ABSTRACT

LIN, LINYU. Development and Assessment of Smoothed Particle Hydrodynamics Method for Analysis of External Hazards. (Under the direction of Dr. Nam Dinh).

Smoothed Particle Hydrodynamics (SPH), as a Lagrangian method, is one of the particle-based Computational Fluid Dynamics (CFD) methods. Because of the unique capability of SPH in handling large-scale fluid simulations with complex interfacial structures, SPH methods are used as the CFD tool for simulating the generation, propagation, and interaction of flooding and high wind with the Nuclear Power Plant (NPP). For the modern NPP risk analysis methodology named Risk-Informed Safety Margin Characterization (RISMC), SPH is first controlled by the statistical software, and multiple simulations are performed by sampling the initial and boundary conditions based on the uncertainty information of external hazards. Next, the simulation results, including the hazard trajectories and impacts on the reactor structure, system, and components, are fed back to the risk analysis tool for identifying the reactor vulnerability under uncertain external hazards and managing the reactor safety margin. The previous study assumes that SPH methods and simulation packages are applicable to the external-hazards risk analysis and that simulation uncertainties do not affect the confidence of safety decision. However, since the product of RISMC aims to inform and support the decision-making regarding the design, operation, and safety of NPP that has high consequences, a convincing validation process is needed to systematically assess the credibility of SPH simulations for designated scenarios. At the same time, considering the potential gaps between established validation framework and application requirements, new methodologies and techniques need to be developed and incorporated for improving the framework. Therefore, this study has two primary objectives: 1) assess the credibility of SPH methods as the RISMC simulation tool for external-flooding and high-wind risk analysis with established frameworks; 2)

improve the assessment framework by identifying and bridging the gaps of practical issues and application requirements.

In this study, the well-established methodology Code Scaling, Applicability, and Uncertainty (CSAU) and its regulatory guide Evaluation Model Development and Assessment Process (EMDAP) are employed to assess the credibility of SPH methods. Expert opinions are elicited from the author's knowledge, and validation data are collected from the literature. Next, adequacy decisions and performance conclusions are made for SPH methods according to the database scaling analysis and the assessment results. At the same time, framework issues and methodology insufficiency are identified accordingly. Attempting to bridge the gaps, three novel methodologies are proposed and formalized according to the application requirements. Meanwhile, case studies are set for demonstrating their capabilities and potential issues. In addition, the scopes of two additionally required methodologies are defined, and a flowchart of the improved CSAU/EMDAP framework is suggested. Finally, a summary of findings and contributions is made together with a discussion on future works.

Development and Assessment of Smoothed Particle Hydrodynamics Methods for Analysis of External Hazards

by
Linyu Lin

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Nuclear Engineering

Raleigh, North Carolina

2019

APPROVED BY:

_____	_____
Dr. Nam T. Dinh Committee Chair	Dr. Igor A. Bolotnov
_____	_____
Dr. Maria N. Avramova	Dr. Abhinav Gupta
_____	
Dr. Robert W. Youngblood	

DEDICATION

To my parents and my love

BIOGRAPHY

The author graduated from University of Michigan, Ann Arbor with a bachelor's degree in Nuclear Engineering in 2013. He joined the nuclear engineering department of North Carolina State University in 2014 and started to work with Dr. Nam Dinh. He obtained his master's degree in 2016. His research focuses on the adequacy assessment for SPH methods and the methodology development for validation. The author is expected to continue his research in nuclear engineering.

ACKNOWLEDGMENTS

First, I would like to appreciate Dr. Nam T. Dinh for guiding my research in the past five years. His knowledge, talent, and character have impressed me so much, and it is always my greatest honor to be one of his students. My English does not allow me to find the most accurate description for his profile in my mind, but I believe that the Chinese word "君子" and its profound meaning could be the most appropriate.

Next, I would also like to thank Dr. Igor Bolotnov, Dr. Maria Avramova, Dr. Robert Youngblood, and Dr. Abhinav Gupta for serving as my committee. Your insights and suggestions are most valuable to me. Especially, I would like to thank Dr. Robert Youngblood for opening the door to a new area.

Next, I would like to thank Dr. Curtis Smith and Steven Prescott for offering an internship opportunity to me in Idaho National Laboratory. My works there mark the start of this study. I would also like to thank Ram Sampath, Niels Montanari, Dr. Nadir Akinci, and Dr. Anh Bui for their supports and suggestions on SPH simulations. Besides, I would like to thank my friends and colleagues at North Carolina State University, Goujing Hou, Han Bao, Chih-Weih Chang, Yangmo Zhu, Yang Liu, Jun Fang, Yuwei Zhu, Hao-Ping Chang, Jinyong Feng, Botros Hanna, Paridhi Athe, Joomyung Lee, Xu Han, Alp Tezbasaran, Anil Gurgen, Nadish Saini, and many others. It has been a pleasure for me to work with you in the past five years.

Finally, I would like to give my special thanks to my parents and my beloved Mengnan for your support and accompany.

The support by the U.S. Department of Energy through the Nuclear Energy University Program (NEUP) and Consortium for Advanced Simulation of Light Water Reactors (CASL) is gratefully acknowledged. The support from the Idaho National Laboratory's National University Consortium (NUC) program is gratefully acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

CDF	Core Damage Frequency
CFL	Courant-Friedrichs-Lewy
CFD	Computational Fluid Dynamics
CSAU	Code Scaling, Applicability, and Uncertainty
DoE	Department of Energy
DNS	Direct Numerical Simulation
DPD	Dissipative Particle Dynamics
EF	Enhanced Fujita
EMDAP	Evaluation Model Development and Assessment Process
EMV	Expected Monetary Value
EOS	Equation of State
ESBWR	Economic Simplified Boiling Water Reactor
EVSI	Expected Value of Sample Information
FoM	Figure of Merit
FV	Finite Volume
GLM	Generalized Langevin Model
HHA	Hierarchical Hazard Assessment
IET	Integrated Effect Test
ISPH	Incompressible SPH
IISPH	Implicit Incompressible SPH
IRP	Integrated Research Project
LES	Large Eddy Simulation
LOOP	Loss of Offsite Power
LWRS	Light Water Reactor Sustainability
MD	Molecular Dynamics
NOAA	National Oceanic and Atmospheric Administration
NPP	Nuclear Power Plant
NRC	Nuclear Regulatory Commission
PCMM	Predictive Capability Maturity Model
PCMQ	Predictive Capability Maturity Quantification
PCMQBN	PCMQ by Bayesian Network
PDE	Partial Differential Equation
PDF	Probability Density Function
PEML	Physics-Evaluated Machine Learning
PTHA	Probabilistic Tsunami Hazard Assessment
PIRT	Phenomenon Identification and Ranking Table
PRA	Probability Risk Analysis
QoI	Quantity of Interest
RANS	Reynold Averaged Navier Stoke
REMDAP	Risk-informed EMDAP
RISMC	Risk Informed Safety Margin Characterization
R/S/U	Relevance/Scaling/Uncertainty
SBO	Station Blackout

SET	Separate Effect Test
SPH	Smoothed Particle Hydrodynamics
SPS	Sub-Particle Scale
SSC	System, Structure and Component
TW	Time Window
UQ	Uncertainty Quantification
V&V	Verification and Validation
VDP	Validation Data Plan
WCSPH	Weakly Compressible SPH

1. INTRODUCTION

On March 11, 2011, Tōhoku earthquake occurred near Japan and triggered a 13-15m high tsunami. In Ōkuma, the wave overtopped the seawall of the Fukushima Daiichi Nuclear Power Plant, flooded and disabled the plant's electrical system. From March 12 to March 15, the insufficient cooling leads to core meltdowns, explosions and release of radioactive material in Unit 1, 2 and 3 (Figure 1-1). Since this accident, increased attentions are paid on to the risk assessment of Nuclear Power Plant due to external hazards. In this chapter, external hazards, including flooding and high wind, are introduced. The well-established Probabilistic Risk Analysis (PRA) method and a novel Risk-Informed Safety Margin Characterization (RISMC) methodology are introduced as the tools of risk analysis and safety management. Since physical models and computational software are applied in risk analysis to capture transient phenomena during accident scenarios, validation framework designed to assess the accuracy of model and software, has been constructed and described in this chapter.

Figure 1-1: Picture of Fukushima accident from BNN [1].

1.1. NPPs' Risk due to External Hazards

In this study, two major external hazards are considered: flooding and high wind. Scenarios of high wind have been analyzed for two decades. Required by U.S. Nuclear Regulatory Commission (NRC), the risk of high wind, including the direct impact and wind-generated missiles, are investigated with PRA and presented in the design control document for the licensing of newly designed NPP [2] [3]. Besides, extensive preparations for the hurricane are required to make for both Nuclear Power Plant (NPP) owner and operator before the hurricane's landfall. While for flooding scenarios, the safety analysis relies on external-hazard PRA with designated event progressions [4].

1.1.1. Risk of Flooding Hazards

Requiring plentiful, reliable sources of water for cooling purposes, many NPPs are located near a body of water, like river or ocean. However, recently the occurrence of Fukushima accident warns regulator and industry that the water source could result in great danger. The floods could damage equipment, knock out the plant's electrical systems, disable its cooling mechanism and lead to core melting. Unlike Fukushima, tsunamis are not the major risk for most NPPs in U.S., natural weather, like heavy rain and snow, can cause rivers to overflow, and the tropic storms can cause storm surges and threaten NPPs near coasts. Another significant source of floods is dam failures. Different from river overflows or hurricanes, dam failures usually occur with little or no advanced warning, which leaves the operator less time to prepare. As stated in an NRC report [5], 700 dam failures have occurred in U.S. since 1975. Since 34 nuclear plants lie downstream from more than 50 dams [6], the cumulative likelihood of at least one plant being flooded due to dam failure is too high to ignore. Responding to the concerns of NPP safety due to external flooding, U.S. NRC has developed recommendations [7] [8] for further regulatory actions in the seismic and flooding designs, and emergency preparations. In addition, U.S. NRC also requests a flooding reevaluation at NPP sites from power reactor licensees and holders of construction permits. Figure 1-2 shows some proposed strategies for post-Fukushima NPP safety enhancement. If the reevaluated hazards are found to be not bounded with current design basis, an integrated assessment is required from the respondents. The integrated assessment is to "evaluate the total

plant response to the flood hazard, considering multiple and diverse capabilities such as physical barriers, temporary protective measures, and operational procedures [9].” Besides, two specific guidance is made for the assessment of flooding hazards due to dam failure [10] and for performing a tsunami, surge, or seismic hazard assessment [11]. Detailed descriptions and proposed safety strategies are further introduced in the following sections.

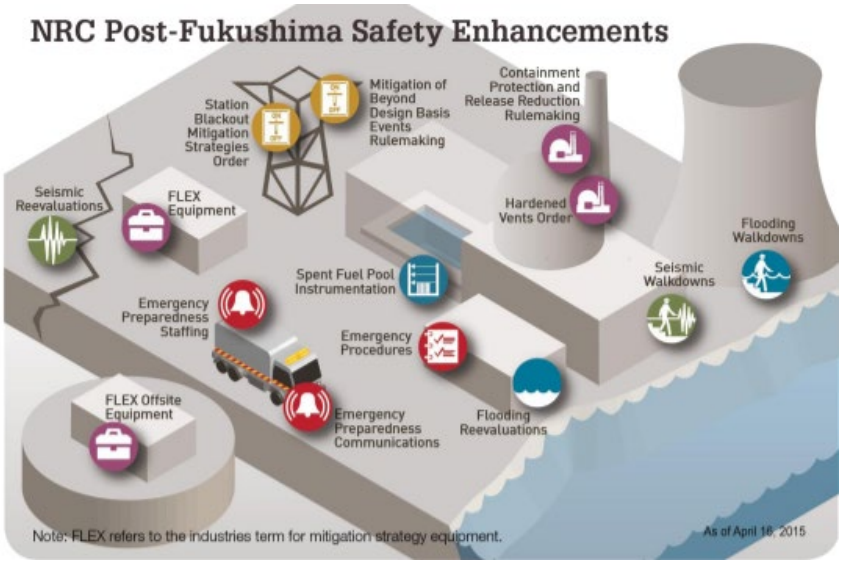


Figure 1-2: NRC post-Fukushima safety enhancements [12].

1.1.1.1. Flooding Hazards due to Dam Failure

There are many reasons that could cause dam failure, including a seismic event, a structural defect, or human-performance-related issues. In fact, not only the reactors located at the downstream of dams will be threatened by such risk, some dams would also impact the plant because of backwater effects. As a result, U.S. NRC first requires a screening of all dams located near an NPP site. Next, a detailed analysis (Figure 1-3) is performed for the potentially critical dams, which includes estimation of demand and loads, assessment of credible failure modes, breach analysis, round routing and inundation mapping.

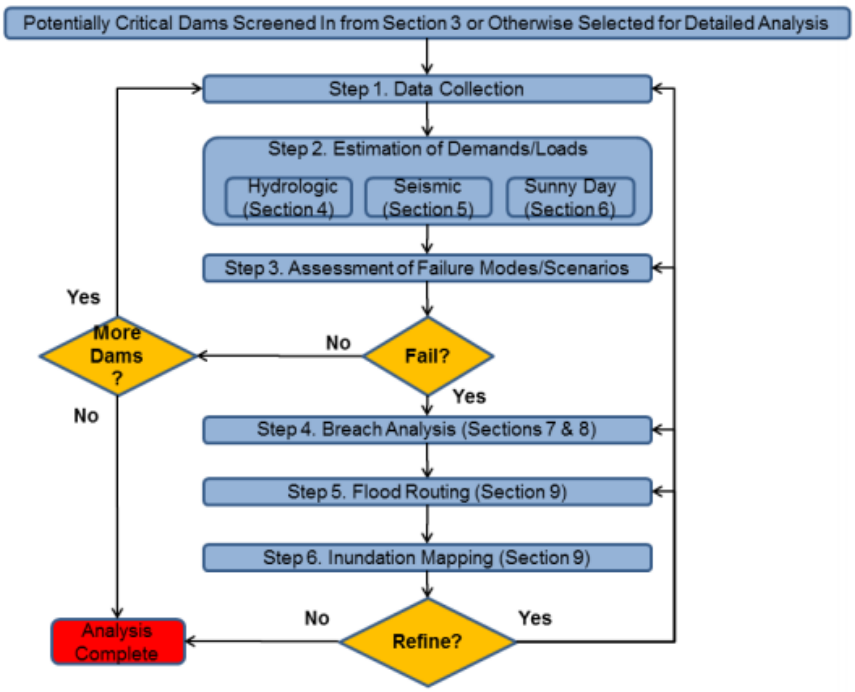


Figure 1-3: Overview of detailed dam failure flood hazard analysis from figure 2 of JLD-ISG-2013-01 [10].

Next, after identifying the critical dams their failure consequence, a probabilistic hazard analysis needs to be performed. However, except for the seismic failure, where Probabilistic Seismic Hazard Analysis (PSHAS) is considered, other types of dam failures, including extreme rainfall and sunny-day dam failures, do not have a widely accepted methodology to estimate their failure probabilities. Figure 1-4 shows a dam failure process prepared by Hanson et al. [13] induced by headcut erosion. The initial overtopping flow results in sheet and rill erosion, and further develops into a series of cascading overfalls. Finally, a large headcut is formed on the slope.

Figure 1-4: An illustration of the dam breaching process by overtopping: (a) rills and cascade of small overfalls at t 1⁄4 7 min; (b) consolidation of small overfalls at t 1⁄4 13 min; (c) headcut at downstream crest at t 1⁄4 16 min; (d) headcut at upstream crest at t 1⁄4 31 min; (e) flow through breach at t 1⁄4 40 min; (f) transition to final breach stage at t 1⁄4 51 min [14].

1.1.1.2. Flooding Hazards due to Tsunami

Tsunami, as one of the most dangerous natural disasters, has been studied for a long time. After the 2004 Indian Ocean (Figure 1-5), U.S. NRC has coordinated a tsunami safety study with the National Tsunami Safety by the National Oceanic and Atmospheric Administration (NOAA). After the Fukushima accident, Probabilistic Tsunami Hazard Assessment (PTHA) starts to be

required as one part of tsunami hazard reevaluations. In some well-established frameworks, the tsunami is generated according to historical data, and it is obtained from the regional or site-specific survey. Next, the computational model is applied to simulate the tsunami propagation. U.S. NRC has suggested several packages based on the shallow-water equation, including MOST, COMCOT and TSUNAMI2. Finally, wave and inundation effects, including the hydrostatic/hydrodynamic forces, debris projectiles, and sediment erosion, are required to be analyzed. However, most of them are analyzed independently by statistical tools, and there is no systematic analysis for capturing flooding progressions and flood-structure interactions.

Figure 1-5: Picture of 2004 Indian ocean tsunami [15].

1.1.1.3. Flooding Hazards due to Surge

Some of the NPP sites are located in the coastal area and near the cooling ponds or reservoirs, which are subject to potential hurricanes, windstorms and squall lines. The changes of water level due to these phenomena could threaten safety-related SSCs in the plant site. To analyze the effects of surge-induced flooding hazards, the U.S. NRC applies the Hierarchical Hazard Assessment (HHA) approach for estimating the generation, propagation and possible interactions of flooding with nuclear SSCs [11]. For the storm generation, historical records are examined to estimate the frequency of extreme winds. In the existing guidance by U.S. NRC, techniques of

Probable Maximum Hurricane (PMH) and Joint Probability Method (JPM) are used for generating synthetic hurricanes (Figure 1-6). Meanwhile, ANSI/ANS-2.8-21992 and Empirical Simulation Technique (EST) are used to generate synthetic extra-tropical storms and squall lines. As for the surge propagation, Advance CIRCulation Surge Model (ADCIRC) model, by USACE Dredging Research Program, is used to simulate tidal circulation and storm-surge propagation over very large computational domains with complex costal configurations [16].

