Probabilistic Validation Methodology for Probabilistic Risk Assessment: Overview, Status, and Future Research Directions

Ha Bui[a*], Hammad Khalid[a], Tatsuya Sakurahara[a], Seyed Reihani[a], Zahra Mohaghegh[a]

[a] Socio-Technical Risk Analysis (SoTeRiA) Laboratory, Department of Nuclear, Plasma, and Radiological Engineering, Grainger College of Engineering, University of Illinois at Urbana-Champaign, USA.

* Email: habui2@illinois.edu; riskanalysis@illinois.edu

Abstract:

The Socio-Technical Risk Analysis (SoTeRiA) Laboratory at the University of Illinois at Urbana-Champaign has been focusing on the advancement of Probabilistic Risk Assessment (PRA), pioneering four key areas of scholarly development: (1) spatiotemporal coupling of physical failure mechanisms with human/social performance and the incorporation of this coupling into PRA using the dynamic-static Integrated PRA (I-PRA) methodology; (2) incorporating Artificial Intelligence (AI), Machine Learning (ML), digital twins, and Virtual Reality (VR) into PRA; (3) Probabilistic Validation in support of PRA; and (4) integrating safety risk and financial risk in support of risk-informed decision-making. The current paper reports on the progress of key area #3, which aims at developing a systematic and scientifically justifiable validation methodology, i.e., the Probabilistic Validation (PV), to facilitate the validity evaluation of advanced simulation models that are used for PRA in support of risk-informed decision-making and regulation, especially when common empirical validation approaches become challenging due to the scarcity of empirical validation data.

In PV, the validity of a simulation prediction used for PRA is determined by: (i) the magnitude of epistemic uncertainty (i.e., the degree of confidence) in the simulation prediction, calculated using a comprehensive uncertainty analysis that quantifies and aggregates all dominant sources of uncertainty involved in the development and usage of the simulation model; and (ii) the results of an acceptability evaluation that determine whether the total uncertainty associated with the simulation prediction is acceptable for the application of interest (e.g., PRA). PV has been applied to a case study for evaluating the validity of a fire simulation model used in Fire PRA of nuclear power plants (NPPs). Later, our team has extended the PV methodology and applied it to the validation of a finite element analysis-based simulation model developed for the explicit physics-maintenance coupling in piping reliability analysis, addressing the challenge of identifying and propagating uncertainties from input parameters to reliability estimates. This approach has been recognized in IAEA TECDOC-1988 and is being further refined to account for correlated inputs in ongoing research. Additionally, PV is being extended to facilitate the validation of AI/ML-based automation systems in NPPs in another ongoing project. This approach is being demonstrated for evaluating the trustworthiness of an AI-based automated fire watch system for use in NPPs. The research progress of our PV development and key areas for future work will be discussed in the current paper.

Keywords: PRA, Probabilistic Validation (PV), Uncertainty Analysis, Modeling and Simulation.

1. INTRODUCTION

Advanced modeling and simulation (M&S) has been increasingly used in the nuclear domain to improve the realism of PRA for existing NPPs as well as to foster the analysis, design, and operationalization of advanced nuclear reactors. Prior to being used in PRA to support risk-informed decision-making, simulation models need to be adequately validated. To support this validation, some of the authors of this paper have initiated a research line to develop the Probabilistic Validation (PV) methodology [1, 2] for systematically and scientifically evaluating the validity of simulation prediction, especially when empirical validation approaches are not feasible, for instance, due to the lack of validation data at the model output level. Four key contributions of this PV line of research are highlighted below.

