2020，18：473-486. https://latamt.ieeer9.org/index.php/transactions/article/view/210

[9] PARHIZKAR T，和 MOSLEH A。针对罕见和极端事件的复杂系统引导概率模拟。2022 年可靠性和可维护性研讨会 (RAMS)，IEEE；2022，第 1-7 页。https://doi.org/10.1109/RAMS51457.2022.9893976

[10] NCUBE C，和 LIM SL。关于系统工程的系统：需求工程视角和研究议程。2018 年 IEEE 第 26 届国际需求工程会议 (RE) 论文集，IEEE； 2018，第 112-123 页。https://doi.org/10.1109/RE.2018.00021

[11] CHATTOPADHYAY S、ROY T、SENGUPTA S 和 BERGER-VACHON C。科学、技术和工程数学中的建模和仿真。卷。 749.第一版。 Cham：施普林格国际出版，2019。https://doi.org/10.1007/978-3-319-74808-5

[12] MALDONADO GRANADOS LF. 工程训练中的数学建模。第 1 版哥伦比亚波哥大：中央大学，2013。https://www.ucentral.edu.co/sites/default/files/inline-files/2015_modelamiento_matematico_001.pdf。

[13] KIM Y, 和 KIM J. 具有复杂误差结构的时程数据的动态系统模型验证。化学计量学杂志，2019，33（5）：e3108。 https://doi.org/10.1002/cem.3108

[14] ARROYO MENÉNDEZ M, 和 HASSAN COLLADO S. 基于软件代理的社会过程模拟。 Empiria 社会科学方法论杂志，2014，28：35。https://doi.org/10.5944/empiria.28.2014.12120

[15] BARLAS Y，和 CARPENTER S。模型验证的哲学根源：两种范式。系统动力学评论，1990，6（2）：148-166。https://doi.org/10.1002/sdr.4260060203

[16] BARLAS Y。系统动力学中模型有效性和验证的形式化方面。系统动力学评论，1996，12：183-210。https://doi.org/10.1002/(SICI)1099-1727(199623)12:3%3C183::AID-SDR103%3E3.0.CO;2-4

[17] SCHOENBERG W，和 SWARTZ J. 系统动力学、机器学习和结构验证。 LIFE：跨学科探究，Intellect Books/芝加哥大学出版社，2024，第 289-308 页。

[18] BARLAS Y. 用于验证系统动力学类型模拟模型的多重测试。《欧洲运筹学杂志》，1989，42：59-87。https://doi.org/10.1016/0377-2217(89)90059-3

[19] BALCI O. 模拟确认、验证和测试的原理和技术。1995 年冬季模拟会议论文集，IEEE，1995，第 147-154 页。https://doi.org/10.1109/WSC.1995.478717

[20] BALCI O. 模拟研究整个生命周期中的确认、验证和测试技术。运筹学年鉴，1994，53：121-173。https://doi.org/10.1007/BF02136828

[21] XIANG X、KENNEDY RC、MADEY GR 和 CABANISS SE。基于代理的科学模拟模型的验证和确认。代理导向模拟会议论文集，2005，第 47-55 页。https://www3.nd.edu/~nom/Papers/ADS019_Xiang.pdf。

[22] SCHOLL HJ (JOCHEN) 和 LUNA-REYES LF。政府的透明度和开放性：系统动力学视角。第五届电子治理理论与实践国际会议论文集，纽约，美国：计算机协会；2011，第 107-114 页。https://doi.org/10.1145/2072069.2072088

[23] LUTZ CS、CARR W、COHN A 和 RODRIGUEZ L。了解母体免疫的障碍和预测因素：通过探索性文献综述确定差距。疫苗，2018，36：7445‑7455。https://doi.org/10.1016/j.vaccine.2018.10.046

[24] NAUGLE A、LANGARUDI S 和 CLANCY T。什么是（定量）系统动力学建模？定义特征及其创造的机会。系统动力学评论，2024，40(2)：e1762。https://doi.org/10.1002/sdr.1762

[25] COYLE G 和 EXELBY D。商业系统动力学模型的验证。系统动力学评论，2000，16：27-41。https://doi.org/10.1002/(SICI)1099-1727(200021)16:1<27::AID-SDR182>3.0.CO;2-1