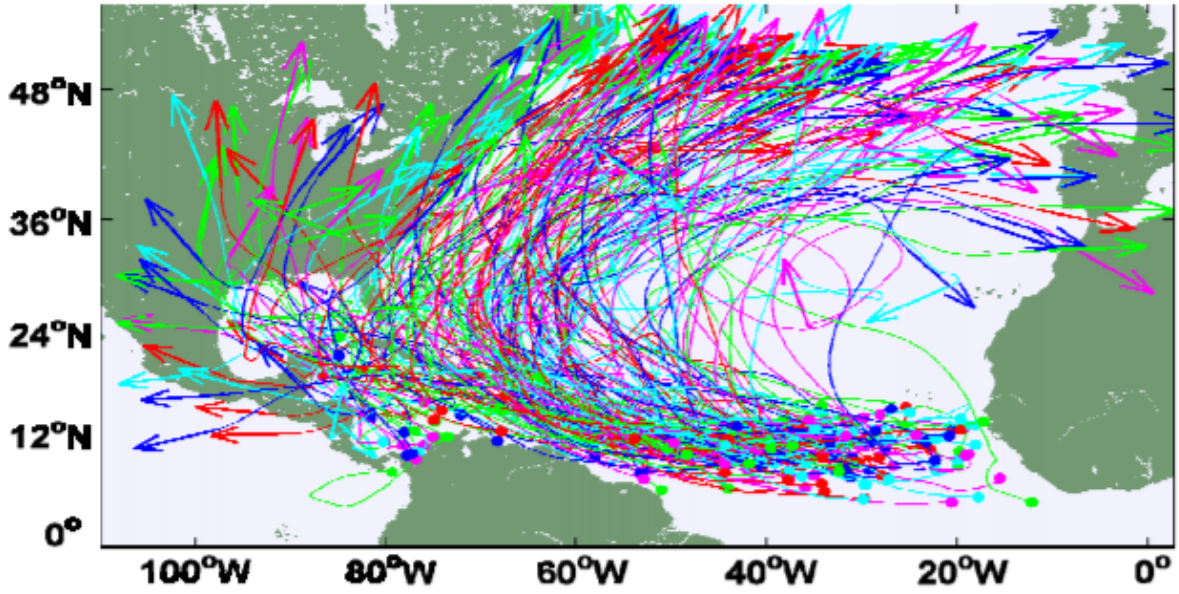


Figure 1-6: Illustration of 200 synthetic storm tracks [17].

In general, U.S. NRC has proposed some guidance to reevaluate the risk due to flooding hazards. Though several statistical and empirical models are recommended, there is no systematic analysis for fully capturing the generation, propagation, and interaction of flooding with NPP sites. Besides, the existing flooding models are developed empirically by statistical tools. And they do not align with the recommendations by U.S. NRC where a combination of probabilistic and deterministic models and tools should be made for analyzing the risk of external flooding hazard [9] [10] [11].

1.1.2. Risk of High Wind Hazards

The effects of high wind on NPP, including hurricanes and tornados, have been investigated since the Turkey Point events on August 24, 1992. Under the strike of Hurricane Andrew (Category 4 tropical cyclone with sustained winds of up to 145mph and gusts of 175mph [18]) on the coast of Florida, Turkey Point NPP loses all offsite power for more than five days. Besides, the facility is also impacted by a complete loss of communication systems, closing of the site access road, and damage to the plant fire protection systems, security systems, and warehouse facilities [19]. Though no damage is found to the safety-related systems except for minor water intrusion [18], significant damage occurred to non-safety-related SSCs, including six turbine canopies, high-water tank for one of the fire protection systems, fossil unit exhaust chimney, (Figure 1-7) and ductwork of radioactive waste building to the vent stack [18]. U.S. NRC and the nuclear industry have been working to identify lessons learned from the event from both the regulatory and the operational sides [18] [20]. As a result, NPP owner/operator has to make extensive preparations for the hurricane before its landfall, PRA analysis is also required by U.S. NRC for NPP licensing. A criterion is set in NUREG-1407 such that the "events pose on the significant threat of a severe accident because the current design criteria for wind are dominated by tornadoes having an annual frequency of exceedance of about 10^{-7}" [21]. Besides, U.S. NRC requires missile protection analysis for NPP licensing [22] has provided information on acceptable designs and acceptance criteria in RG-1.76 for wind-generated missiles [23]. This document defines the range of projectiles with a spectrum of design basis missiles and corresponding impact velocities that have been determined to be acceptable by the U.S. NRC.

Figure 1-7: Structural damage to Turkey Point unit 1 exhaust stack [24].

For identifying and managing safety margins of NPP during external hazards, risk analysis is required for measuring the probabilities and consequences of designated scenarios. The following section reviews the existing risk analysis methodologies in the nuclear discipline.

1.2. Risk Analysis for External Hazards

To analyze the risk due to external hazards, risk-analysis methodologies have been developed to determine the safety margin of NPP during external hazards. Probabilistic Risk Analysis (PRA) is usually recognized a formal and well-established method for NPP risk assessments. However, as discussed in previous sections, the involved physics of external hazards are complex, and the regular PRA are not able to completely capture the progression of hazards and their interactions with NPPs. Therefore, a new methodology called Risk-Informed Safety Margin Characterization (RISMC) is developed by the U.S. Department of Energy (DoE) Light Water Reactor-Sustainability (LWR-S) program to support NPP safety assessments and management. Advanced three-dimension simulations are performed for systematic and

comprehensive descriptions of accident initiating, progression, and consequence. And the results are expected to be more informative and effective for risk management and mitigation purposes.

1.2.1. Probabilistic Risk Analysis (PRA)

In recent years, the impact of external events, including flooding and high wind, has drawn attention from both the regulatory and research departments. For a long time, PRA, as one of the ways to assess the risk of the complex system, has proven to be both efficient and effective in assessing the capability and identifying the safety margin of reactor designs. However, regular PRA has only been applied to the scenario of internal flooding, risk analysis of external flooding is still under investigation. While for the case of high wind, regular PRA methodology usually treats high wind as initial conditions for the status of several SSCs. PRA analysis with four major elements is employed and the schematic procedure is shown Figure 1-8 [27]:

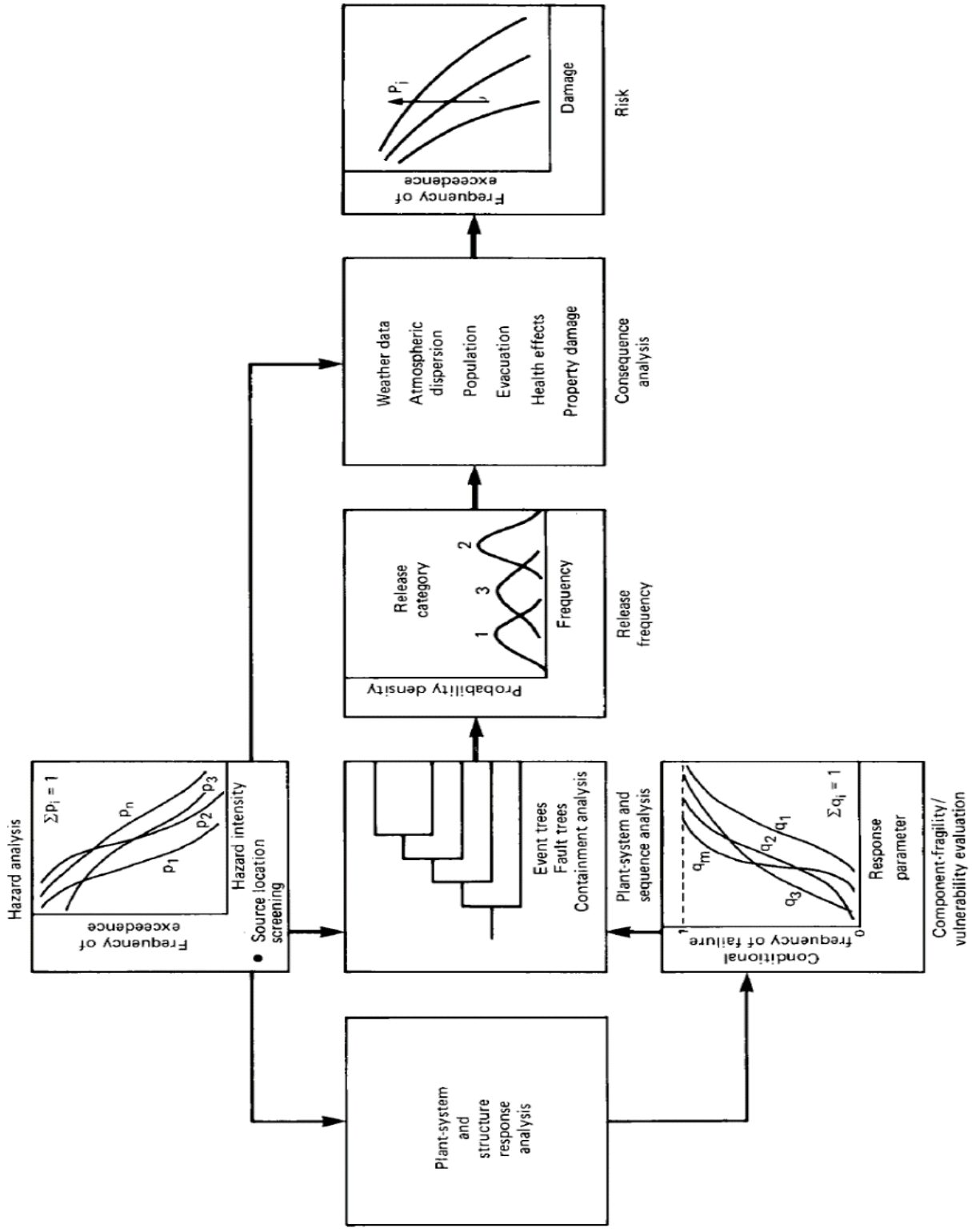


Figure 1-8: Risk-assessment procedure for external events from figure 10-1 of NUREG/CR-2300 [27].

Five major steps are included in the PRA analysis for external hazards, detailed descriptions can be found as follows:

- Hazard analysis – to determine the frequency of occurrence for different intensities of high wind. The result is to generate one or more high wind hazard curves that are applicable to the NPP;
- Plant-system and structure response analysis – to translate the hazard input into the responses acting on a component. The result is frequency distribution of the responses (e.g. force acting on SSCs by winds) for each high wind intensity,
- Evaluation of the fragility and vulnerability of components – to determine the conditional frequency of components' failure given a value of the response parameter. The result is to generate a fragility curves for the various SSCs that are susceptible to the effects of the winds or wind generated missiles;
- Plant system and sequence analysis – to evaluate the risk initiated by high wind or LOCA initiating event induced or caused by high wind. The result is to develop hazard specific fault tree/event tree model,
- Consequence analysis.

Of the five steps of high wind-PRA, the most challenging part is assessing the possibility and impact of the hazard, including the direct impacts of the wind and the impacts of wind-generated missiles. For the PRA study for ESBWR, a tornado strike initiating event frequency is calculated with the methodology in NUREG/CR-4461 [28] and the model for U.S. Central region (Eq. 1).

$$P(u \geq u_0) = P_p \exp\left(-\left(\frac{u_0 - 65}{a_p}\right)^{b_p}\right) + P_l \exp\left(-\left(\frac{u_0 - 65}{a_l}\right)^{b_l}\right) \quad \text{Eq. 1}$$

where $P_p = 3.58 \times 10^{-4}$ and $P_l = 1.74 \times 10^{-4}$ are striking probabilities for point structure and finite structure respectively, $a_p = 26.47$ $b_p = 1.238$ $a_l = 40.84$ $b_l = 1.545$ are constant obtained from descriptive statistics of tornadoes in region of interests. The resulting tornado strike frequencies for ESBWR high wind events can be calculated for each level of tornado events (Table

1-1) and substitute into the event tree for Loss of Preferred Power Event Tree. The Enhanced Fujita (EF) represents the tornado category classified by the NOAA's National Weather Service [29].

Table 1-1: Tornado strike frequencies for ESBWR in high-wind events [25].

Tornado Event	**Tornado Strike Frequency (events/yrs)**
EF2/ EF3	9.68E-05
EF4	4.42E-05
EF5	5.02E-07

It turns out that the total Core Damage Frequency (CDF) for both at-power and shutdown conditions induced by high wind is $4.80 \times 10^{-8}/yr$ with a CDF of $8.51 \times 10^{-9}/yr$ for at-power operations and a CDF of $3.95 \times 10^{-8}/yr$ for shutdown operations. These results are acceptably low and meet the NRC goals.

The missile strike problem is described in NUREG-0800, in which spectrums of missiles are provided [30]. Recently, some initial work has been done by EPRI and computer codes like TORMIS [31] and TORRISK [32] are developed based on the Monte Carlo Method. However, U.S. NRC has described some issues that have occurred in the use of these codes and the interpretation of results, TORMIS/TORRISK still needs detailed analysis to obtain regulatory approval [33]. Another development is made by Westinghouse Electric Corporation for a tool named "Tornado Missile Strike Calculator (TMSC)". It is an application based on Microsoft Excel, and it uses statistical correlations to estimate the missile strike probability [34]. However, it does not explicitly model the physics of high wind, and it might only be applicable for addressing issues within its documented validation basis.

As described in the previous sections, the assessment of high-wind impacts on NPPs require extensive efforts. The scenario is characterized by complex physics, which have different potential interactions with NPP SSCs. Also, high wind represents a set of transient phenomena with multi-physics and multi-scale characteristics. For example, in a tornadic event, the intensity of the wind varies with both the temporal duration of the storm and the physical dimensions of the tornado itself. Twisdale and Dunn indicate that [32], based on an assessment of the resultant debris

field, an F5 tornado will only experience F5 force winds over roughly 15% of its track. Furthermore, the direct effects of the wind are combined with other correlated hazards such as locally intense precipitation and wind generated missiles. However, it turns out that such resolution and complexity are hard to be achieved with the existing methods and tools in the regular high-wind PRAs. New approaches that are capable of accurately capturing multi-physics interactions and dynamically tracking the failure probability are needed.

1.2.2. Risk Informed Safety Margin Characterization (RISMC)

Risk-Informed Safety Margins Characterization (RISMC) Pathway, also known as Risk-Informed Systems Analysis (RISA), is conducting research and development for advanced methods and tools to support Nuclear Power Plant (NPP) safety assessments and management [35]. Within an identified issue space, RISMC analyzes the system's Figures of Merit (FoMs) and the corresponding safety margins probabilistically, where a broader range and types of uncertainties are considered. At this point, RISMC is applied to address the NPP safety margin during external hazards, including flooding, high winds and so on. Considering the complexity of external hazards, advanced simulation tools are needed for a comprehensive scenario representation. Because the RISMC approach explicitly couples probabilistic approaches (the "scenario") with phenomenological representations (the "physics") through a modeling-and-simulation-based approach, it is ideally suited to serve as a framework to address the interactions of external hazards on NPPs and their potential impacts on the NPP safety. Figure 1-9 shows a simple representation of load and capability in both regular PRA (left) and RISMC (right). Originally, the load and capability on the Structures Systems Components (SSCs) are calculated deterministically by the system code with less uncertainty information. In ideal cases, uncertainty of the load and capability are considered, and the safety margin is characterized by the probability of falling in the region where the load exceeds the capacity.

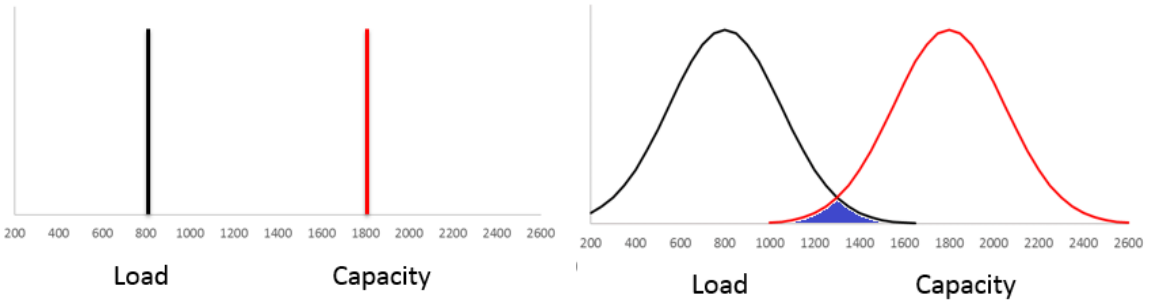


Figure 1-9: Simple representation of load and capability in regular PRA (left) and RISMC (right) analysis. The shaded region in right plot is where the load exceeds capacity and risk could possibly happens.

Note that for the RISMC analysis in practical situations, the simple load-capacity approach turns to be oversimplified since load and capability can be correlated in some scenarios. It has been argued that though the risk-informed concept that margin needs to be quantified with probability remains valid, it is not applicable to compute load and capacity curve separately as if they are independent [36]. As a result, it is suggested that limiting surface should be employed to determine the failure probability [37].

In previous works [26] [38], RISMC has been applied to analyze the risk induced by flooding hazards, and Figure 1-10 shows the corresponding flowchart for tsunami events. RAVEN (Risk Analysis in a Virtual Control Environment) is the statistical package that samples, executes, and evaluates the simulation with uncertain parameters. NEUTRINO, one of the Smoothed Particle Hydrodynamics (SPH) software packages, is applied as the Computational Fluid Dynamics (CFD) tool for simulating the flooding generation, propagation, and interaction with NPP sites. RELAP-7 is the system code designed for the reactor thermal-hydraulics simulation. For the CFD simulation, three-dimensional terrains and building models are constructed by SPH particles for the entire NPP site. Next, multiple full-size and real-time fluid simulations with various initial conditions characterized by RAVEN are performed with NEUTRINO. Event sequences and plant responses, including the dynamic water height, force acting on a certain SSC, and so on, can be extracted for tsunami events. Reference document [26] has demonstrated detailed procedures and findings.