First, this research line performed a cross-disciplinary literature review to generate scientific justification for using epistemic uncertainty as a supporting, quantitative validity measure for validating simulation model prediction, especially when empirical validation is challenging [1]. This literature review was conducted to examine the evolution of the validation concept across various academic and professional communities. This historical analysis traced the changing perspectives on the validation concept from the early, more vague interpretations to more nuanced definitions, notably those by the U.S. Department of Defense (DOD), American Institute of Aeronautics and Astronautics (AIAA), and American Society of Mechanical Engineers (ASME). The review critically assessed various approaches to validation, such as the ASME standards (ASME V&V 20-2009, ASME V&V 10-2019, ASME V&V 10.1-2012) for computational fluid dynamics and solid mechanics, as well as the Probability Bound Analysis and Bayesian approaches to verification, validation, and uncertainty quantification (VVUQ). Through this analysis, the growing consensus on the importance of uncertainty quantification, particularly epistemic uncertainty, in validating simulation models was highlighted. This synthesis of diverse methodologies and standards underscored the critical gaps and challenges in existing practices, such as the reliance on experimental data and the difficulties in accurately extrapolating model-form uncertainty. These insights from the literature review reinforced the necessity of incorporating epistemic uncertainty as a fundamental and quantitative measure in the validation process, especially in contexts where empirical validation is limited, infeasible, or subject to large uncertainty. This thorough and cross-disciplinary investigation, therefore, provided a robust scientific justification for the use of epistemic uncertainty in the PV methodology, underlining its relevance and applicability in enhancing the credibility of simulation model predictions. Although the PV methodology was explained in previous publications within the context of PRA of the nuclear industry, the methodology is grounded on this cross-disciplinary review of literature and, thus, facilitates its applicability to validation problems that are not necessarily associated with PRA or nuclear applications.

Second, this PV research line developed a theoretical foundation for PV [1]. Common empirical validations become challenging when validation data are limited. In the PV methodology, the validity of a simulation prediction used for PRA is determined by: (1) the magnitude of epistemic uncertainty (i.e., representing the degree of confidence) in the simulation prediction, calculated using a comprehensive uncertainty analysis that can quantify and aggregate all dominant sources of uncertainty involved in the development and usage of the simulation model; and (2) the results of an acceptability evaluation that determines whether the total uncertainty (including both aleatory and epistemic uncertainties) associated with the simulation prediction is acceptable for the specific application of interest (e.g., PRA). The PV methodology is a unique integration of five key characteristics:

(i) Characteristic (i): The PV methodology offers a multilevel, multi-model-form validation analysis that can integrate data and uncertainty analysis at multiple levels of the system hierarchy to support the degree of confidence evaluation for the system-level simulation prediction.
(ii) Characteristic (ii): The PV methodology separates aleatory and epistemic uncertainties and, when possible, differentiates between two forms of epistemic uncertainty (i.e., statistical variability and systematic bias) while considering their influence on the uncertainty in the simulation prediction.
(iii) Characteristic (iii): The PV methodology uses a risk-informed acceptability criteria, along with a predefined guideline, to evaluate the acceptability of the simulation prediction.
(iv) Characteristic (iv): The PV methodology combines uncertainty analysis with a two-layer sensitivity analysis to streamline the validity assessment and to efficiently improve the degree of confidence in the simulation prediction.

(v) Characteristic (v): The PV methodology is equipped with a theoretical causal framework that supports the comprehensive identification and traceability of uncertainty sources influencing the uncertainty in the simulation prediction.

Third, this research line developed a methodological platform for PV that offers a 14-step "base algorithm" for operationalizing the five key characteristics of PV [2]. This algorithm was designed with three modules that feature a comprehensive uncertainty analysis framework, acceptability evaluation, and importance ranking analysis [2].

Fourth, this research line developed a computational platform for the PV methodology and integrated it into the Integrated PRA Methodological Framework (previously developed by some of the authors of this paper) to assist in risk analysis and risk-informed decision-making activities [2]. The PV methodology and its computational platform were applied to several case studies that are highlighted in Sections 4 and 5 of this paper.

The remainder of this paper is structured as follows. Section 2 highlights the theoretical foundation of PV. Section 3 briefly describes the methodological platform of PV. Sections 4 and 5 provide a summary of current applications, ongoing work, and some future research ideas for PV.

2. THEORETICAL FOUNDATION OF THE PV METHODOLOGY

The theoretical foundation of the PV methodology is centered around a clear and focused goal: to assess the validity of simulation predictions, especially in contexts where empirical validation data are scarce or entirely unavailable. At its core, PV is built on the standard of "Credibility of Accuracy," which acknowledges the often unknown or uncertain nature of true values in simulation predictions. This acknowledgment is crucial in addressing the inherent uncertainties and complexities in predictive modeling. Furthermore, PV demonstrates remarkable adaptability and applicability across a broad spectrum of validation data availability. It is uniquely designed to be effective in various scenarios, ranging from those completely lacking empirical data to those where such data is plentiful. This wide-ranging applicability ensures that PV can be a robust tool in diverse validation contexts, accommodating the varying degrees of data availability and uncertainty that characterize different simulation applications. The theoretical foundation of PV is supported by five key characteristics, each bringing unique contributions and addressing specific facets in the validation process [1].