[26] QUDRAT-ULLAH H. 系统动力学和基于代理的仿真模型的结构验证。欧洲范围内的仿真 - 第 19 届欧洲建模与仿真会议，ECMS，2005。https://scs-europe.net/services/ecms2005/pdf/abs-07.pdf。

[27] HOMER JB. 部分模型测试作为系统动力学的验证工具 (1983)。 系统动力学评论，2012，28：281-294。https://doi.org/10.1002/sdr.1478

[28] QUDRAT-ULLAH H. 关于系统动力学类型模拟模型的验证。电信系统，2012，51：159‑166。https://doi.org/10.1007/s11235-011-9425-4

[29] VAN NISTELROOIJ LPJ、ROUWETTE EAJA、VERSTIJNEN IM 和 VENNIX JAM。旁观者之眼：通过参与模型验证过程改变客户观点的案例。系统研究与行为科学，2015，32：437‑49。https://doi.org/10.1002/sres.2336

[30] LAIMON M、MAI T、GOH S 和 YUSAF T。系统动力学建模以评估可再生能源系统和能源效率对能源部门绩效的影响。可再生能源，2022，193：1041-1048。https://doi.org/10.1016/j.renene.2022.05.041

[31] KANNAN U 和 SWAMIDURAI R。系统动力学网络安全模型的实证验证。2019 年 SoutheastCon 论文集，IEEE，2019，第 1-6 页。https://doi.org/10.1109/SoutheastCon42311.2019.9020607

[32] HUANG H, CHEN W, XIE T, WEI Y, FENG Z, 和 WU W. 基于系统动力学和交叉验证的个体行为和政府指导措施对疫情引发公众情绪的影响。国际环境研究与公共健康杂志，2021，18：4245。https://doi.org/10.3390/ijerph18084245

[33] JIA S., LI Y. 和 FANG T。COVID-19 防控策略的系统动力学分析。环境科学与污染研究，2022，29：3944‑3957。 https://doi.org/10.1007/s11356-021-15902-2

[34] SHALAHUDDIN M、SUNINDYO WD、EFFENDI MR 和 SURENDRO K。模糊集定性比较分析用于验证系统动力学模型中的因果关系。工程报告，2024，6：e12855。https://doi.org/10.1002/eng2.12855

[35] LI Q、WANG Z、LI L、HAO H、CHEN W 和 SHAO Y。使用图神经网络进行结构动态响应的机器学习预

测。计算机和结构，2023，289：107188。 https://doi.org/10.1016/j.compstruc.2023.107188

[36] FORRESTER JW。系统动力学、系统思维和软OR。系统动力学评论，1994，10：245-256。 https://doi.org/10.1002/sdr.4260100211

[37] STERMAN J。系统动力学：复杂世界的系统思维和建模。马萨诸塞州剑桥：2002。 https://archive.org/details/businessdynamics0000ster

[38] BELETE GF、VOINOV A 和 LANIAK GF。模型集成过程概述：从预集成评估到测试。环境建模与软件，2017，87：49-63。 https://doi.org/10.1016/j.envsoft.2016.10.013

[39] SCHRUBEN LW。建立模拟的可信度。模拟，1980，34：101-105。 https://doi.org/10.1177/003754978003400310

[40] TEDESCHI LO。评估数学模型的充分性。农业系统，2006，89(1-2)：225-247。 https://doi.org/10.1016/j.agsy.2005.11.004

[41] ROBINSON S 和 BROOKS RJ。工业模拟模型的独立验证和确认。模拟，2010，86：405-416。 https://doi.org/10.1177/0037549709341582

[42] PACE DK。模拟模型的验证、确认和认证。在：OBAIDAT，M.S.，和 PAPADIMITRIOU，G.I.（编辑）应用系统模拟。应用系统模拟：方法和应用，Springer，波士顿，马萨诸塞州：克鲁维尔学术出版社；2003，第21章：487-506。

Word count: 8,775 words, excluding references.

Peer review information:
The reviewer reports submitted in first round: 4 reports
The reviewer reports submitted in 2nd round: 2 reports
The revision rounds stages: 3 rounds
Final revised version submitted: January 17, 2025.