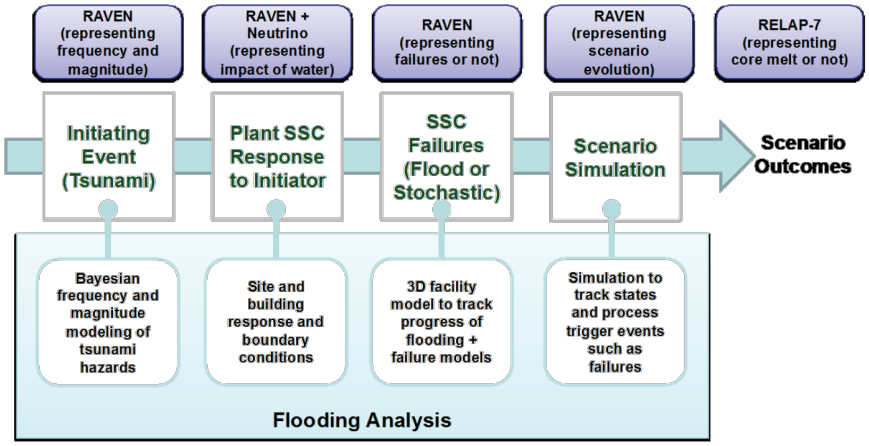


Figure 1-10: Flowchart of applying RISMC in external flooding analysis by C. Smith, et al. [26].

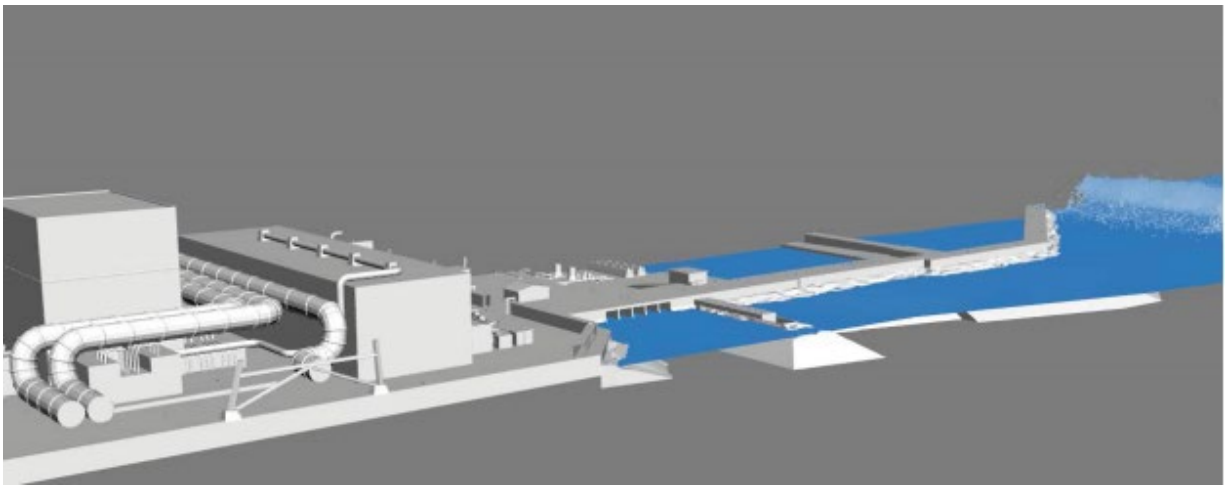


Figure 1-11: NEUTRINO simulation for the Fukushima tsunami event [39].

Figure 1-11 shows a snapshot of tsunami simulation using NEUTRINO with a ¼ slice of the nuclear facility. The impact of the tsunami, whose wave height ranges from 14 to 38 meter, are determined based on the SPH simulation. Based on the information from NEUTRINO, the system code can predict the thermal-hydraulic status inside the reactor core. The scenario-based and risk-informed safety margin can be obtained for the reactor and facilities using the statistical analysis. Reference document [40] shows the safety analysis for a Pressurized Water Reactor during the tsunami event. Though NEUTRINO has shown great capability in performing the large-scale fluid simulation, the simulation's credibility or uncertainty remains to be a problem. Some initial

assessments have been performed for SPH and NEUTRINO [26], however, due to the complexity of physics and phenomenon, a comprehensive validation is still needed to accurately characterize the uncertainty and credibility of simulations. At the same time, the factor of computational efficiency also needs to be considered. Although SPH is comparable to DNS with sufficiently refined particles, it is not practical to perform RISMC analysis with DNS configurations. Besides, as a simulation tool for engineering and industrial applications, the validation goal should be adapted to the contextual requirements, including safety margins, probability of occurrence, consequence, and so on. Therefore, the validation frameworks and standards that guide SPH assessments should be application-oriented and error-tolerant. Meanwhile, the validation results from the framework should be directly used by the applications, like engineering design, safety analysis, etc.

In conclusion, comparing to the statistical analysis by the well-established PRA, RISMC relies on advanced three-dimension and multi-physics simulation tools. In return, RISMC is expected to provide comprehensive and detailed descriptions for the accident initiation and progression. The product of RISMC is also expected to be more effective and informative for the purposes of risk management and mitigation. However, the simulation tools need to be systematically validated such that the predicted System Quantity of Responses (SQR) from simulations are convincing and reliable. Besides, considering the purposes of safety analysis and management, the validation process needs to investigate how model uncertainties affect the safety decisions or facility designs such that the validation goal can be properly designated.

1.3. Model Validation and Uncertainty Quantification (VUQ)

As described in the previous section, both regular PRA and the novel RISMC methods need computational tools to simulate the accident scenarios. And the products of PRA and RISMC aim to inform and support the decision-making regarding the design, operation, and safety of NPP that has high consequences. Therefore, the accuracy and confidence for simulation predictions become crucial, and such assessment process is usually known as validation uncertainty quantification (VUQ) in engineering discipline. This chapter reviews several methodologies and frameworks designed for model VUQs in various disciplines, including aerospace engineering,

mechanical engineering, and nuclear engineering. Moreover, since validation is essentially a decision-making process that aims to determine the credibility of simulation models, this chapter briefly discusses the concept of decision analysis and uncertainties in decision-making. The goal of these reviews is to gain insights from various disciplines and identify potential resolutions that could help the achievement of objectives for this study.

1.3.1. Model Validation Framework

In addition to the AIAA Guidance for CFD code validation [41], U.S. NRC suggests a validation framework called Evaluation Model Development and Assessment Process (EMDAP), which is mainly designed for the system code for transient and accident analysis of NPP [42]. Besides, to support the decision-making that relies on the model credibility, a framework of Predictive Capability Maturity Model (PCMM) is developed to assess the quality of M&S activities with maturity levels.

1.3.1.1. Validation of Computational Fluid Dynamics (CFD)

In order to assess the accuracy and reliability in scientific computing, a method of Verification and Validation process was proposed by W.L. Oberkampf and formulated by AIAA as a guidance in 1998 [41]. Figure 1-12 shows the flowchart of CFD V&V and Uncertainty Quantification (UQ) process. The goal is to estimate credibility and confidence for the simulation model and code. For the validation process, the goal is to determine "the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model [43]" and the fundamental strategy is to "assess how accurately the computational results compare with the experimental data, with quantified error and uncertainty estimates for both [43]". A hierarchical methodology is employed, which separates and simplifies the physical and coupled phenomenon in the complex systems. Figure 1-13 shows the hierarchy structure of the validation tiers. A complete system is divided into three (or more) simpler tiers, including subsystem cases, benchmark cases, and unit problem. It can be noted that the experimental data for benchmarks cases and unit problems is usually from Separate Effect Tests (SETs), in which a special-designed

hardware is fabricated and inspected. The purpose is to isolate each element of a complex system so that "critical evaluations of mathematical models or sub-models can be evaluated [43]". Besides, it is also suggested that repeated experiments on different facilities are needed to ensure the identification of systematic errors. However, some parameters that are needed for the CFD simulation of benchmark cases and unit problems are not measured. Then the analyst should assume these quantities. As for the large-scale testing of the complete system and subsystem cases, also known as Integral Effect Tests (IETs), only the Quantity of Interests (QoIs) for engineering purposes are measured. Because the large-scale tests are usually expensive and complex, complete measurements of all physical parameters are hard to achieve. In addition, for the risk analysis related to NPP and the external hazards, validation experiments of the complete system are impossible to achieve.

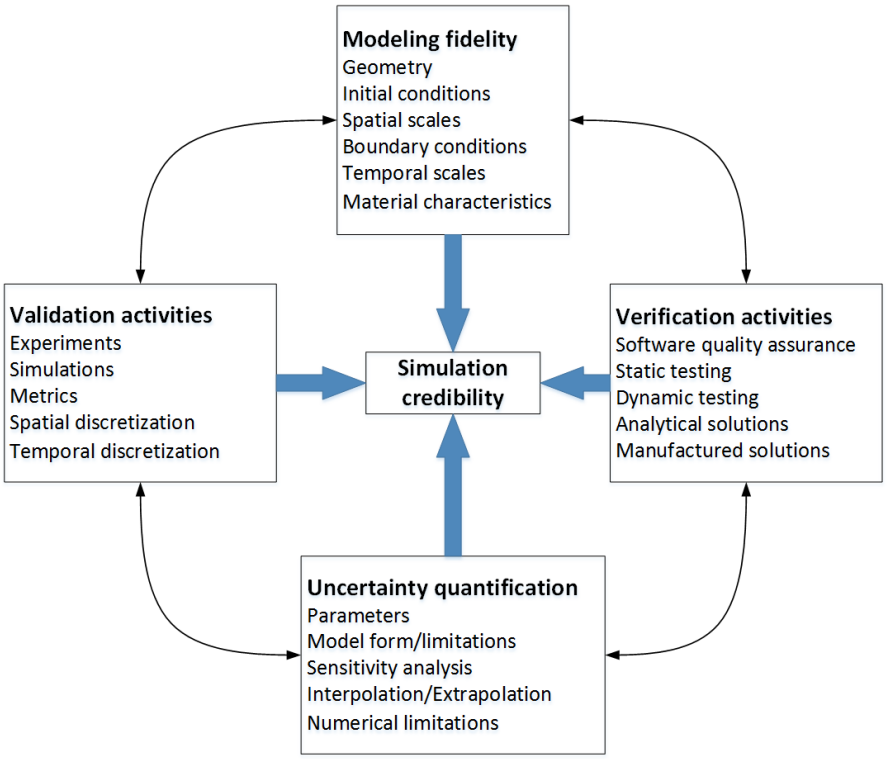


Figure 1-12: Flowchart of CFD V&V and Uncertainty Quantification (UQ) Process.

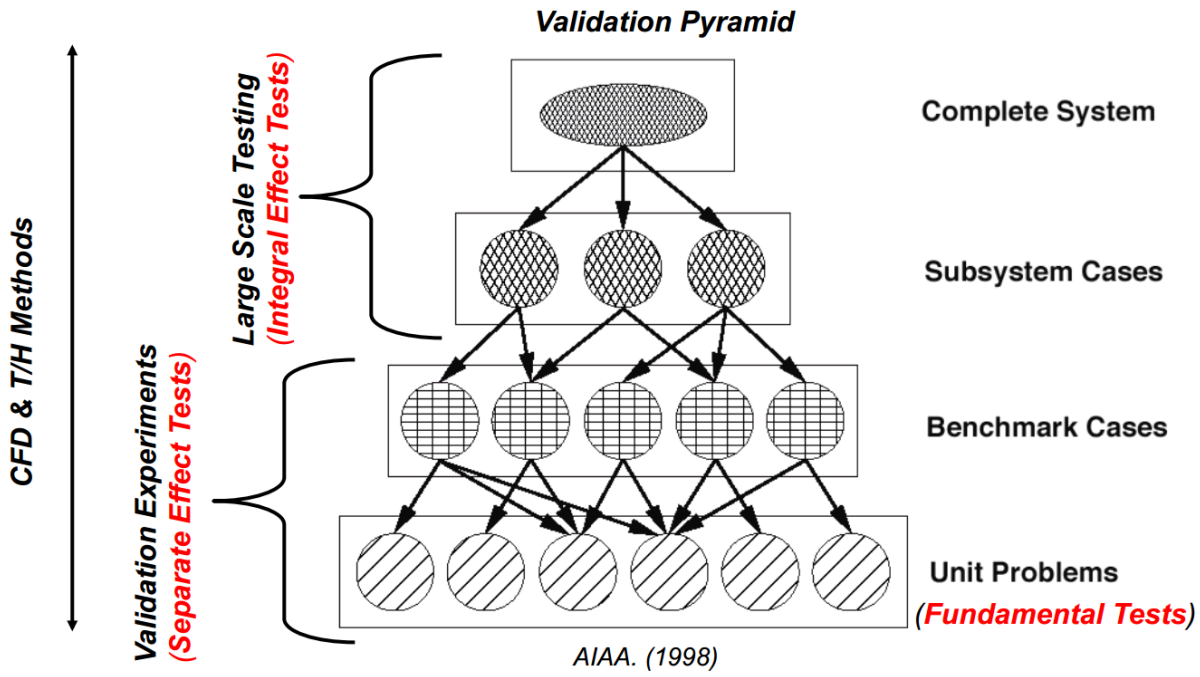


Figure 1-13: Demonstration of validation tiers.

In general, a "divide-and-conquer" strategy is used in the CFD validation with simplistic treatment for multi-physics problems. Note that the concept of scale separation is applied for the hierarchy structures. Due to the simplistic treatment of the multi-scale phenomena, such framework is suitable for the weakly non-linear system. For the application of nuclear engineering, where the scale could range from kilometer (flooding) to micrometer (boiling), the technique of divide-and-conquer may not very applicable. Besides, the CFD validation framework proposed by AIAA is mainly designed for aerospace engineering, where the accuracy requirements are very high. While for the risk analysis in nuclear engineering, the acceptance range for the accuracy of nuclear reactor system code is not as high as the aerospace. As a result, a validation framework designed for nuclear risk analysis is developed by U.S. NRC.

1.3.1.2. CSAU/EMDAP

Code Scaling, Applicability, and Uncertainty (CSAU) evaluation methodology was introduced in 1989 [44] to accommodate the revised rule on the acceptance of Emergency Core Cooling System (ECCS) entitled "Emergency Core Cooling System; Revision to Acceptance

Criteria". The objective is to demonstrate a method that "can be used to quantify uncertainties as required by the best-estimate option described in the NRC's 1988 revision to the ECCS Rule (10 CFR 50.46) [42]". CSAU methodology is mainly composed by three elements [45]: (1) requirements and code capability; (2) assessment and ranging of parameters; (3) sensitivity and uncertainty analysis; The first element aims to identify the code applicability and potential code limitations for the particular scenarios. This is usually achieved with a phenomenon identification and ranking process, (e.g., Phenomenon Identification and Ranking Table (PIRT) [46]), such that the assessment process can be sufficient and efficient. The product of the first element is a hierarchy of relevant tests for the code validation, including the SETs, mixing effect tests (METs), and IET. The second element aims to assess the capability of the code by comparing simulation results against experimental data. The scaling methodology is needed in this element to guide the development of simulation codes and assess the code scale-up capability. The third element aims to assess the uncertainty due to the code limitations, scaling distortions, data quality, and so on. Usually, the second and third element are performed at the same time and the ultimate product is a simple and direct statement of the code uncertainty or credibility in the primary safety criterion (e.g. the Peak Cladding Temperature (PCT)). For obvious and inevitable limitations in knowledge, a conservative margin is added to compensate for the effect of epistemic uncertainty.