(i) The first characteristic of PV is its innovative multi-level, multi-model-form validation analysis, designed to integrate data and uncertainty analysis across various levels of a system's hierarchy. This characteristic is particularly crucial when dealing with complex systems, where empirical validation data might be scarce or absent at the system level but available at sub-system levels. The PV approach acknowledges that the quantity, quality, and relevance of sub-system data can vary significantly, being influenced by factors like the complexity of the system's physical interactions and the cost of data acquisition. Models of complex systems typically consist of interconnected sub-models, each contributing to the overall uncertainty in the system-level prediction. The PV methodology addresses this by adopting a multi-level validation analysis, similar to the building block approach in fluid dynamics. This approach divides the system into simpler tiers, allowing for a more detailed and tier-specific assessment of model accuracy. The methodology employs advanced techniques, such as Bayesian methods, to link data and models across different system levels, thereby facilitating effective uncertainty propagation. Unlike previous methodologies, the scope of uncertainty analysis in PV is more comprehensive, guided by a theoretical causal framework that identifies all crucial sources of uncertainty impacting the system-level prediction. Additionally, PV introduces a unique algorithm that leverages the results of uncertainty analysis and predefined criteria to assess the validity of simulation predictions for specific applications. This multifaceted approach to validation analysis marks a significant advancement in handling the complexities and uncertainties inherent in modeling large, intricate systems.

(ii) The second characteristic of the PV methodology is the distinct separation of aleatory and epistemic uncertainties, along with further differentiating between two forms of epistemic uncertainty (i.e., statistical variability and systematic bias) while considering their influence on the uncertainty in the simulation prediction. This is because these sources of uncertainty have different impacts on the total uncertainty of the simulation prediction and are addressed differently if one is seeking uncertainty reduction strategies. The separation of aleatory and epistemic uncertainties is a common practice in uncertainty analysis within domains like nuclear and aerospace but is uniquely applied in PV for the validation assessment purpose. In PV, all important sources of aleatory and epistemic uncertainty, existing in all stages of the M&S process, are identified, characterized separately, and appropriately propagated up to the simulation prediction level (or to the application output level). As a result, the total uncertainty associated with the simulation prediction is a mixed combination of aleatory uncertainty and epistemic uncertainty inherited from the M&S process and is represented using the Probability Bound Analysis approach, i.e., the total uncertainty is represented by a probability box (p-box). This approach enables a more comprehensive and nuanced understanding of the uncertainties involved in simulation predictions, enhancing the accuracy and transparency of the validation process in PV.

(iii) The third characteristic of the PV methodology is the use of risk-informed acceptability criteria, coupled with predefined guidelines, to evaluate the acceptability of the simulation prediction. Typically, in scenarios with sufficient validation data, metrics like the area between the model-predicted and experimentally measured cumulative distribution functions of the System Response Quantity (SRQ) are used to gauge the accuracy of the simulation model. These metrics are then assessed against predefined acceptability criteria or accuracy requirements. In situations lacking sufficient validation data, however, PV introduces the "degree of confidence" concept, derived from the magnitude of epistemic uncertainty, as an alternative to the conventional validation metric. This approach is uniquely developed for PV to enable the evaluation of simulation prediction acceptability based on its current degree of confidence, even in the absence of empirical data. The acceptability criteria, tailored to the specific application's needs and goals, are established before validation to align with the model's expected performance for that application. In PV, if the risk estimates obtained from PRA, with the simulation prediction being an input, are safely within regulatory risk limits (e.g., risk acceptance guidelines in Regulatory Guide 1.174 [3]), the validity of the simulation prediction should be accepted for the risk assessment application of interest, provided that the degree of confidence has already been calculated properly and thoroughly (i.e., all important sources of uncertainty affecting the uncertainty in the simulation prediction have been properly accounted for in calculating the degree of confidence). If acceptability criteria are available at the simulation prediction level, the acceptability evaluation can be performed at this level in a similar manner.