Evaluation Model Development and Assessment Process (EMDAP) is a regulatory guide developed by U.S. NRC [42] for code development and assessment. The objective is to describe an acceptable process of developing and assessing the evaluation models that are used to analyze transient and accident behavior within the design basis of a nuclear power plant. The principle of EMDAP is developed based on the CSAU methodology, while EMDAP has formal and explicit descriptions for most of the assessment process, including the PIRT, evaluation model, assessment base, scaling analysis, and so on. After the validation, the system code will be "frozen" and applied to accident scenario for reactor transient and risk analysis. Though CSAU and EMDAP have a logical and comprehensive structure, the decision process for the adequacy is not explicitly specified, and the judgment is made based on experts' opinion. Therefore, for an individual researcher, it is difficult to conduct EMDAP independently and transparently, and the epistemic uncertainty are easily overlooked. For a group of researchers, problems of expert elicitation become serious and challenging. In addition, though CSAU/EMDAP emphasizes the importance

of VUQ, they have a vague classification between the verification and validation. CSAU has been successfully applied to the RELAP-5 simulation code for the Small Break Loss-of-Coolant Accidents of AP600 [47], while EMDAP has not been widely applied. Figure 1-14 shows the overall diagram of the EMDAP process and the relationships among its elements [42]:

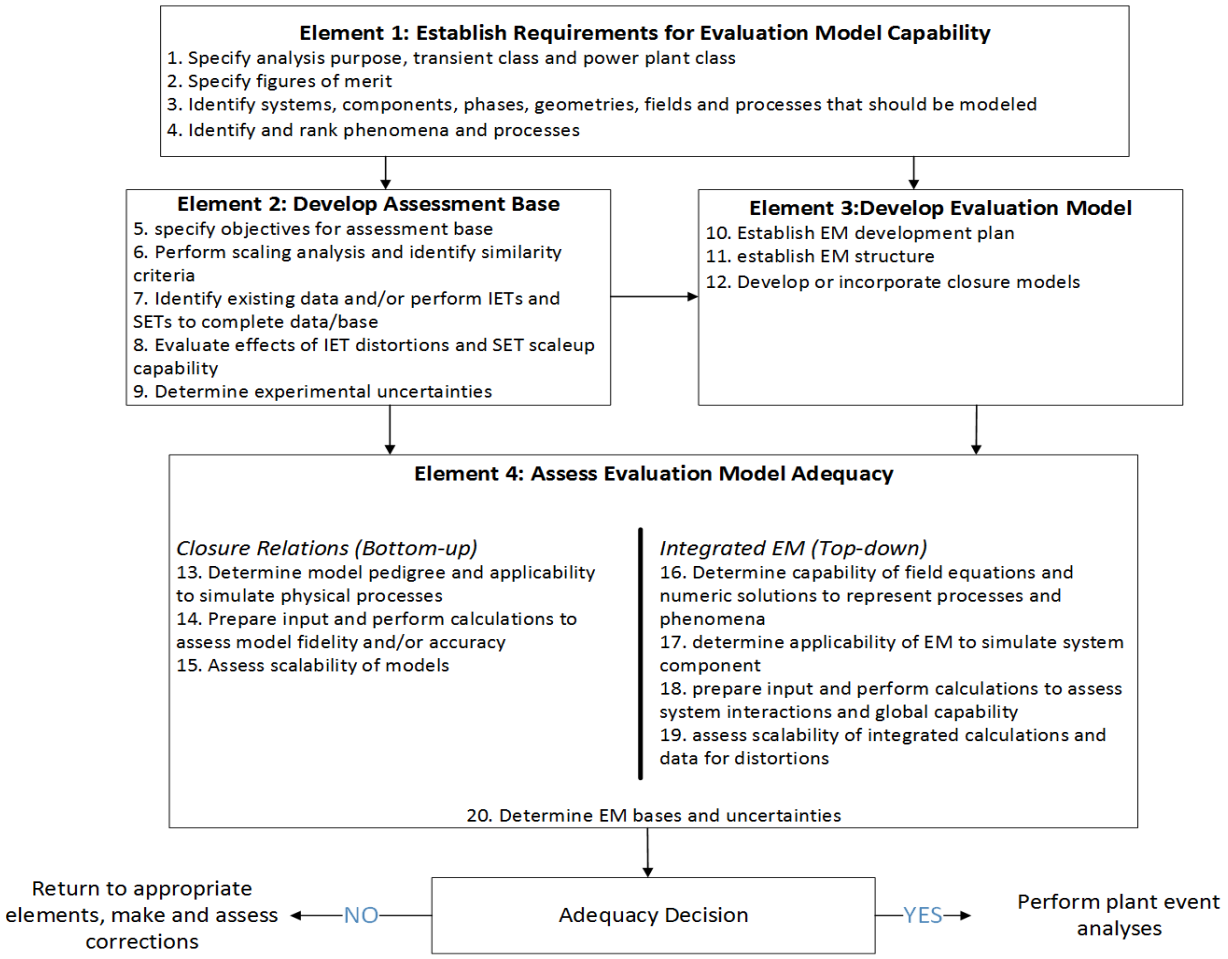


Figure 1-14: EMDAP validation framework.

1.3.1.3. Predictive Capability Maturity Model

Predictive Capability Maturity Model (PCMM) [48] was developed by Sandia National Laboratory in 2007, and it aims to assess the credibility of M&S tools based on the decision consequence. Comparing to CSAU/EMDAP, PCMM treats the model credibility/uncertainty assessment as a decision-making process with explicit structures. First, a specific application and

the corresponding model is selected. Also, based on the nature and consequence of the application, requirements and targets can be designated. Next, for the chosen scenario, six attributes are designed and assessed separately: representation and geometric fidelity; physics and material model fidelity, code variation, solution verification, model validation, and uncertainty quantification and sensitivity analysis. At the same time, a qualitative assessment for each attribute is performed based on a PCMM matrix, and it characterizes each attribute with maturity levels. Finally, the model's achieved level is compared against the target level and the decision of validation adequacy can be made. Since the final decision can either be made based on a stringent requirement or a trade-off between the fulfillment and limitation, PCMM can effectively guide the development and validation of M&S tools. In addition, PCMM presents a formal definition for credibility assessment using the maturity level and it also explicitly distinguishes the difference between verification and validation. Though the maturity descriptor of PCMM includes statements of IETs and SETs, the scaling analysis and the hierarchical structure are not explicitly discussed. As a result, the capability of PCMM is limited in validating complex systems with multiple phenomena. In 2013, PCMM has been specifically discussed in the report by U.S. NRC [49], which emphasizes the use of maturity for the model credibility assessment. CASL also adopts PCMM for assessing Multiphysics computational tools [50].

1.3.2. Uncertainty Quantification

In the process of Verification and Validation (V&V), Uncertainty Quantification (UQ) is an inherent procedure that refers to the activity of identifying and understanding all possible uncertainties within the system of interest. From the fundamental essence, uncertainty can be classified into two categories: epistemic uncertainty and aleatory uncertainty [51] [52] [53]. The aleatory uncertainty is induced by the inherent variation or randomness. Although aleatory uncertainty is not reducible, it can be accurately described with a sufficiently large number of samples. The epistemic uncertainty arises due to a lack of knowledge, and if sufficient knowledge is added, the epistemic uncertainty can be eliminated [54]. In scientific applications, it is usually hard to distinguish them. Depending on the questions, uncertainty classifications can be different for the same problem. It is also argued in some literatures [53] that the aleatory can be interpreted as epistemic uncertainty by the principle of direct inference. In this study, discussions will be

limited to the treatment for epistemic uncertainty, and there are two approaches for representing an epistemic uncertainty: interval with no associated Probability Distribution Function (PDF); or degree of belief by a PDF. The latter approach is usually known as the Bayesian approach to UQ, while the former one is usually based on the interval arithmetic, where no value is truer than any other value. Detailed discussion is beyond the scope, and this study mainly investigates the Bayesian approach.

In the scientific computing, it is important to identify all sources of uncertainty. If a fixed value is known precisely, the simulation uncertainties can be treated as deterministic. Otherwise, they are represented with probability. There are three sources of uncertainty: model inputs, model form, and numerical approximation. Model inputs mainly include the model parameters, initial & boundary conditions, and geometry. Usually, the input uncertainties are characterized or calibrated based on expert opinion, measurements, theories, etc. If the input uncertainties are represented by a distribution function, they will be propagated by performing a number of individual simulations with sampled inputs. The number of simulations depends on the problem specifications, including the linearity of the equation systems, dependence among different inputs, sampling techniques, etc. The model form uncertainty mainly comes from the assumptions that models relies on [43], and it is usually characterized by the validation process. Various validation frameworks have been discussed in previous sections, and extensive discussion will be made in chapter 4 and 5. For models that are built upon Partial Differential Equations (PDE), numerical approximations become an important source of uncertainty. The process of determining uncertainties associated with numerical approximations is known as verification. Specifically, the uncertainties of numerical approximations can be further divided into four components: discretization error, iterative convergence error, round-ff error, and programming mistakes. Various methods have been developed for estimating the numerical error, including higher-order estimator [55] [56] and residual-based estimator [57]. Engineering standards, like GCI method, have also made to guide the mesh convergence studies [58] [59]. Figure 1-15 shows a schematic flowchart for the UQ process, and examples of inputs and outputs for each component are shown.

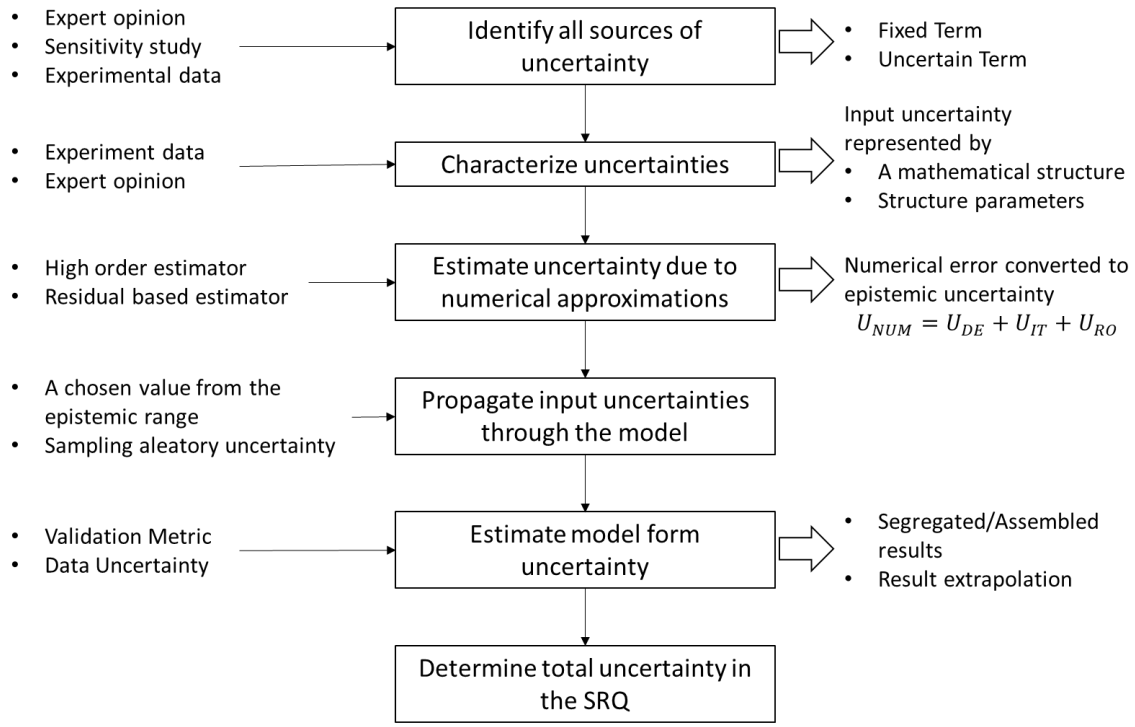


Figure 1-15: Schematic flowchart of uncertainty analysis, including examples of inputs and outputs for each component.

1.3.3. Decision Analysis

The concept of decision analysis is formally defined by R.A. Howard as "a body of knowledge and professional practice for the logical illumination of decision problems [60]." Nowadays, decision analysis has been applied widely used in the field of business, environmental remediation, health care, and management, etc. Generally, decision analysis aims to create value for decision makers facing difficult decisions [61], which could involve multiple decision makers, multiple objectives, complex alternatives, important uncertainties, and significant consequences, etc. As a result, depending on different axioms and applications, the created "values" to the decision maker can be different. For the process of validation, the value can be created by effectively planning the design of validation experiments (Validation Data Plan), assessing the model credibility based on available evidence, or determining the adequacy criteria for validation, etc. To be more specific, decision analysis provides a systematic procedure for transforming opaque decision problems into transparent decision problems by a sequence of transparent steps.

Also, by formally representing the decision problems, decision analysis can produce alternatives that are logically consistent and therefore persuasive [62]. In this study, for model validation purposes, decision analysis aims to improve the assessment for model credibility and identification for validation criteria such that their processes are transparent, consistent, and robust. Various techniques have been developed to represent the validation decision with formal and mathematical languages, including risk-based method [63] and hypothesis testing [63] [64] [65].

1.3.4. Uncertainty in Decision Analysis

The notion of uncertainty has been discussed and emphasized in various fields, including engineering, statistics, psychology, and economics, etc. In each discipline, a specific context and scope can be provided in terms of problem, theory, methods, and tools, etc. In this chapter, uncertainty will be discussed from a broader point of view, where uncertainty can be much larger. Generally, uncertainty can be defined as limited knowledge. At the same time, in decision-making process, the extent of uncertainty involves levels of subjective confidence since it is related to the satisfaction with existing knowledge. In addition, the uncertainty is colored by the underlying values and perspectives of the decision maker involved in the decision-making process, and the decision options available to them.

For the purpose of determining ways of dealing with uncertainty in decision-making process, two extreme levels of uncertainty and five intermediate levels [66] [67] can be made. Table 1-2 prepared by W.E. Walker, et al. [68] defines each level of uncertainty with respect to the knowledge assumed about the various aspects of a decision problem, including future state, the model of relevant system for that future state, the outcomes from the system, and the weight that various decision maker will put on the outcomes. Complete certainty represents the situation where everything is known precisely, and the probability of such event is either 0 or 1. Level 1 uncertainty represents situations where decision-makers admit the uncertainty, but not willing or able to measure the degree of uncertainty in a completely certain way [69]. Aleatory uncertainty is the major source of Level 1 uncertainty, but the uncertainty distribution can be well characterized, and the range of uncertainty is so small that it won't affect the forecast of future state. In scientific computing, the forms and parameters of uncertainty distribution can be well characterized with a

sensitivity analysis for all model parameters. Level 2 uncertainty is the situation that decision-makers are not certain about the future state. And such uncertainties are usually captured in the form of multiple forecasts (scenarios) with associated probabilities. Level 2 uncertainty can be adequately described in statistical terms. In risk analysis, event/fault tree is usually used for determining the possible outcomes and the corresponding likelihoods. Level 3 uncertainty represents the situation where decision-makers can create multiple alternatives and rank the alternatives in terms of perceived likelihood rather than probabilities [70]. The alternatives represent some different parametrizations of the system model, alternative sets of outcomes, and/or different conceivable sets of weights. At the same time, the likelihood is usually created based on available knowledge and information for these alternatives. For Level 3 uncertainty, the future states can usually be represented by more than one alternative, which is often developed based on various assumptions. Furthermore, at least one or the range of model predictions can cover the real future states. In scientific computing, the alternatives are characterized with mathematical models. Level 4 and Level 5 uncertainty is usually defined as "deep uncertainty" under which conditions that "analysts do not know or the parties to a decision cannot agree upon (1) the appropriate models to describe interactions among a system's variables, (2) the probability distributions to represent uncertainty about key parameters in the models, and/or (3) how to value the desirability of alternative outcomes" [71]. To be specific, the phrase of "do not know" refers to level 5 uncertainty, while "cannot agree upon" refers to Level 4 uncertainty. In this study, focuses have been put onto Level 2 and 3 uncertainties.

To deal with large uncertainties in decision-making, four general approaches are found from literatures [68]: resistance, resilience; static robustness, and adaptive robustness. The resistance approach aims to plan for the worst conceivable case or future states. And the technical approach includes conservative treatments by introducing sufficient safety margins. The resistance approach is usually expensive, and its resistance against unexpected states, commonly known as "black swan" [72], is relatively poor. The second approach accepts short-term pain, such as negative system performances), but focuses on recovery. The third and fourth approaches do not aim to determine the "best" predictive model and solve for the decisions that are optimal since they are fragilely dependent on assumptions. Instead, the robust approach aims to seek for the most robust decisions, which is able to "achieve a given level of goodness across the myriad model and

assumptions consistent with known facts" [73]. Before making decision, the static robust approach will explore how different assumptions about future values of uncertain assumptions that affect the decisions. Scenario planning [74] is an example of such approach. It assumes that there are a range of plausible futures that can be specified well enough to identify a static decision, and it will produce acceptable outcomes in most of them. The fourth approach aims to generate decisions that can adapt to changing conditions over time. Since this approach is proposed mainly for Level 5 uncertainty, which is beyond the scope of this study, no detailed discussion will be further made.

Table 1-2: Definition of uncertainty levels with respect to the knowledge assumed about the various aspects of a decision-making problem [68].

Complete Certainty		**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**	Total ignorance
	Future State	A clear future with sensitivity	Alternate future with probabilities	Alternate futures with ranking	A multiplicity of plausible futures no ranking	Unknown future	
	System Model	A single system model	A single system model with a probabilistic parameterization	Several system models, one of which is most likely	Several system models, with different structures	Unknown system model; know we do not know	
	System Outcomes	Point estimates with sensitivity	Several sets of point estimates with confidence intervals, with a probability attached to each set	Several sets of point estimates, ranked according to their perceived likelihood	A known range of outcomes	Unknown outcomes; know we do not know	
	Weights on the outcomes	A single estimate of the weights	Several sets of weights with a probability attached to each set	Several sets of weights, ranked according to their perceived likelihood	A known range of weights	Unknown weights; know we do not know	

In conclusion, this chapter discusses and reviews state of the art for model validations. It is found that well-established validation frameworks have a pyramid structure for integrating evidence and argument related to model credibility/uncertainty. After the integration, an adequacy decision is needed to judge if the simulation tool is adequate for performing confirmatory analysis for a given application and recommended guidance. Depending on the inherent methodologies, different frameworks have their integration schemes and performance standards. In this study, two established frameworks in nuclear discipline have been reviewed. The CSAU/EMDAP framework emphasizes the scaling methodology for bridging the gaps between the validation database and application scenarios. The PCMM formalizes the model validation as a decision-making process and assesses the model adequacy according to six attributes. Meanwhile, maturity levels are defined for each attribute, and the adequacy is decided by comparing the achieved maturity levels against the target levels. If features from both frameworks are considered, validation is essentially a decision-making process under uncertainties. As a result, the concept of decision analysis and the involving uncertainty are reviewed. According to the aspects of future state, system model, system outcomes, and weights on the outcome, the uncertainty is classified into five levels. It is found that the well-established UQ process for the scientific computing belongs to level 1, and the subjective assessments become inevitable when uncertainty is getting large (level 2 and 3). In model validations, such large uncertainties are induced by data gaps and imperfect model for describing the system behaviors. Therefore, when uncertainties are large, to validate a model transparently and consistently, the subjective components need to be formalized appropriately.

1.4. Summary Remarks

This chapter first describes the significant safety challenge for the NPP due to the external hazard and emphasizes the necessity of the corresponding safety/risk analysis. Next, section 1.2 reviews the methodology for external-hazard safety analysis, including regular PRA and RISMC. It is emphasized that though RISMC can characterize the reactor safety margin with higher resolution and broader coverage, its reliability depends on the uncertainty of simulation tools. And a comprehensive estimation for simulation's credibility and uncertainty is needed. Next, section 1.3 discusses the state of the art for model validations and reviews two validation frameworks in the nuclear discipline. It is concluded that the model validation is essentially a decision-making

process under uncertainties, which are introduced by the insufficient databases and imperfect models. Section 1.3 also reviews the decision analysis under uncertainty and categorizes the uncertainties according to four key aspects. Figure 1-16 illustrates a general validation framework with pyramid structures, where evidence and argument of model credibility/uncertainty are integrated for informing the adequacy decision. If the model turns to be adequate, the credibility/uncertainty information will be fed into the safety analysis framework any safety-related decision.

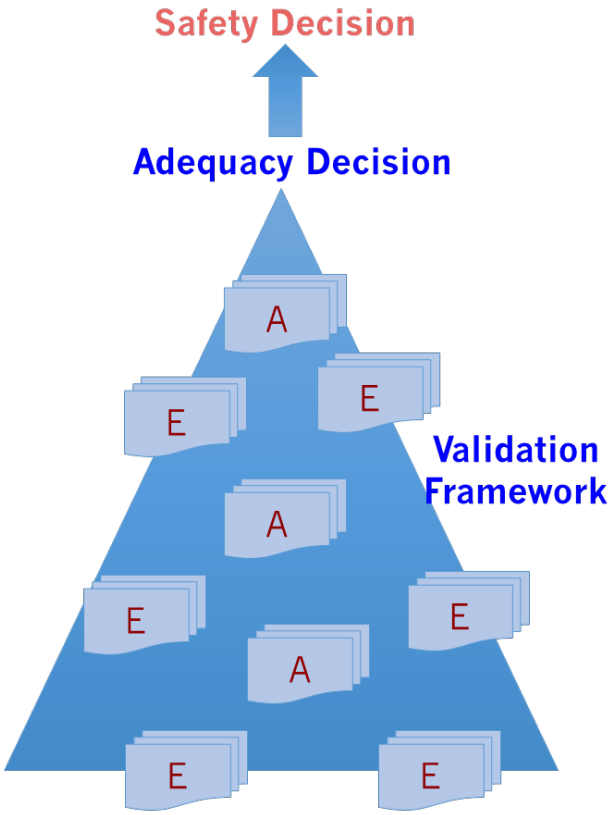


Figure 1-16: Illustration of pyramid-type validation framework, and its connection to adequacy decision, model credibility, and the safety decision. Letter "E" stands for **E**vidence, while letter "A" stands for **A**rgument. Both evidence and argument are for model credibility/uncertainty. After the integration by validation framework, an adequacy decision will be made with respect to the selected model and application. If the model turns to be adequate, the credibility/uncertainty information will be fed into the safety analysis framework any safety-related decision.

2. DISSERTATION BACKGROUND

This chapter aims to describe the workflow of this work according to its motivation, objective, and technical approach. At the same time, important conditions and assumptions are listed for refining the investigation scopes. At last, expected outcomes and products are described.

2.1. Motivation, Objective, and Technical Approach

Motivated by the application of RISMC for the reactor safety analysis, this study aims to assess the credibility of SPH methods in simulating the generation, propagation, and interaction of the external flood and high wind with the NPP. At the same time, considering the state of the art of the validation frameworks and the SPH methods, the processes of assessment and development are expected to be perform iteratively: (1) assessment for SPH methods is first performed with established framework; (2) issues and challenges are identified, at the same time, new methodologies and capabilities are developed and incorporated into the existing framework or the SPH tools. Next, similar assessment and subsequent framework developments are performed iteratively until the SPH tool is adequate for performing confirmatory analysis for the given applications and recommended guidance. As a result, this study has two primary objectives: 1) Assess the credibility of SPH methods as the RISMC simulation tool for external-flooding and high-wind scenarios with established framework and existing SPH tools; 2) Improve the assessment framework by identifying and bridging the gaps between practical issues and application requirements in a transparent, consistent, and robust manner.

Objective #1 and objective #2 are tightly coupled: the reliability of assessment processes depends on the quality of the framework, while the quality should be assured by minimizing the gaps between application requirements and achieved levels. Therefore, the technical approach of this study is (1) to identify the requirements of RISMC applications; (2) to assess the SPH credibility with regular and well-established validation frameworks; (3) if a confident conclusion on code adequacy cannot be drawn, an improvement will be made to the established framework in a transparent, consistent, and robust manner. For those improvements that have not been

formalized, clarification for their scopes or requirements should be made for further developments. This process will be repeated until a confident conclusion is obtained.

2.2. Dissertation Conditions and Assumptions

To properly identify the scope of this study, important conditions and assumptions are listed in Table 2-1. Category A aims to refine the investigation scope; category B focuses on the formalization of validation process; category C focuses on the novel UQ technique based on local information of SPH simulations.

Table 2-1: Important conditions and assumptions with respect to aspects of investigation.

ID	Conditions/Assumptions
A	**Investigation Scope**
A1	The SPH-based computer codes, including LAMMPS-SPH and NEUTRINO, have been verified
A2	Current work only focuses on the RISMC application in external hazards, including flooding and high wind;
A3	When uncertainty is large, the subjective assessment for technical questions cannot be avoided
A4	The uncertainty for accepting a simulation tool is not deep
A5	Validation frameworks of CSAU/EMDAP are consistent and improvable
A6	Attentions have been payed to the assessments of separate phenomena or behavior
B	**Validation Formalization**
B1	The validation process is a structured argument process supported by a body of evidence;
B2	The uncertainty regarding model credibility can be measured with a probability and updated with evidence;
B3	Human preference can be represented analytically by a function;
B4	The subjective value associated with the uncertain outcomes is associated with the expectation of the decision maker's valuations of the outcomes;
B5	Standard Bayesian analysis is applied to all possible combinations of prior distributions and likelihood functions selected from classes of priors and likelihoods considered empirically plausible by the analyst [75];
B6	The attributes of maturity assessment are conditionally exclusive;
C	**Local Data-Driven UQ**
C1	The simulation uncertainties are so large that they cannot be precisely represented by parametric distributions;
C2	Sub-scale phenomena taken together produce macro-scale behavior;
C3	Training and testing cases are dominated by the same physics
C4	Each SPH particle has a stable size and the volumes are very close to the mean value

In scientific computing, verification is usually divided into two types: code verification and solution verification. The code verification aims to ensure that the code is a faithful representation of the underlying mathematical mode, while the solution verification aims to ensure that the simulation (i.e. numerical approximation) of a mathematical model is sufficiently accurate for its intended use. Assumption A1 mainly assumes that the code verification has been performed. The confidence on such assumption is built on the theory manual of LAMMPS-SPH [76] and NEUTRINO [77], together with their past applications in various disciplines. Assumption A2 is prepared for the subsequent VUQ process for defining the intended use and problem of interest. Assumption A3 and A4 constrain the investigation to level 2 and 3 uncertainty in decision analysis. In this study, such uncertainty is mainly induced by the lack of large-scale data and knowledge and poor estimation for simulation uncertainties. Assumption A5 rationalizes the technical approach of this study, where issues and challenges of the CSAU/EMDAP validation framework are suggested to be resolved by incorporating new methodologies. Specifically, if the validation process by CSAU/EMDAP is treated as a deductive process, the consistency suggests that such process can be represented by a model, and there is an interpretation under which all formulas in the theory are true. As a result, it becomes reasonable to improve the CSAU/EMDAP framework by formalizing the process, interpreting the formulations, and verifying the models. However, since this study does not focus on experimental design or data acquisition, and the database is acquired from literature reviews, a proper theory and database verification is beyond the scope. Meanwhile, assumption A6 restricts the major investigations to accuracy assessment of separate phenomena due to the limitation of the database.

As discussed in section 5.2, category B aims to formalize the validation process as an argument process (B1) and to represent the validation argument with Bayesian decision theory. Assumption B2, B3 and B4 rationalize the application of probabilistic reasoning for arguing the credibility of simulation codes. Assumption B3 suggests a utility function or monetary value for representing the preference of decision-makers, and B4 suggests decision-maker to use expected-utility or expected-monetary-value criteria as rule of choice. Assumption B5 is made for the Bayesian sensitivity study as a convenient approach to deal with imprecision issue. Besides, assumption B5 is the basic idea for robust Bayes by J. Berger [75], where standard Bayesian analysis is applied to all possible combinations of prior distributions and likelihood functions. The

prior distributions are selected from a class of priors, while any likelihood function that are plausible is used. However, it is argued that robust Bayes approach is inconsistent with the Bayes' theorem, where the probability and utility should be measured precisely by functions [53]. Admitting the difficulty of identifying precise prior distributions or likelihood functions, robust Bayes or Bayesian sensitivity analysis is accepted as a convenient approach that extends the application of traditional Bayesian analysis. Since this study employs Bayesian Network for mathematically representing evidence gathering and adequacy decision process, Assumption B6 is suggested as one of the requirements of the Bayesian network.

Category C aims to rationalize the technical approach of local data-driven UQ discussed in section 5.3. Assumption C1 suggests that the total SPH simulation errors should neither be assessed with respect to each uncertainty source nor represented accurately by parametric distributions. Assumption C2 suggests that the macroscopic properties and coefficients, like the total uncertainty, can be estimated by integrating all local uncertainties. Such an assumption is inspired by concepts from the material science and biology discipline, where information from molecular dynamics simulations (microscale particle-based simulations) are used for determining the global material quantities. However, the integration and averaging process can be very complicated, and a powerful algorithm is needed for resolving highly nonlinear interactions. Assumption C4 assumes that the SPH fluids are weakly compressed such that all SPH particles have stable sizes, which can be approximated as the cube of average particle diameter.

For clarification, the term "subjective" has been frequently used in this report, however, this study does not intend to focus on the philosophical argument of subjectivity against objectivity or rational against emotional. Instead, this study uses the term "subjective" to represent rational human assessments for technical questions, which have insufficient evidence and poor formalizations. Also, the term "subjective" is an important part of Bayesian interpretation for probability, which will be rationally changed to account for availability of related evidence. Definitions for the term "subjective" is also provided in the glossary table at the end of this document. Meanwhile, the term "credibility" is defined as a combination of objective and subjective measurements of model quality for intended uses.

2.3. Dissertation Overview and Outline

To explain the technical approach and demonstrate the progress, this study is organized as follow: First, chapter 3 reviews the current state of the art for SPH methods and analyzes the truncation and discretization error of SPH methods. Next, chapter 4 demonstrates the credibility assessment for SPH methods with the Code Scaling Applicability and Uncertainty (CSAU) methodology and its regulatory guide Evaluation Model Development and Assessment Process (EMDAP). Conclusions regarding the SPH model adequacy for each scenario are made according to the assessment results and scaling analysis. At the same time, issues and challenges of validation frameworks are identified. Next, chapter 5 formalizes three proposed methodologies for the corresponding framework limitations. Besides, chapter 5 demonstrates the capabilities of new methodologies and identifies possible limitations through various simple case studies. At the end of chapter 5, an improved validation framework named is suggested by incorporating the new methodologies into the CSAU/EMDAP framework. At the same time, this study aims to contribute to the development of risk-informed EMDAP framework (REMDAP). Recommendations are provided for the validation of SPH and system simulations according to the developments and achievements of REMDAP components. Meanwhile, potential conflicts and issues are suggested from the perspective of fundamental concepts and practical implementations. Finally, chapter 6 summarizes this study by listing key findings and major contributions. Recommendations for the further work are also provided. Since new concepts and technical terms are frequently used in this study, a glossary is prepared at the end of this report for better understanding. Figure 2-1 shows the scheme of the structures of this dissertation.

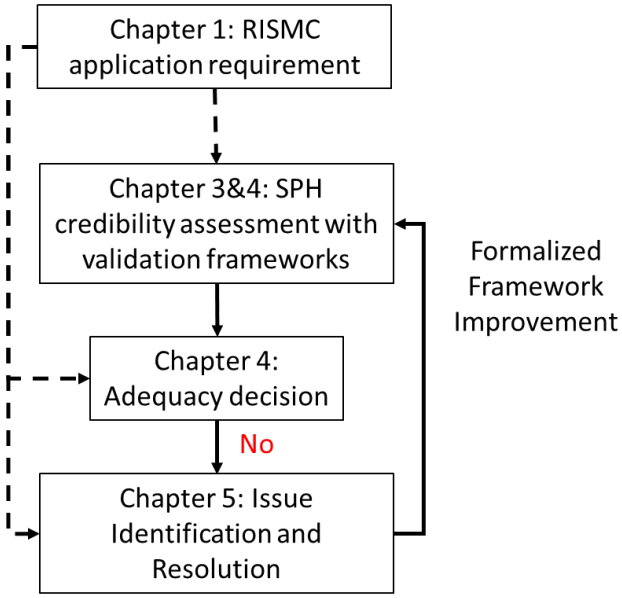


Figure 2-1: An illustration of dissertation structures and flowcharts.

3. SMOOTHED PARTICLE HYDRODYNAMICS (SPH)

As one of the mesh-free methods for solving Partial Differential Equation (PDE), SPH is able to simulate highly non-linear fluid problem, such as multi-phase simulation, fluid-structure interaction, shock waves and so on. Besides, compared to the mesh-based methods, the computational cost for simulating such problem is much lower, which is found to be crucial to the RISMC analysis. Therefore, SPH is selected as the CFD tool in simulating flooding and high wind phenomenon, and their interactions with NPP sites and SSCs. Though the SPH simulation is usually crude with relatively low accuracy, the error is still acceptable for the simulation of external hazards. However, the careful investigation is needed for the SPH simulation error such that the results from SPH are acceptable for the risk analysis. As a result, a validation framework with uncertainty quantification is needed for SPH simulation. In the current study, an SPH-based simulation software named NEUTRINO [78] is employed. Before the recommendation of the new framework, SPH is first introduced based on the past literature. Then SPH's mathematical formulation and error analysis are performed for better characterizing the model and simulation error.

3.1. State of the Art

In 1977, SPH was first invented to deal with astrophysical problems in three-dimensional open space [79] [80], where the collective movement of particles is similar to the movement of liquid or gas flow. Since then, SPH has advanced tremendously in hydraulics with a large amount of work being performed both theoretically and computationally. Though mesh-based methods are still the dominant tool, SPH has built its fidelity in simulating fluid problems with complex and irregular interfaces [81]. Recently, SPH has grown popular and become an essential part of the numerical arsenal of several industrial and laboratory institutions. For RISMC analysis of flooding and high wind, as a Lagrangian method, SPH can predict the particle motion through space and time with no requirement for any underlying mesh. This brings some key advantages with SPH in the field of long wave production and development [82] [83] [84] [85], sloshing [86], floating objects [87] [88] [89] [90], water wave impacts and interacting force estimate [91] [92] [93], which is perfectly suitable for the RISMC simulation in flooding and high wind. In the following section,

the PIRT process is conducted for a designated scenario in flooding and high wind respectively. The adequacy of SPH in model, code, and validation data is assessed for the intended uses. In general, the success of SPH in these fields relies on its capabilities in simulating scenarios with nonlinear and multi-scale phenomena. In addition, as suggested by D. Violeau and B.D. Rogers, [94]: "SPH has been able to generate results in close agreement with reference solutions/data in validation tests, without highly sophisticated algorithms required in mesh-based schemes." Since SPH is likely to provide more promising futures in complex free-surface and very large-scale flow simulation than mesh-based methods, it is selected as the major simulation tool for RISMC analysis in flooding and high wind.