(iv) The fourth characteristic of PV is the integration of uncertainty analysis with a two-layer sensitivity analysis, streamlining the validity assessment and efficiently improving the degree of confidence in the simulation prediction. This approach starts by quantifying epistemic uncertainty as a key metric for validity assessment. Addressing validity issues involves reducing the epistemic uncertainties contributing to the total uncertainty in the simulation prediction. Sensitivity analysis, a more mathematically complex and resource-intensive task than uncertainty analysis, is crucial for understanding how total uncertainty changes with various contributing factors and their interrelations. While some VVUQ methods recognize the importance of sensitivity analysis, they often lack a formal, quantitative screening approach that can screen out insignificant sources of uncertainty to reduce the dimension of the input space and efficiently manage the complexity. PV addresses this by incorporating a two-layer sensitivity analysis. The first layer employs computationally efficient methods, like one-at-a-time (OAT) techniques, to screen out less significant uncertainty sources. The second layer undertakes a comprehensive global sensitivity analysis to prioritize and rank the sources of epistemic uncertainty. This approach not only streamlines the validation process but also aids in efficiently allocating resources for uncertainty reduction, making it a crucial component of the PV methodology.

(v) The fifth characteristic of PV, which is uniquely developed for PV, is its theoretical causal framework that aids in thoroughly identifying and tracing the sources of uncertainty influencing the uncertainty in simulation predictions. This framework is vital for the PV methodology, as the methodology depends heavily on uncertainty analysis for validating simulation predictions. This theoretical causal framework ensures the comprehensive identification, characterization, and propagation of all dominant sources of uncertainty that emerge during the M&S process. The framework categorizes these uncertainties into three main phases: conceptual, mathematical, and computational modeling. Each phase contributes its unique uncertainties, which are systematically analyzed and linked to understand their collective impact on the simulation prediction. For instance, in the conceptual phase, some uncertainties arise from the validity of assumptions and specifications based on subjective interpretations. In the mathematical modeling phase, uncertainties are related to factors such as input conditions and the chosen model form. In the computational phase, uncertainties may stem from sources such as numerical algorithm selection, discretization methods, and programming correctness. This high-level causal framework serves as a qualitative guide in the PV process, ensuring that all significant factors and their associated uncertainties are considered in the uncertainty quantification for the simulation prediction, thereby enhancing the robustness and integrity of the validation process in PV. The development and evaluation of this theoretical causal framework conclude that uncertainties that have arisen during the conceptual and mathematical modeling phases are carried over to the computational modeling phase through the causal relationships among the influencing factors of the causal framework. Thus, to validate the simulation prediction using its associated uncertainty, the theoretical causal framework indicates that the uncertainty can be quantified by characterizing, propagating, and aggregating sources of uncertainty considered in the computational modeling phase. Modeling and measuring techniques for operationalizing this theoretical causal framework are subject to future research.

Among these five characteristics, Characteristics (iii) and (v) are uniquely developed for the PV methodology. Although each of the other characteristics (i, ii, and iv) exists in some of the current studies, the integration of these characteristics under one methodology is a unique contribution of this PV research. Characteristic (i) has not been used in the PRA domain (because complex hierarchical system simulation models have not been integrated into PRAs until recently) and is adopted from other disciplines such as computational fluid dynamics and structural reliability analysis. Characteristic (ii) is common in PRA of NPPs for uncertainty analysis, but not for the validation purpose as seen in other disciplines. Regarding characteristic (iv), literature outside the PRA domain acknowledges the importance of sensitivity analysis for the validity improvement but lacks a formal, quantitative screening technique to make it computationally feasible; and this gap is addressed in the PV methodology.

3. PROBABILISTIC VALIDATION METHODOLOGY

The PV methodological platform was developed to operationalize the PV theoretical foundation and evaluate the validity of simulation predictions of complex, hierarchical systems [2]. The PV methodological steps are organized into three main modules (Modules A, B, and C in Figure 1). Module A is grounded on Characteristics (i), (ii), and (v) of the PV theoretical foundation. Module B and Module C are grounded on Characteristic (iii) and Characteristic (iv), respectively. A shared database is created to facilitate communication among these modules such that it provides necessary input information for execution of the modules and receives and stores results/updated information obtained from these modules.