In SPH, the elements are represented as particles and the properties possessed are distributed and smoothed across a spatial distance (smoothing length) by certain rules (e.g., kernel function). Therefore, the physical quantity of any particle can be obtained by summing the relevant properties of all the particles lying within the range of kernel. In SPH, the summing of property or function f is governed by:

$$f(\vec{r}) \cong \int f(\vec{r}')W(\vec{r}-\vec{r}',h)d\vec{r}' \qquad \text{Eq. 2}$$

where h is the smoothing length and $W(\vec{r}-\vec{r}',h)$ is the kernel function of delta-function type. It weighs the properties during summing process and when it becomes a Delta-function, the left-hand side and right-hand side of Eq. 2 will be exactly the same. The smoothing function W has a finite range of κh, where κ is a constant that defines the support domain of the smoothing function. In general, the properties of the weight function can be summarized as:

$$\int W(\vec{r}-\vec{r}',h)d\vec{r}' = 1 \qquad \text{Eq. 3}$$

$$\lim_{h\to 0} W(\vec{r}-\vec{r}',h) = \delta(\vec{r}-\vec{r}') \qquad \text{Eq. 4}$$

$$W(\vec{r}-\vec{r}',h) = 0 \; when \; |\vec{r}-\vec{r}'| > \kappa h \qquad \text{Eq. 5}$$

Eq. 3 represents the normalization or unity property, Eq. 4 represents the Delta function property and Eq. 5 represents the compact condition. If the equation is further approximated with particles, which means to represent the problem domain with a set of particles and estimate the field variables based on these particles. As a result, the equation becomes:

$$f(\vec{r}) \cong \sum_b f\left(\vec{r}_b \quad\right) W\left(\vec{r} - \overrightarrow{r_b} \quad, h\right) \Delta V_b \qquad \text{Eq. 6}$$

where b represents any discrete region within the affecting region. If each discrete region has mass, which can be represented as ρdV, the equation will then be written as:

$$f(\overrightarrow{r_a}) \cong \sum_b \frac{m_b}{\rho_b} f\left(\vec{r}_b \quad\right) W\left(\overrightarrow{r_a} - \vec{r}_b \quad, h\right) \qquad \text{Eq. 7}$$

m_b and ρ_b represents the mass and density of each particle. Eq. 7 is usually referred as "summation interpolation", which is the basis of all SPH formalisms. Note that mass and density are introduced during the particle approximation. This makes SPH very suitable for application in hydrodynamics problems where mass and density are very important parameters to determine. Similar technique can also be applied for approximating the spatial derivatives.

$$\nabla \cdot f(\vec{r}) \cong \int [\nabla \cdot f(\vec{r}')] W(\vec{r} - \vec{r}', h) d\vec{r}' \qquad \text{Eq. 8}$$

Based on the divergence product rule and the divergence theorem, Eq. 8 can be written as:

$$\begin{aligned} \nabla \cdot f(\vec{r}) &= \int [\nabla \cdot f(\vec{r}')] W(\vec{r} - \vec{r}', h) d\vec{r}' \\ &= \int \nabla \cdot [f(\vec{r}') W(\vec{r} - \vec{r}', h)] d\vec{r}' - \int f(\vec{r}') \cdot \nabla W(\vec{r} - \vec{r}', h) d\vec{r}' \qquad \text{Eq. 9} \\ &= \oint f(\vec{r}') W(\vec{r} - \vec{r}', h) \cdot \vec{\boldsymbol{n}} d\vec{r}' - \int f(\vec{r}') \cdot \nabla W(\vec{r} - \vec{r}', h) d\vec{r}' \end{aligned}$$

For the surface integral derived from divergence theorem, because the smoothing function W is defined to be zero at the surface (Eq. 5), the first term vanishes. With particle discretization technique, the equation will become Eq. 10, which is also known as the symmetric [95] form for divergence calculation.

$$\nabla \cdot f(\vec{r}) = \sum_b [f(\vec{r}_b) - f(\vec{r})] \cdot \nabla W\left(\vec{r} - \overrightarrow{r_b}\ , h\right) \Delta V_b \qquad \text{Eq. 10}$$

Another anti-symmetric notation is also proposed by placing the density inside the divergence operator as shown in Eq. 11. The calculation by anti-symmetric formulation is more stable. However, the convergence rate of anti-symmetric equation is not even 0th order, which will greatly deteriorate the accuracy.

$$\begin{aligned} \nabla \cdot f(\vec{r}) &= \frac{1}{\rho}\left[\int \nabla \cdot [\rho f(\vec{r}\,') W(\vec{r} - \vec{r}\,', h)] d\vec{r}\,' - \int f(\vec{r}\,') \cdot \nabla \rho \, d\vec{r}\,'\right] \\ &= \frac{1}{\rho} \sum_b m_b \left[\left(\frac{f(\vec{r}_b)}{\rho_b^2}\right) - \left(\frac{f(\vec{r})}{\rho_a^2}\right)\right] \cdot \nabla W\left(\vec{r} - \overrightarrow{r_b}\ , h\right) \end{aligned} \qquad \text{Eq. 11}$$

However, it is noted that when the continuum equation is discretized as in Eq. 7, the functions are assumed to be smoothly varying on the smallest length scale, which is the smoothing length h. This means that discontinuities on such scales are not resolved by the numerical method. Though SPH naturally preserves mass and momentum, which grants SPH unique power in dealing with complex interfaces and shocks without mass loss, such assumption also causes many challenges. For the application in high wind and flooding, the major difficulties include:

- Lack of consistent and reliable theory in mathematical foundation (error analysis, convergence, stability) [96] [94];
- Inaccurate in pressure prediction (fluctuation, particle vacancies) [97];
- Long computational time for 3D large scale simulation;
- Difficult in dealing with variable particle resolution for incompressible flows.

Issue a) and b) is caused by the Lagrangian formulation of SPH. The discretization error is extremely hard to analyze because of the disordered particle configurations once the simulation starts. However, such disorder is not purely random, it is determined by the discrete equations of motion. Currently, there is no consistent method available for such error analysis. And such deficiency can cause certain SPH scheme to deteriorate to 0^{th} order of accuracy or worse [98] [99] [100]. Because of such inconsistency, SPH cannot correctly reproduce even a constant (zeroth-order) field, unless the particles obey exactly certain very specific geometric arrangements. Many works have been done to increase the order of accuracy through high-order kernel function (Wendland kernels [101]), some new derivation of SPH equation, moving-least-square [102] [103]) or manually regulating the particle distribution (artificial viscosity [104], XSPH [105] [106], Particle Shifting Algorithm [107] [108]). Table 3-1 shows a summary of some well-known SPH treatments. Note that many incompressible SPH (ISPH) solver is proposed recently. Avoiding Equation of State, ISPH is divergence-free and can maintain constant density [109] [110] [111] [112]. Comparing to WCSPH, ISPH can accurately predict the pressure field and retain more regular particle distribution. However, additional boundary treatments, including free surface and interfaces, is needed [96] and this requires more computational power.

Table 3-1: Summary of some popular SPH treatments in Computational Fluid Dynamics (CFD).

Method Name	Consistency/Order of Accuracy	Conservation (mass/energy/momentum)	Number of Neighbors (2D)	Known Troubles
Original SPH (artificial viscosity, XSPH)	0~1	Yes	~32	Poor Pressure Prediction, Instability
High-Order Kernel SPH	0~1	Yes	~128-442	Expensive, Instability
Corrected SPH (MLS, Particle Shifting)	0	No	~32	Mass Loss
ISPH (Divergent Free)	0~1	Yes	~32	Boundary Treatment (free surface)

When Eq. 7 is solved in SPH with kernel function, 30 ~ 50 particles are needed inside the region. Such amount is significantly larger than the mesh-based methods (5~13 neighboring data in FV methods). Based on Leroy's conclusion [113], SPH is more computationally expensive than mesh-based methods for single-phase flow (laminar and turbulence), while for simulation with complex surface structures or multiple-phase, SPH has shown to be both efficient and accurate.

3.2. SPH in Fluid Simulations

This section describes the theory and formulation of SPH methods in fluid simulations. The goal is to address the question of whether the SPH approximation of the fluid-dynamic equations is sufficiently accurate for its intended use. Such process is commonly known as solution verification process, and it is performed according to the assumption A1 and A2. The solution activities include: (1) assuring the accuracy of input data for the problem of interest; (2) estimating the numerical solution error; (3) assuring the accuracy of output data for the problem of interest. Section 3.2.1 to 3.2.4 introduces the derivations of continuity and momentum equations in SPH, and the objective is to ensure that SPH formulations can correctly represent the fluid dynamic models. In addition to the common formulations and capabilities, this section also introduces the development of SPH turbulence model. Section 3.3 mathematically estimates the numerical solution error, including truncation error, discretization error, and simulation error. At the same time, the accuracy of the output data is assessed for a simple laminar problem with known solutions. Finally, findings and suggestions are made in the summary remarks.

3.2.1. Continuity Equation

First, the summation interpolation (Eq. 7) leads to a simple form for the evaluation of density:

$$\rho_a = \sum_b m_b W_{ab} \qquad \text{Eq. 12}$$

where W_{ab} denotes $W\left(\vec{r_a} - \vec{r}_b \ \ , h\right)$ as in Eq. 7 If time derivative is taken for Eq. 12 based on Eq. 35:

$$\frac{d\rho_a}{dt} = \sum_b m_b(\vec{\boldsymbol{v}}_{\boldsymbol{a}} - \vec{\boldsymbol{v}}_{\boldsymbol{b}}) \cdot \nabla_a W_{ab} \qquad \text{Eq. 13}$$

where ∇_a denotes the gradient calculated with repected to the coordinates of particle a and the density derivatives can be translated back to continuum form based on summation (Eq. 7) interpolation:

$$\frac{d\rho}{dt} = \vec{\boldsymbol{v}} \cdot \nabla\rho - \nabla \cdot (\rho\vec{\boldsymbol{v}}) = -\rho(\nabla \cdot \vec{\boldsymbol{v}}) \qquad \text{Eq. 14}$$

Both Eq. 12 and Eq. 14 are SPH expressions for the continuity equation, and Eq. 14 can also be derived from the Navier-Stoke mass conservation equation. When used with Equation of State, Eq. 12 will create an artificial density drop near the free surface because no all support regions are filled with particles. NEUTRINO uses Eq. 12 and implicitly solves Poisson equation, which avoids the density deficiency near the boundary, but increases the computational cost.

3.2.2. Laminar Flow Modeling

Since the SPH approximation is one of the scientific computing techniques for PDE solving, the SPH formulation of the Navier-Stoke momentum conservation for compressible flow can be found as:

$$\frac{D(\rho u_i)}{Dt} = -\frac{1}{\rho}\frac{\partial\left(p\delta_{ij}\right)}{\partial x_i} + \frac{\partial\sigma_{ij}}{\partial x_j} \qquad \text{Eq. 15}$$

where σ_{ij} is given by

$$\sigma_{ij} = 2\mu S_{ij} - \frac{2}{3}\mu\delta_{ij}S_{kk} \quad \text{Eq. 16}$$

The strain rate tensor S_{ij} is given by

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \quad \text{Eq. 17}$$

For laminar flow, where fluids flow in parallel layer and there is no disruption between layers [114], the gradients perpendicular to the direction of flow is almost zero. As a result, the momentum equation (Eq. 15) can be simplified as:

$$\frac{D(\rho u_i)}{Dt} = -\frac{1}{\rho}\frac{\partial(p\delta_{ij})}{\partial x_i} + \nu\frac{\partial^2 u_i}{\partial x_i \partial x_j} \quad \text{Eq. 18}$$

With SPH approximation as Eq. 17, continuity and momentum equation can be approximated as:

$$\frac{d\vec{v}_a}{dt} = -\sum_b m_b\left(\frac{P_b}{\rho_b^2} + \frac{P_a}{\rho_a^2}\right)\vec{\nabla}W_{ab} + \sum_b m_b\left(\frac{4\nu\vec{r}_{ab}\vec{\nabla}W_{ab}}{(\rho_a + \rho_b|\vec{r}_{ab}|^2)}\right)\vec{v}_{ab} \quad \text{Eq. 19}$$

Note that though SPH employs compressible formulation, the density variation is required to be lower than 3% [115]. Therefore, SPH is claimed to be weakly compressible and the value of $\frac{D\rho}{Dt}$ is almost zero.

3.2.3. Turbulent Flow Modeling

In the early applications of SPH to fluid dynamics, turbulence is ignored because SPH is found to be appropriate for representing violent flows, where inertia force plays a major role. Therefore, SPH has been widely and successfully applied in flooding analysis. However, when SPH is applied to confined flows, turbulence model starts to be addressed. For the RISMC analysis of high wind, where very large velocity (>30m/s) and complex surface structure is involved without free surface, a turbulence model is needed for both accuracy and efficiency consideration. Though SPH is capable of modeling the turbulence flow with a very refined resolution [97] [116]

[117], it is only applied in a restricted area. Also, because SPH is way more computationally expensive than mesh-based methods as a DNS tool, the meaning of such simulation is limited. In SPH, each particle carries scalar information including density, pressure, velocity components, etc., and such information is assumed to be homogeneous "inside" particles. Note that such interpolation is similar to the LES filter, it is straightforward to apply Sub-particle scaling (SPS) method to SPH for turbulence modeling, where governing equations are spatially averaged over a length scale comparable to the computational elements. However, no correlation has been found that could relate the numerical diffusion and turbulence viscosity in SPH. Similar to the coarse-mesh CFD tool, this case shows the importance of parameter and model selection in SPH. As a result, for large-scale eddies, resolved by the grid or particle sizes, the averaged equations are sufficient to solve for these, for the smaller turbulent eddies, smaller than the particle size, a closure scheme is needed to model their effects on the flow field. With a formulation very similar to original WCSPH, additional turbulence viscosity term is introduced with SPS technique [95]. Though some vortexes are claimed to be captured, classical WCSPH formulation shows tremendous noise and fluctuations [95] [113]. It is also found to be difficult in dealing with wall boundary layers with such techniques due to the incorrect reproduction of velocity-pressure interactions [118]. Some other efforts, including Stochastic models [119], LES models with particle regulation [120] and Reynolds-averaged Navier-Stokes (RANS) [113] [121] [122] have been made. Except for the stochastic models with Generalized Langevin Model (GLM), other models show certain improvement compared to original LES model, especially in predicting the velocity in near-wall region. Table 3-2 shows a summary of these popular SPH turbulence model. Compared to LES and Stochastic model, RANS models have the best accuracy, and the wall function plays a crucial role in the prediction accuracy.

Table 3-2: Review and comparison of LES and RANS k- ε model applied in SPH.

Leading Work	Arai, J., et al. [123]	Leroy, A. [113]	Violeau, D., et al. [124]	Welton, W.C., et al. [119]
Technique	LES Filter	RANS		Stochastic Models
Mechanism	Filter	Time/Space Averaging		Langevin equations
Models Applied	Smargorinsky Model	$k-\epsilon$ model; Buoyancy.	Mixing Length	GLM
Special Boundary Treatment	Wall function	Wall function; Discrete Laplacian boundary term.	NA	NA

In general, SPH is capable of simulating problems with multiphase and complex solid structures, the computational time and accuracy are supreme compared to mesh-based methods. Therefore, SPH can serve as a simulation tool for the mechanistic study of RISMC, especially for the scenarios of external events. However, simulation parameters need to be carefully varied such that the accuracy of QoIs can be ensured. Besides, additional models may be needed.

3.2.4. Fluid-Structure Interactions

Nowadays, the major simulation tools for solid mechanism are mostly based on Finite Element Methods, while SPH has been applied to some nonlinear mechanics [125] [126], like fractures, cracks, and collisions. At the same time, the SPH method is popular in shell formation, deformation and fracture [127] [128], where large deformation occurs. In this study, BulletPhysics [129] is applied as the solver of Solid Mechanics and two-way coupling is achieved between Neutrino IISPH solver and Bullet solver. IISPH solver will calculate the pressure force exerted on solids by SPH particles and transfer it into Bullet Solver for solid motion. At the same time, Bullet will also feed updated information back to SPH. In this study, all moving solids are treated as rigid bodies without deformations or fractures. It's also found from various literature that SPH method is accurate in estimating forces and predicting floating-body movement [97] [130].

In NEUTRINO, the forces between solid and fluid particles are calculated by fluid pressure. It is assumed that the applied pressure from fluid to the rigid does not have any kinematic influence

on the nearby fluid particles. Eq. 20 shows the equation for calculating pressure force applied from the fluid to the rigid. The symmetric pressure force from the fluid to the rigid will then have the same value, but opposite direction.

$$F_{fluid \to rigid} = m_{fluid} \cdot \rho_0 V_{rigid} \left(\frac{P_{fluid}}{\rho_{fluid}^2} \right) \nabla W_{fluid-rigid} \qquad \text{Eq. 20}$$

where the volume of the rigid is estimated by Eq. 21.

$$V_{rigid} = \frac{m_{rigid}}{m_{rigid} \sum_k W_{rigid-k}} \qquad \text{Eq. 21}$$

where *k* represents the neighbors of that rigid particle. This approach avoids the stability issues when there are large density ratios across the interface [131]. However, this approach is found to be very sensitive to the selection of particle size, which will be shown in later chapter.

3.3. Theoretical Error Analysis

As mentioned previously, SPH approximation is mainly composed by two parts: Interpolation (truncation Eq. 2) and Discretization (Eq. 6). The main purpose of interpolation is to build the potential function based on continuous equation. The discretization is to assign properties to each single particle. With properly assigned properties and potential function that specify the interaction between each particle, summation is made to determine the evaluation of the particles and properties.