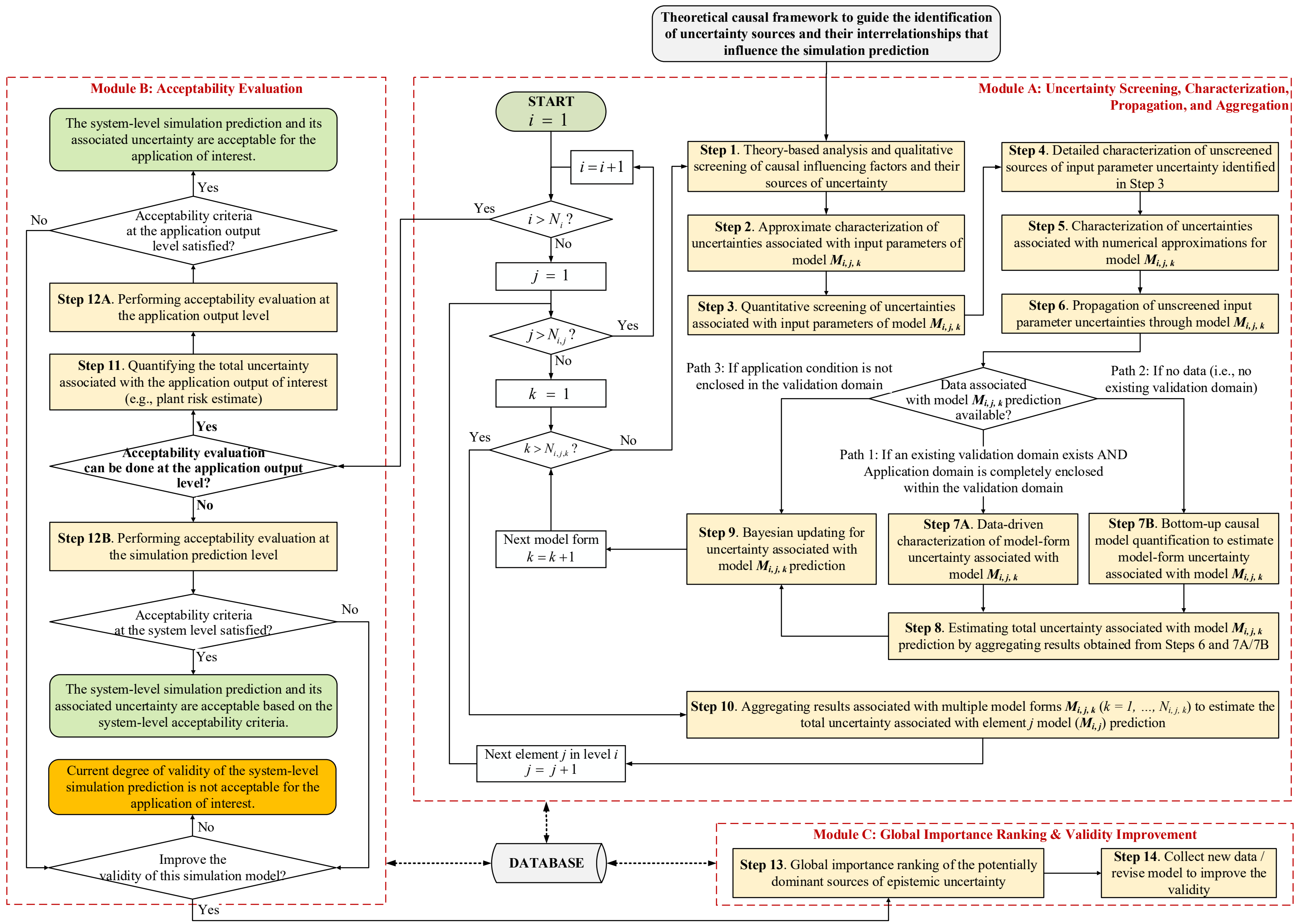

Figure 1. Probabilistic Validation Methodological Platform [1]

Module A is dedicated to a comprehensive uncertainty analysis that spans different levels of the system hierarchy. It begins with a theory-based analysis and qualitative screening of influencing factors and their uncertainties (Step 1 in Figure 1) using a theoretical causal framework (discussed in Characteristic (v) of PV theoretical foundation) and qualitative ranking approaches such as the Phenomena Identification and Ranking Table (PIRT) and Analytic Hierarchy Process (AHP). This is followed by an approximate characterization of uncertainties associated with input parameters (Step 2), typically using available data and expert judgment. A quantitative screening (Step 3) is then employed, often via sensitivity analysis like the Morris Elementary Effects method [4, 5], to identify non-influential parameters and reduce the dimensionality of the input space. Subsequently, a detailed characterization of unscreened sources of input parameter uncertainty (Step 4) is carried out, including a separation of aleatory and epistemic uncertainties. The module also addresses the characterization of uncertainties due to numerical approximations (Step 5) in the simulation model. Uncertainties from important parameters are propagated through the model using methods such as Monte Carlo sampling (Step 6). Model-form uncertainty, arising from assumptions and formulations in the model, is then characterized using either data-driven approaches or bottom-up causal model quantification (Step 7). The total uncertainty associated with the model prediction is estimated by aggregating results from all these steps (Step 8). Bayesian methods are applied to update the model prediction uncertainty using available empirical data (Step 9). Finally, results from multiple plausible model forms are aggregated to estimate the overall uncertainty (Step 10).

Module B evaluates the acceptability of the simulation prediction for a specific application. It involves integrating the simulation model with the computational platform for the application (e.g., PRA application) to quantify the total uncertainty in the application output (Step 11). The module then assesses whether the 'degree of confidence' in the simulation prediction is acceptable by comparing the total uncertainty against predefined acceptability criteria. This evaluation (Step 12) can be made at two levels: the application output level and the simulation prediction level.

Module C, intended for scenarios where the validity of the simulation predictions is insufficient, focuses on global importance ranking and validity improvement. It identifies key sources of epistemic uncertainty impacting the simulation prediction. This includes ranking the magnitudes of epistemic uncertainty components and ranking epistemic input uncertainties using various methods (Step 13). Based on insights from this importance ranking analysis, practical steps for data collection and model revision are suggested to improve the validity of the simulation prediction (Step 14).

Key advantages of the PV methodology include the capabilities to utilize all available data and information at different levels of the system hierarchy to (1) identify important sources of uncertainty that contribute to the uncertainty in the simulation model prediction; and (2) evaluate the uncertainty in the simulation prediction and assess the validity of the simulation model even when system-level validation data associated with the simulation prediction are limited. To perform a comprehensive uncertainty analysis, the PV methodology, however, requires synchronized efforts from a multidisciplinary team of modelers, code developers, and system analysts.

4. CURRENT APPLICATIONS

The PV methodology was computationalized and embedded in the Integrated PRA (I-PRA) methodological framework, previously developed by some of the authors [6], and applied for evaluating the validity of a hierarchical fire simulation model used in the setting of NPP Fire PRA [2]. The step-by-step illustration and results of this fire case study were discussed in Ref. [2] for a hypothetical Fire PRA model. The validity of a fire simulation model in this case study was determined by evaluating: (1) the degree of confidence in the simulation prediction, measured by the magnitude of its epistemic uncertainty; and (2) the result of an acceptability evaluation that determines whether the total uncertainty (including both aleatory and epistemic uncertainties) associated with the simulation prediction and the corresponding degree of confidence are acceptable for the Fire PRA application of interest. A comprehensive uncertainty analysis, embedded in Module A of the PV methodology, was used to identify, characterize, propagate, and aggregate all dominant sources of aleatory and epistemic uncertainties contributing to the uncertainty in the fire simulation prediction and were combined with other sources of uncertainty in the hypothetical Fire PRA model to quantify the uncertainty in the Fire PRA output, i.e., core damage risk estimate. Comparison between the uncertainty in the core damage risk and an acceptability criterion was made to demonstrate the applicability of the acceptability evaluation feature (Module B) of the PV methodology. The case study also demonstrated that an importance ranking of contributing sources of epistemic uncertainty can be conducted (using Module C of the PV methodology) to inform decision-makers as to where additional research and empirical study should be prioritized to reduce the epistemic uncertainty most efficiently. While the PV methodology was illustrated for the Fire PRA context, the methodology itself is applicable for various simulation applications in different domains.