3.3.1. Truncation Error Analysis

First of all, if we have an arbitrary scalar (or vector) field represented as a spatial integral over the Dirac distribution:

$$A(\boldsymbol{r},t)=\int_{\Omega}A(\boldsymbol{r}',t)\delta(\boldsymbol{r}-\boldsymbol{r}')d\boldsymbol{r}' \qquad \text{Eq. 22}$$

Where Ω represents the whole continuous region. In real application, the Dirac distribution is usually approximated by a finite function, like Gaussian or any other Gaussian-like function with a compact support. Then the equation is written as:

$$A(\boldsymbol{r},t)\approx A_c(\boldsymbol{r},t)=\int_{\Omega}A(\boldsymbol{r}',t)w_h(\boldsymbol{r}-\boldsymbol{r}')d\boldsymbol{r}' \qquad \text{Eq. 23}$$

Where w_h represents the finite function with parameter h indicating its supporting length. For Gaussian type function, h will be infinite, while for Gaussian-like function with a compact support, h can be any value or 0. In order to assess the accuracy of $A_c(\boldsymbol{r},t)$ comparing to $A(\boldsymbol{r},t)$, the Taylor series expansion of $A(\boldsymbol{r}',t)$ is written around the point $\boldsymbol{r}$ to second order term:

$$A(\boldsymbol{r}',t)=A(\boldsymbol{r},t)-\frac{\partial A}{\partial \boldsymbol{r}}\cdot\Delta\boldsymbol{r}+\frac{1}{2}\Delta\boldsymbol{r}\frac{\partial^2 A}{\partial \boldsymbol{r}^2}\Delta\boldsymbol{r}+O(|\Delta\boldsymbol{r}|^3) \qquad \text{Eq. 24}$$

Where $\Delta\boldsymbol{r}$ is defined as $(\boldsymbol{r}-\boldsymbol{r}')$. Substituting Eq. 24 into Eq. 23, we obtain:

$$\begin{aligned}A_c(\boldsymbol{r},t)\approx A(\boldsymbol{r},t)\int_{\Omega}w_h(\Delta\boldsymbol{r})d\boldsymbol{r}'-\frac{\partial A}{\partial \boldsymbol{r}}\cdot\int_{\Omega}\Delta\boldsymbol{r}w_h(\Delta\boldsymbol{r})d\boldsymbol{r}'\\+\frac{1}{2}\frac{\partial^2 A}{\partial \boldsymbol{r}^2}\int_{\Omega}\Delta\boldsymbol{r}\cdot\Delta\boldsymbol{r}w_h(\Delta\boldsymbol{r})d\boldsymbol{r}'+\int_{\boldsymbol{\Omega}}\boldsymbol{O}(|\Delta\boldsymbol{r}|^3)w_h(\Delta\boldsymbol{r})d\boldsymbol{r}'\end{aligned} \qquad \text{Eq. 25}$$

In order to obtain $A_c(\boldsymbol{r},t)\approx A(\boldsymbol{r},t)$, it is required that:

$$\int_{\Omega}w_h(\Delta\boldsymbol{r})d\boldsymbol{r}'=1 \qquad \text{Eq. 26}$$

$$\int_{\Omega}\Delta\boldsymbol{r}w_h(\Delta\boldsymbol{r})d\boldsymbol{r}'=0 \qquad \text{Eq. 27}$$

Eq. 26 indicates that the zeroth-order moment of kernel function should be 1. Eq. 27 indicates that the kernel first-order moment should be zero. Besides, if an even kernel is selected, then the kernel function should be symmetric:

$$w_h(\boldsymbol{r}) = w_h(-\boldsymbol{r}) \quad \text{Eq. 28}$$

$$\nabla w_h(-\boldsymbol{r}) = -\nabla w_h(\boldsymbol{r}) \quad \text{Eq. 29}$$

Also, from Eq. 28, the kernel moment of all odd orders is identically zero:

$$\int_{\Omega} \Delta\boldsymbol{r} \cdot (\dots 2m+1 \dots) \cdot \Delta\boldsymbol{r} w_h(\Delta\boldsymbol{r}) d\boldsymbol{r}' = \boldsymbol{0} \quad \text{Eq. 30}$$

Therefore, the error due to interpolation can be obtained as:

$$A_c(\boldsymbol{r}, t) = A(\boldsymbol{r}, t) + O(h^2) \quad \text{Eq. 31}$$

If we apply gradient to Eq. 31, we can also obtain the error for gradient operator:

$$(gradA_c)(\boldsymbol{r}, t) = (gradA)(\boldsymbol{r}, t) + O(h^2) \quad \text{Eq. 32}$$

On the other hand, we can apply Eq. 24 to following equation and obtain the same form of Eq. 32:

$$\begin{aligned}(gradA_c)(\boldsymbol{r}, t) &= \int_{\Omega} \frac{\partial A(\boldsymbol{r}', t)}{\partial \boldsymbol{r}'} w_h(\Delta\boldsymbol{r}) d\boldsymbol{r}' \\ &= \int_{\Omega} \frac{\partial}{\partial \boldsymbol{r}'} \cdot [A(\boldsymbol{r}', t) w_h(\Delta\boldsymbol{r}) d\boldsymbol{r}' - \int_{\Omega} A(\boldsymbol{r}', t) \frac{\partial w_h(\Delta\boldsymbol{r})}{\partial \boldsymbol{r}'} d\boldsymbol{r}'\end{aligned} \quad \text{Eq. 33}$$

$$(gradA_c)(\boldsymbol{r}, t) = \oint_{\partial\Omega} \frac{\partial}{\partial \boldsymbol{r}'} \cdot [A(\boldsymbol{r}', t) w_h(\Delta\boldsymbol{r}) \boldsymbol{n}(\boldsymbol{r}') d\boldsymbol{r}' + \int_{\Omega} A(\boldsymbol{r}', t) \frac{\partial w_h(\Delta\boldsymbol{r})}{\partial \boldsymbol{r}'} d\boldsymbol{r}' \quad \text{Eq. 34}$$

Because the kernel function w_h usually has finite support radius and equal to zero at surface, the contour integral vanishes and only term with gradient of kernel function is left. If we insert Eq. 24 to Eq. 34,

$$\begin{aligned}(gradA_c)(\boldsymbol{r},t) \\ &= A(\boldsymbol{r},t)\int_\Omega \frac{\partial w_h(\Delta\boldsymbol{r})}{\partial \boldsymbol{r}'}d\boldsymbol{r}' - \frac{\partial A}{\partial \boldsymbol{r}}(\boldsymbol{r},t)\int_\Omega \frac{\partial w_h(\Delta\boldsymbol{r})}{\partial \boldsymbol{r}'}\cdot\Delta\boldsymbol{r}d\boldsymbol{r}' \\ &+\frac{1}{2}\frac{\partial^2 A}{\partial \boldsymbol{r}^2}(\boldsymbol{r},t)\int_\Omega \frac{\partial w_h(\Delta\boldsymbol{r})}{\partial \boldsymbol{r}'}\cdot\Delta\boldsymbol{r}\cdot\Delta\boldsymbol{r}d\boldsymbol{r}' + \int_\Omega \frac{\partial w_h(\Delta\boldsymbol{r})}{\partial \boldsymbol{r}'}O(\Delta\boldsymbol{r}^3)d\boldsymbol{r}'\end{aligned} \quad \text{Eq. 35}$$

In a second-order accuracy is expected, following relations are required:

$$\int_\Omega \nabla w_h(\Delta\boldsymbol{r})d\boldsymbol{r}' = 0 \quad \text{Eq. 36}$$

$$\int_\Omega \nabla w_h(\Delta\boldsymbol{r})\cdot\Delta\boldsymbol{r}d\boldsymbol{r}' = -\boldsymbol{I_n} \quad \text{Eq. 37}$$

$$\int_\Omega \nabla w_h(\Delta\boldsymbol{r})\cdot\Delta\boldsymbol{r}\cdot\Delta\boldsymbol{r}d\boldsymbol{r}' = 0 \quad \text{Eq. 38}$$

where $\boldsymbol{I_n}$ is the dimensional identity tensor. Eq. 36 and Eq. 38 can be derived from Eq. 29, what is more, any even order moments of the kernel gradient equal zero. As for Eq. 37, we first apply integral separation, then use Gauss theorem and the fact that surface integral vanishes due to infinite support region of kernel function to derive such relation.

$$\begin{aligned}\int_\Omega \frac{\partial w_h(\Delta\boldsymbol{r})}{\partial \boldsymbol{r}'}\times\Delta\boldsymbol{r}d\boldsymbol{r}' &= \int_\Omega \frac{\partial}{\partial \boldsymbol{r}'}\cdot[w_h(\Delta\boldsymbol{r})\Delta\boldsymbol{r}]d\boldsymbol{r}' - \int_\Omega w_h(\Delta\boldsymbol{r})\left(\frac{\partial\Delta\boldsymbol{r}}{\partial\Delta\boldsymbol{r}}\right)d\boldsymbol{r}' \\ &= \oint_{\partial\Omega}[w_h(\Delta\boldsymbol{r})\Delta\boldsymbol{r}\cdot\boldsymbol{n}(\Delta\boldsymbol{r})]d\boldsymbol{r}' - \int_\Omega w_h(\Delta\boldsymbol{r})\left(\frac{\partial\Delta\boldsymbol{r}}{\partial\Delta\boldsymbol{r}}\right)d\boldsymbol{r}' = -\boldsymbol{I_n}\end{aligned} \quad \text{Eq. 39}$$

For the last term of Eq. 35, because of the integral, the third order term becomes second order. The interpolation error of gradient operator can be obtained:

$$(gradA_c)(\boldsymbol{r},t) = \frac{\partial A}{\partial \boldsymbol{r}}(\boldsymbol{r},t) + O(h^2) \qquad \text{Eq. 40}$$

3.3.2. Discretization Error Analysis

The discretization is to distribute particles inside a support region with properties and potential functions. Though the particles may behave like material points, but they are actually macroscopic entities carrying kinematic and thermodynamic properties. Each particle has a fixed density and volume and all properties are assumed to be homogeneous inside a particle. Taking a 2-D example as in Figure 3-1, we take any one of the particle volume to be V_p, then the mean diameter of a particle can be determined as: $\delta r = V_p^{1/2}$. Notice that δr also stands for the unit grid size. Another important parameter is $h/\delta r$, with h defined as smoothing length of the kernel in previous section. In most current application, this parameter depends on the selection of kernel function and usually equals to 1 to 2. As the particles evolve, the properties assigned to each particle are also subject to a variation. Noted that such variation is purely Lagrangian, there is no advection term.

The discretization operator is based on Riemann sum and in application, the summation points are finite number of adjacent particles b.

$$\int_{\Omega} C(\boldsymbol{r}',t)d\boldsymbol{r}' \approx \sum_b C(\boldsymbol{r}_b,t)V_b \qquad \text{Eq. 41}$$

If we apply this method to the continuous approximation done before (Eq. 23), we can get:

$$A_c(\boldsymbol{r},t) = \int_{\Omega} A(\boldsymbol{r}',t)w_h(\boldsymbol{r}-\boldsymbol{r}')d\boldsymbol{r}' \approx A_d(\boldsymbol{r},t) = \sum_b A(\boldsymbol{r_b},t)w_h(\boldsymbol{r}-\boldsymbol{r_b})V_b \qquad \text{Eq. 42}$$

Where A_d defines the discrete interpolation of the region. Note the sum covers all the particles inside the support domain and the size of the domain is determined by the radius of kernel function.

For a kernel function with support radius as $h_t = K \cdot h$, we can calculate the number of particles N inside the support domain as:

$$N = \frac{V_n h_t^2}{V_p} \qquad \text{Eq. 43}$$

Where V_n is the volume of the unit 2D circle (2π). For some popular kernel function applied these days, magnitude of K could equal to 3 (B-Spline M=5), 2.5 (B-Spline M=4), 2 (Wendland) or other values. As a result, we need 30 ~ 50 particles inside the region. This is significantly larger than the mesh-based methods. And this is considered to be a weakness in SPH method. For example, in Finite Volume methods, solving a second order differential equation will only need 5-13 neighboring data.

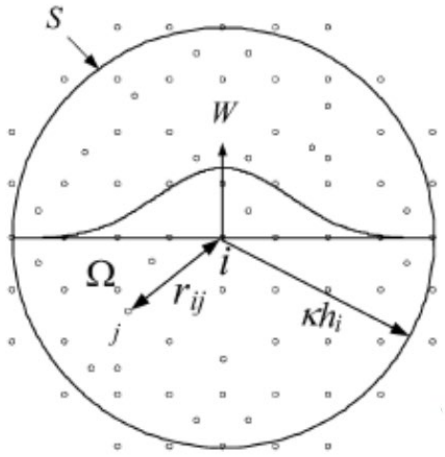


Figure 3-1: Demonstration of particles inside one support region with radius h_t.

Same analogy can be applied to gradient vector:

$$[gradA]_c(\boldsymbol{r}, t) \approx [gradA]_d(\boldsymbol{r}, t) = \sum_b A(\boldsymbol{r_b}, t) \nabla w_h(\boldsymbol{r} - \boldsymbol{r_b}) V_b \qquad \text{Eq. 44}$$

However, estimating the error made by the discrete operator A_d is complicated because it depends on the particle positions, and the particles will distribute in a disordered pattern. However, such

disorder is not completely random cloud and approaches of estimating theoretical mean are not applicable. Researcher has proposed that because such disorder is deterministically determined by discrete equations and if such disorder can be measured by some quantities, the assessment of the accuracy of the discrete interpolation may become possible [132]. In this work, we only estimate the error at the beginning of the simulation, where all particles are placed at the grid point of Cartesian grid [96]. If we assume the discretization error is represented as:

$$E_d[A(\boldsymbol{r_b})w_h(\boldsymbol{r}-\boldsymbol{r_b})] = \sum_b A(\boldsymbol{r_b},t)w_h(\boldsymbol{r}-\boldsymbol{r_b})V_b - \int_\Omega A(\boldsymbol{r}',t)w_h(\boldsymbol{r}-\boldsymbol{r}')d\boldsymbol{r}' \qquad \text{Eq. 45}$$

Then we translate the system (not affecting error Eq. 46) and apply Fourier series transformation with derivatives considered at point $\boldsymbol{r}$ (Eq. 47):

$$E_d[A(\boldsymbol{r_b})w_h(\boldsymbol{r}-\boldsymbol{r_b})] = E_d[A(\boldsymbol{r}-\boldsymbol{r_b})w_h(\boldsymbol{r_b})] \qquad \text{Eq. 46}$$

$$A(\boldsymbol{r}-\boldsymbol{r_b}) = A(\boldsymbol{r}) - \frac{\partial A}{\partial \boldsymbol{r}}\cdot\boldsymbol{r_b} + \frac{1}{2}\boldsymbol{r_b}\frac{\partial^2 A}{\partial \boldsymbol{r^2}}\boldsymbol{r_b} + O(|\boldsymbol{r_b}|^3 \qquad \text{Eq. 47}$$

Then we apply Eq. 47 to the error term and get:

$$\begin{aligned} &E_d[A(\boldsymbol{r_b})w_h(\boldsymbol{r}-\boldsymbol{r_b})] \\ &\quad = A(\boldsymbol{r})E_d[w_h(\boldsymbol{r_b})] - \frac{\partial A}{\partial \boldsymbol{r}}E_d[w_h(\boldsymbol{r_b})\cdot\boldsymbol{r_b}] \\ &\quad + \frac{1}{2}\frac{\partial^2 A}{\partial \boldsymbol{r^2}}E_d[w_h(\boldsymbol{r_b})\cdot\boldsymbol{r_b}\cdot\boldsymbol{r_b}] + E_d[O(|\boldsymbol{r_b}|^3)w_h(\boldsymbol{r_b})] \end{aligned} \qquad \text{Eq. 48}$$

As we mentioned before in Eq. 26, an odd function $(w_h(\boldsymbol{r_b})\cdot\boldsymbol{r_b})$ will have a zero-error term if all particles are regularly distributed on the Cartesian grid. As for $E_d[w_h(\boldsymbol{r_b})]$ and $E_d[w_h(\boldsymbol{r_b})\cdot\boldsymbol{r_b}\cdot\boldsymbol{r_b}]$ term, Eq. 49, proposed by Violeau, D [133], is used with k_i as non-zero integer.

$$E_d[C(\cdot)] = A\sum_{k_i}\hat{C}(k_1K_{\delta r},\ldots,k_nK_{\delta r}) \qquad \text{Eq. 49}$$

Where $\hat{C}$ is the Fourier transform of the function C and $K_{\delta r}$ is the fundamental wavenumber that is defined as $K_{\delta r} = 2\pi h/\delta r$. Finally, we can write Eq. 48 as:

$$E_d[A(\boldsymbol{r_b})w_h(\boldsymbol{r}-\boldsymbol{r_b})] = A(\boldsymbol{r})n\widehat{w_h}(K_{\delta r}) - \frac{1}{2}\frac{\partial^2 A}{\partial \boldsymbol{r^2}}\frac{\partial^2 \widehat{w_h}}{\partial K^2} + O(h^3) \qquad \text{Eq. 50}$$

Therefore, the relative error of $E_d[A(\boldsymbol{r_b})w_h(\boldsymbol{r}-\boldsymbol{r_b})]/A(\boldsymbol{r})$ is subject to an approximation of the order of h^2. Besides, it is important to note that $\widehat{w_h}(K_{\delta r})$, which is the kernel Fourier transform, plays a crucial role in the error analysis. It can be concluded that a higher order of $h/\delta r$ will give a lower $\widehat{w_h}(K_{\delta r})$ and a smaller error [106].

3.3.3. Simulation Error Analysis

In this section, SPH simulations are carried out to determine the modeling error and compare the observed error to theoretical error. Simple Poiseuille flow between stationary infinite plates at y = 0 and y = L is simulated. The fluid is initially set to be stationary and driven by a body force F parallel to x-axis. The results are compared to analytical solution, which has a parabolic velocity profile and has the form of Eq. 51.

$$u(y) = \frac{FL^2}{8\nu}\left(1 - \left(y - \frac{L}{2}\right)^2 \Big/ \left(\frac{L}{2}\right)^2\right) \qquad \text{Eq. 51}$$

where ν is the kinematic viscosity. In this simulation, flow with $\nu = 0.1m^2/s$, $L = 1m$, $F = 0.2m/s^2$, $\rho = 1000\ kg/m^3$ is set. The maximum velocity is found to be $2.5m/s$ and the Reynold number is calculated as 25 (Eq. 52). The particle size is enlarged or refined to determine the convergent rate.

$$Re = \frac{uL}{\nu} \qquad \text{Eq. 52}$$

Figure 3-2 shows the comparison plot of SPH simulation and analytical solution. A pretty good agreement is found. Besides, the L_2 relative error norm is calculated by Eq. 53 to determine the error comparing to exact solution.

$$L_2 = \sqrt{\frac{\sum_{i=1}^{N}\left(u_{SPH}(i) - u_{EXACT}(i)\right)^2}{\sum_{i=1}^{N} u_{EXACT}^2(i)}} \qquad \text{Eq. 53}$$

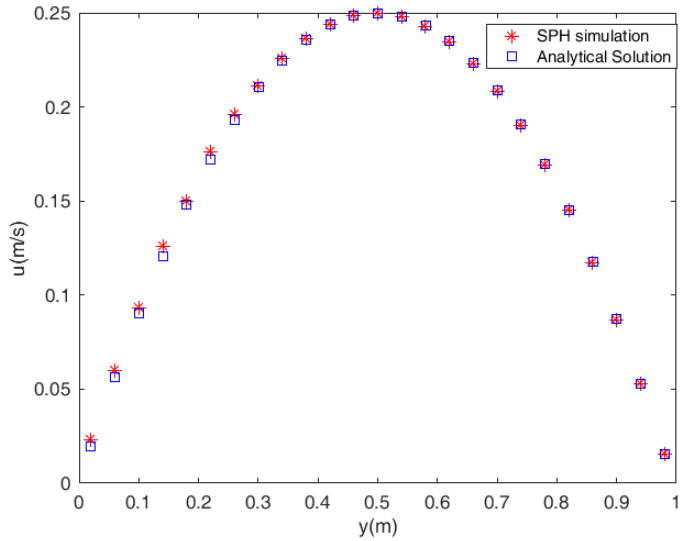


Figure 3-2: Comparison of SPH simulation and analytical solution of Poiseuille flow.