The fire case study in Ref. [2] also brought unique contributions and benefits for the current practice of Fire PRA of NPPs. Most significantly, the work demonstrated a systematically and scientifically justifiable methodology to facilitate the validity assessment for simulation models used in Fire PRA in the context where experimental validation data are scarce (and empirical validation is not applicable). In addition, the work advanced the uncertainty analysis practice in Fire PRA by providing a comprehensive uncertainty analysis framework. The study was the first to (i) include a two-step sensitivity analysis approach for Fire PRA, (ii) screen out insignificant sources of uncertainty using the Morris Elementary Effects Analysis method, and (iii) rank the importance of the unscreened sources of epistemic uncertainty to identify the most significant contributors. Notably, the importance ranking was able to handle different types of uncertainty, including pure aleatory, pure epistemic, and mixed aleatory-epistemic uncertainties. From the PRA quantification perspective, the Fire I-PRA computational platform developed in this Ref. [2], by utilizing the simulation-informed approach, could contribute to more explicit and accurate treatment of dependencies at multiple levels of Fire PRA (e.g., physical inputs, multiple targets, multiple spurious actuations, and PRA basic events). In all previous I-PRA studies by the authors, the PV methodology was not fully integrated into the probabilistic interface of the I-PRA framework to help validate the underlying simulation models. This integration was done for the first time in this Ref. [2] for validating fire simulation models in Fire I-PRA. To computationalize Fire I-PRA, equipped with the PV methodology, this research developed an automated RAVEN-based computational environment (utilizing the open-source RAVEN toolkit developed by the Idaho National Laboratory [7]) that can integrate the plant PRA scenarios, underlying fire simulations, and features of the PV methodology. The RAVEN-based computational platform helped facilitate the sampling-based uncertainty quantification for Fire PRA. Overall, the methodological and computational developments in this Ref. [2] helped improve the realism of Fire PRA results and better support decision-making in the risk-informed regulatory framework in which PRA results are an important input.

5. ONGOING AND FUTURE RESEARCH DIRECTIONS

After the first application of PV for the Fire PRA context, the authors have worked to advance its methodological platform and applied it for validation of a finite element analysis-based simulation model developed for the explicit physics-maintenance coupling in piping reliability analysis [8]. The existing validation studies in the field of NPP piping reliability either compares the estimated reliability metrics (e.g., annual leak and rupture frequencies) with historical pipe failure data or conducts benchmark studies where the reliability estimates from one method are compared with those from other methods. While these existing validation methods focus on the output level, our PV approach focuses on identifying and characterizing major sources of uncertainties at the level of each input parameter and model element and then propagating them up to the reliability estimates. The PV approach was recently adopted as one of the recommended validation approaches for piping reliability analysis of NPPs in the IAEA TECDOC-1988 [9]. An ongoing journal manuscript [10] is being prepared to report on the advancement of the PV approach to account for correlated input parameters in finite element analysis for piping degradation mechanisms.

In an ongoing project sponsored by the U.S. Department of Energy (DOE), the PV methodology is also being extended for the validation of AI/ML-based automation models at NPPs. Preliminary results on this effort have been published in recent conference papers [11, 12]. Our literature review [11] highlighted two challenges for deploying AI/ML-based automation models in NPPs that need further investigation: (i) the lack of a consensus among existing definitions of "automation trustworthiness" and (ii) validation of the AI/ML-based automation, considering its interactions with human operators and plant systems, given the lack of automation performance data from actual plant environments. To address the first challenge, some authors of this paper conducted a systematic, cross-disciplinary literature review to establish the definitions of the relevant key terms for use in the nuclear domain [11]. Based on the literature insights, our research proposed a generic definition of automation trustworthiness as "the degree of confidence that an automation system will function as expected." This definition connects the automation trustworthiness concept to the "degree of confidence," acting as a bridge to a probabilistic measure of epistemic uncertainty. This bridge offers a foundation for the advancement of PV to address the second challenge related to the validation of the AI/ML-based automation models. In this advancement effort [12], the static-dynamic I-PRA framework [6] was extended by developing a coupled simulation model of the underlying AI/ML-based automation-human-physics interactions and integrating this coupled simulation model with the plant PRA model. The PV methodology was then advanced and used in I-PRA to characterize and propagate the uncertainties associated with the AI/ML-based automation-human-physics coupling [12]. Automation trustworthiness can be measured by the epistemic uncertainty associated with the automation output. By analyzing the impact of this epistemic uncertainty on the probabilities of PRA events through the AI/ML-based automation-human-physics coupling, the overall plant risk impact of the automation trustworthiness can be assessed. This research offers a novel risk-informed approach to the evaluation and enhancement of automation trustworthiness. The approach is being demonstrated through a case study where the trustworthiness of an AI-based automated firewatch system for use in NPPs is evaluated. The results of this research are summarized in a conference paper [12] and a journal manuscript under review [13].

The steps of the PV methodology, at this state, constitute a "base algorithm" that is applicable for simulation models of complex, hierarchical systems where the hierarchical system elements can be subject to multiple plausible model forms. The base PV algorithm currently assumes that the element models on the same hierarchical level do not interact with each other in a way that an output from one element model is an input to another element model and vice versa. Such interactions often appear in simulation models with coupled element/sub-models, such as those with the spatiotemporally coupled human-physics models used in the External Control Room (Ex-CR) settings. Ongoing work is being done to advance the "base algorithm" in PV to deal with these spatiotemporally coupled human-physics models. One example is the advancement of PV for the validation of AI/ML-based automation models at NPPs in the context where these models interact with other human and physics models (discussed above). Another example is an ongoing work, sponsored by the U.S. Nuclear Regulatory Commission (NRC) on advancing the PV methodology for validation of a maintenance work process model, considering its interaction with a probabilistic physics of failure. These ongoing efforts are being done and will be reported in upcoming publications.

The PV methodology is equipped with a theoretical causal framework (Characteristic (v) of PV theoretical foundation) that guides the comprehensive identification of uncertainty sources. The current development of this theoretical causal framework is based on a review of the principles and procedures for modeling and simulation formalized in previous studies and recommended in the ASME standards for verification and validation. Future work will update this theoretical causal framework and quantify it by leveraging the data-theoretic methodology developed by the SoTeRiA team [14]. This causal quantification, once developed, will also allow for characterizing model-form uncertainty when validation data associated with the model prediction are unavailable (Sub-step 7B in the PV methodology; Figure 1).

6. REFERENCES

[1] Bui H, Sakurahara T, Reihani S, Kee E, Mohaghegh Z. Probabilistic Validation: Theoretical Foundation and Methodological Platform. ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg. 2023;9.
[2] Bui H, Sakurahara T, Reihani S, Kee E, Mohaghegh Z. Probabilistic Validation: Computational Platform and Application to Fire Probabilistic Risk Assessment of Nuclear Power Plants. ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg. 2024;10.
[3] U.S. Nuclear Regulatory Commission. Regulatory Guide 1.174 (Revision 3): An Approach for Using Probabilistic Risk Assessment in Risk-informed Decisions on Plant-specific Changes to the Licensing Basis. 2018.
[4] Morris MD. Factorial sampling plans for preliminary computational experiments. Technometrics. 1991;33:161-74.
[5] Campolongo F, Cariboni J, Saltelli A. An effective screening design for sensitivity analysis of large models. Environmental Modelling & Software. 2007;22:1509-18.
[6] Bui H, Sakurahara T, Pence J, Reihani S, Kee E, Mohaghegh Z. An algorithm for enhancing spatiotemporal resolution of probabilistic risk assessment to address emergent safety concerns in nuclear power plants. Reliability Engineering & System Safety. 2019;185:405-28.
[7] Alfonsi A, al e. RAVEN Theory Manual. Idaho National Laboratory; 2019.
[8] Cheng W-C. Finite element-based probabilistic physics-of-failure analysis to estimate pipe failure rates for risk assessment of nuclear power plants [Dissertation]: University of Illinois at Urbana-Champaign (UIUC); 2023.
[9] International Atomic Energy Agency. Technical Insights from Benchmarking Different Methods for Predicting Pipe Failure Rates in Water Cooled Reactors. 2021.
[10] Cheng W-C, Sakurahara T, Reihani S, Mohaghegh Z. Probabilistic Validation of Spatiotemporal Probabilistic Physics-of-Failure Model with Correlated Input Parameters Manuscript in Preparation.
[11] Khalid MH, Bui H, Farshadmanesh P, Al Rashdan A, Mohaghegh Z. Automation Trustworthiness in Nuclear Power Plants: A Literature Review. 2023.
[12] Khalid MH, Ahmed MI, Roy S, Bui H, Reihani S, Rashdan AA, et al. Evaluation of Automation Trustworthiness in Nuclear Power Plants: A Risk-Informed Approach using Probabilistic Validation and Integrated Probabilistic Risk Assessment. 2023.
[13] Khalid H, Ahmed I, Bui H, Reihani S, Mohaghegh Z. Evaluation of Automation Trustworthiness in Nuclear Power Plants: A Risk-Informed Approach using Probabilistic Validation and Integrated Probabilistic Risk Assessment. Manuscript Under Review.
[14] Pence J, Sakurahara T, Zhu X, Mohaghegh Z, Ertem M, Ostroff C, et al. Data-Theoretic Methodology and Computational Platform to Quantify Organizational Factors in Socio-Technical Risk Analysis. Reliability Engineering and System Safety. 2019;185:240-60.