Figure 3-3 shows the observed convergence plot of SPH simulation. the slope of fitting curve is found to be 1.088 and is pretty close to 1. From previous section, the theoretical rate is 2 for both interpolation and discretization. However, the discretization error analysis only applies at initial state of simulation when particles are distributed regularly on grid points. Besides, the artificial viscosity and compressibility will introduce additional error. Fatehi has pointed that [134] theoretical order of accuracy of leading term in viscosity formulation of Eq. 19 is between first and second order and their simulation also show a first order convergent rate.

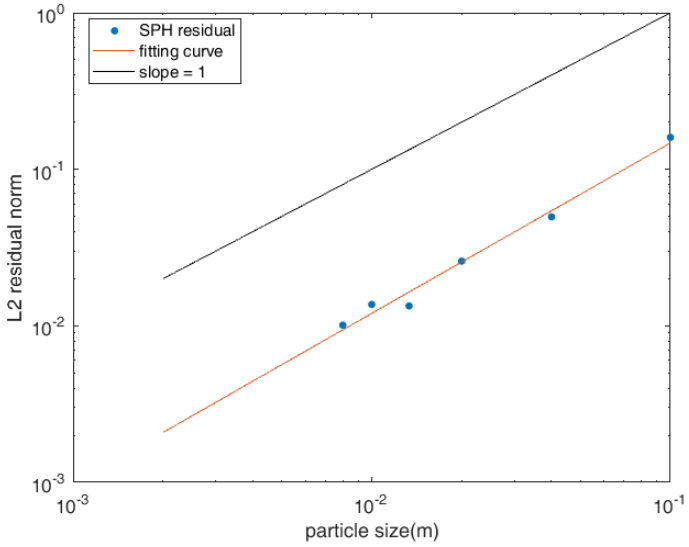


Figure 3-3: Plot of convergent rate against particle size in SPH simulation.

3.4. Summary Remarks

This chapter first reviews the state of the art for the SPH method and analyzes the theoretical error for SPH methods. The goal is to identify the source of error for SPH simulations. At the same time, this chapter can be used as an input for the decision regarding the adequacy of the SPH methods for an external-flooding and high-wind application. It is found that SPH methods have a rigorous mathematical basis as a PDE solving algorithm, and the Lagrangian operator has a second order of accuracy for approximating an arbitrary scale (or vector) field and gradients of any order. Also, the particle approximation is tightly related to the law of thermodynamics, and the equation of motions can be derived directly from the second law of thermodynamics. Most errors come from the discretization process, and the math turns to be more complicated than mesh-based methods. Meanwhile, a theoretical derivation is performed given that the particles are uniformly distributed, and the second order of accuracy can be achieved with SPH approximations. However, it is found from a simple laminar scenario, where particle distributions are relatively uniform, that SPH methods with WSPH formulation have the first order of accuracy. As a result, for violent flows with “chaotic” particle distributions, the discretized field starts to fluctuate, and order of accuracy will degrade, especially for high-order operators like viscous and pressure terms. Note that SPH can achieve good accuracy with very fine particle sizes, its computational cost is

much higher since the number of neighboring elements is 3-10 times more than the mesh-based simulations. Therefore, SPH's performance is expected to be bad for estimating turbulent phenomena, like boundary layers or vortex shedding.

However, SPH is found to be successful in "large-scale" scenarios with very complex interface structures and violent flows, and the computational costs are far less than mesh-based simulations. Such findings are results of SPH's numerical dissipations, which is known as implicit large eddy simulations in conventional CFD. In consequence, the SPH method is found to be the only applicable method for the large-scale external-flooding and high-wind simulations. To confidently apply SPH's simulation results to nuclear safety analysis, the credibility of SPH needs to be analyzed systematically for the intended uses. In this study, the term "credibility" instead of "accuracy" is used since it aligns with the definition of validation as a decision-making process. In addition, the term credibility suggests the importance of acceptance criteria and the role of decision-makers. And the adequacy of simulation tools should be decided according to the safety goal and risk levels.

4. SPH CREDIBILITY ASSESSMENT WITH EMDAP FRAMEWORK

This chapter aims to validate the SPH methods based on EMDAP framework, and such process will lead to a decision regarding the adequacy of the SPH methods for a specific application. It has been discussed in previous chapter that the objective of validation is to assess the credibility of a (collection of) model(s), and for many nuclear safety-related applications, the credibility (sometimes known as fidelity) is subjectively defined. However, neither CSAU or EMDAP has formalized or explicitly discussed the subjective components of model credibility (or fidelity). For clarification, this chapter defines credibility as accuracy that is characterized by validation metrics only. The process, demonstrate in Figure 4-1, involves four key activities:

- Identify the importance and adequacy of math, code and data for all related systems, components, processes, and phenomena.
- Establish standards against which code models and performance can be measured
- Construct code-assessment case studies and database
- Perform code assessment by comparing the code results against the standards
- Perform scaling analysis for database in terms of their relevancy and sufficiency for the intended application

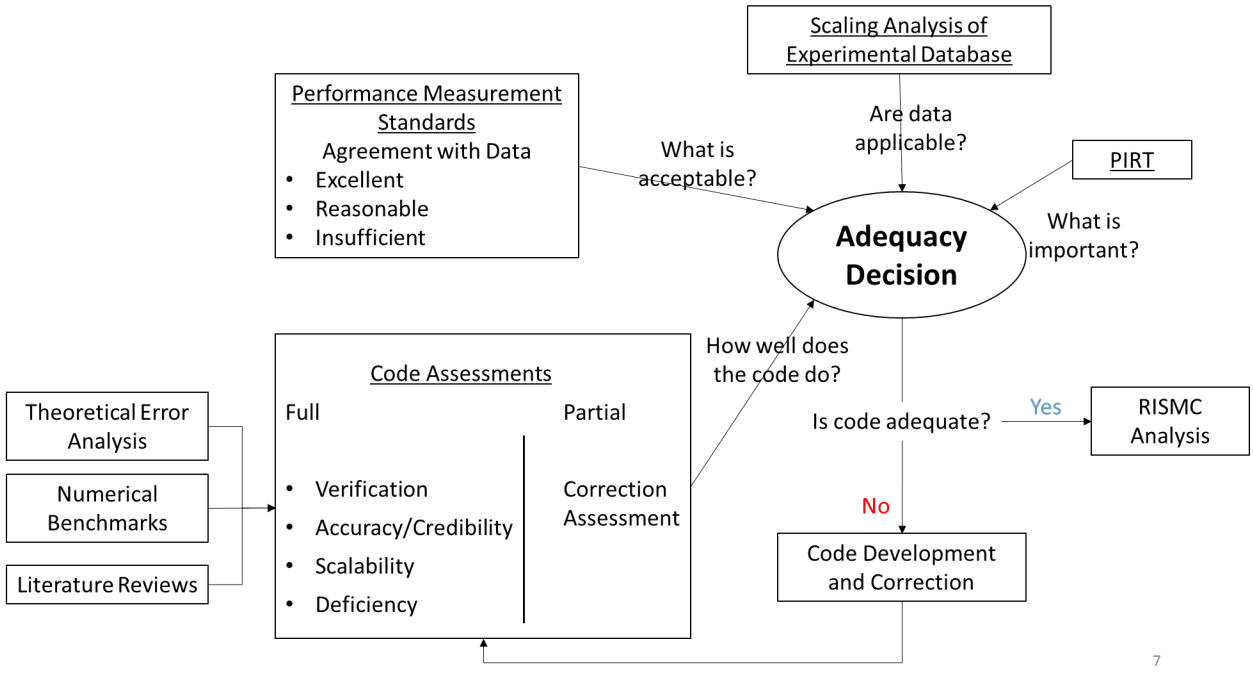


Figure 4-1: Demonstration of adequacy decision process.

Both the verification by theoretical error analysis and the literature reviews for SPH have been performed in chapter 3. The rest components will be discussed in this chapter. Finally, a summary of adequacy findings and lessons learned will be discussed. In addition, Figure 4-1 also suggests an iterative evaluation and development process, where efforts are continuously required for issues of insufficient database and imperfect models.

4.1. Establish Requirement for Evaluation Model Capability

Element 1 aims to identify and agree upon the importance of constituent phenomena, processes, and key parameters with respect to the purpose of applications. In EMDAP, Element 1 is composed by four steps: (1) Specify analysis purpose, transient class and power plant class; (2) Specify figures of merit; (3) Identify systems, components, phases, geometries, fields and processes that should be modeled; (4) identify and rank phenomena and processes. Following paragraphs will briefly discuss the progress of each step.

4.1.1. Identify and Rank Key Phenomena and Processes

In this study, scenarios are selected for the external-flooding and the high-wind application respectively. For external-flooding applications, the analysis purpose is to assess if the performance of SPH in simulating the scenario of "floods damage the building structures, enter the diesel generator (DG) room, and cause diesel generator malfunctioning" is acceptable. The corresponding FoMs include the response time and the structural loads on SSCs by floods. In this study, the response time is the time for the external floods to reach the DG building and to potentially fail the DGs. The response time is crucial for the nuclear safety analysis, including estimating the decay heat, initiating emergency procedures and identifying responses of SSCs, etc. The structural loads on stationary SSCs are important for determining the structural integrity and operational status of SSCs during external-flooding scenarios. For example, if the load exceeds the regulatory limits, certain SSCs need to be shut down; if the force exceeds the operational limits, certain SSCs can be malfunctioning and even failed. Besides, the structural loads on moving objects are important for the flooding-induced debris analysis, including the debris transport and its impacts to SSCs.

For high-wind applications, the purpose is to assess if the performance of SPH in simulating the scenario of "the long-last wind damages the offsite facilities including exhaust stack and electric wire" is acceptable. The corresponding FoMs include characteristics of the local wind fields experienced at the NPP site and the structural loads on SSCs by winds. In this study, the investigation on wind characteristics mainly includes the wind turbulence, vortex shedding, and wind generations, which are important for the structural analysis of SSCs. Besides, the wind field analysis is also important for the statistical analysis of wind-generated-missiles trajectories, which are important for the safety analysis of SSCs and plant operation personnel. Similar to the external-flooding scenario, the structural load analysis aims to determine the structural integrity and operational status of SSCs.

The Phenomenon Identification and Ranking Table (PIRT) is a ranking table to align applications' technical requirements with associated physics and materials models, simulation code capabilities and verification, validation and uncertainty quantification activities [135]. It is developed to identify and prioritize important physical phenomena in an application and to assess

the adequacy/gaps of the current simulation capabilities. The PIRT helps to ensure both sufficiency and efficiency by expert elicitation and prioritization of the analysis of the simulation and experimental capabilities. In this study, the PIRT is designed to assess the capability of SPH and NEUTIRNO in simulating the selected external-flooding and high-wind scenarios according to the evidence and expert opinion. The PIRT table contains two major attributes: **Imp**ortance and adequacy. The importance attribute aims to rank the importance of corresponding phenomena according to their effects to FoMs. Three levels of "**H**igh", "**M**edium", and "**L**ow" are used for qualitatively representing the importance, and this study only considers high-rank phenomena. The adequacy attribute is measured from three aspects: the mathematical model, computational code, and data. The math model usually refers to a mathematical representation for the phenomenon, while the code refers to a numerical reproduction for the phenomenon. The data refer to the database for existing model input and validation data. Similar to the importance ranking, there are three levels for each adequacy ranking. The notation of "**H**igh" indicates an adequate resource for a certain aspect, the "**L**ow" suggests that there are few resources available. Detailed definitions for each attribute and its ranking can be found in the literature [46] [135]. A complete PIRT process usually requires inputs from a group of experts based on their best knowledge. However, the following PIRT is tentatively performed according to author's knowledge.

For the scenario of floods breaking and entering diesel generator room, the FoMs include the response time and the structural loads on SSCs by floods, while the response time is defined as the time for the external floods to reach the DG building and to potentially fail the DGs. The response time is crucial for the nuclear safety analysis, including for estimating the decay heat, initiating emergency procedures, and identifying responses of SSCs, etc. The force exerted on SSCs is important for determining the structural integrity and operational status of SSCs during external-flooding scenarios. For example, if the force exceeds the regulatory limits, certain SSCs need to be shut down; if the force exceeds the material limits, certain SSCs can be malfunctioning and failed. To accurately predict these FoMs with Computational Fluid Dynamics (CFD) tools, the PIRT process consider all hydrodynamic phenomena, processes, and conditions (in this study, all of these are referred to as "phenomena" for brevity) that affect the FoMs. It is stressed that the PIRT process itself does not decide whether a model or code is adequate. Rather, the information in the PIRT aims to guide the evaluation of whether the code contains the essential capabilities for

modeling phenomena important for the proposed scenarios. In another word, code models that are necessary for the accurate simulation of high-rank phenomena must fully satisfy the adequacy requirements. In addition, the information of PIRT is crucial for the adequacy and safety decisions with a sufficient confidence. Moreover, the rankings of PIRT should be adaptive throughout the adequacy assessment process. In another word, If the models are demonstrated to have less impact on the simulation, they can be moved to a lower standard, such as the accuracy can partially satisfy the requirements.

Table 4-1 shows the PIRT high-rank phenomena for the FoMs of response time and structural loads, where the shaded areas represent the FoMs. Since SPH has been applied to simulate the violent flow for a long time [94], its model and code are adequate for simulating the wave generator [136] [137], wave propagation, and force estimation on static objects [97]. Comprehensive reviews for SPH methods and codes in simulating these phenomena can be found in the reference documents [138]. Therefore, most rankings for these phenomena are "**H**igh" except for the math model and code adequacy of structural loads. Although SPH methods have shown good agreements in predicting the hydrodynamic forces, the results are usually unstable, and additional smoothing techniques are usually needed. Therefore, they are ranked as "**M**edium" since these issues are considered moderate. For the phenomenon of vortex shedding, the math model and code in SPH is very limited, and they are all ranked as "**L**ow". In fact, to achieve a comparable accuracy to mesh-based simulations, SPH either needs to refine its particle to DNS scales or to be incorporated by closures. However, since the number of neighboring elements (particles) for an SPH element (particle) is typically 3-4 times more than that for a mesh cell in mesh-based simulations, SPH methods are more computationally expensive than mesh-based methods in DNS scales. Besides, since the nature of particle-based methods deviates from the mesh-based methods, most well-known closure models in established CFD studies may not be perfectly suitable for SPH. Therefore, better SPH closures are still needed for such phenomenon from the aspects of math and model adequacy. Finally, the data of full-scale solidarity wave can be obtained analytically, while the data from aerospace engineering for large-scale vortex shedding phenomenon are quite adequate. Therefore, the ranking of data adequacy for solitary wave and vortex shedding is High. But for the rest of phenomena, there is a lack of large-scale data.

Table 4-1: PIRT and lists of numerical benchmarks designed for each phenomenon in external flooding scenario.

ID	Phenomenon Description	Imp	Adequacy		
			Math Model	Code	Data
A	Response Time				
	Solitary Wave	H	H	H	H
	Wave Propagation	H	H	H	L
	Vortex Shedding	H	L	L	H
B	Structural Loads				
	Hydrodynamic force on stationary structures	H	M	M	L
	Hydrodynamic force on moving structures	H	M	M	L

For the high wind scenario, the corresponding FoMs include characteristics of the local wind fields experienced at the NPP site and the structural loads on SSCs by winds. In this study, the investigation on wind characteristics mainly includes the wind turbulence, vortex shedding, and wind generation, which are important for the structural analysis of SSCs. Besides, the wind field analysis is also important for the statistical analysis of wind-generated-missiles trajectories, which are important for the safety analysis of SSCs and plant operation personnel. Similar to the external-flooding scenario, the structural load analysis aims to determine the structural integrity and operational status of SSCs during high-wind scenarios.

Table 4-2 shows the PIRT high-rank phenomena for the FoMs of local wind characteristics and structural loads. The phenomena of vortex shedding and hydrodynamics force on stationary structures have been discussed in the PIRT for external flooding. For a local high-wind simulation, the wind generation is usually accomplished by a separate model, like climate models, and represented by the inflow boundary conditions in CFD models. However, the inflow boundary is known to be a challenging problem for SPH methods since it is hard to maintain the continuity during the particle injection. However, some feasible models are available in the NEUTRINO

package, where a flow emitter can emit SPH particles according to the volumetric or mass flow rate. Though, such a rough model is likely to cause serious mass loss problem, it is a convenient approach appropriate for violent scenarios. As a result, adequacy rankings for the inflow BC with respect to the math model and code are "**M**edium" since these issues are considered moderate. At the same time, the ranking for data adequacy is "High" since a lot of large-scale data have been generated from the climate and atmospheric engineering. For the turbulence phenomenon, SPH also experiences some accuracy and efficiency issues. As discussed previously, particle refinements can improve the accuracy, however, the computational costs make SPH methods far less competitive against mesh-based methods. The well-established turbulence models are not perfectly suitable for SPH methods with different fundamentals. As a result, the adequacy ranking for both math model and code of SPH is "**M**edium" since these issues are considered moderate. Since an adequate amount of data have been produced by aerospace and mechanical engineering, like large-scale wind tunnel experiments, the data adequacy is ranked as "**H**igh" for the turbulence phenomenon.

Table 4-2: IRT and lists of numerical benchmarks designed for each phenomenon in high wind scenario.

ID	Phenomenon	Imp	Adequacy		
			Math Model	Code	Data
A	Wind Characteristics				
	Inflow boundary	H	M	M	H
	Turbulence	H	M	M	H
	Vortex shedding	H	L	L	H
B	Structural Loads				
	Hydrodynamics force on stationary structures	H	M	M	H

4.1.2. Performance Measurement Standards

To decide if the code is adequate for the specific application, standards need to be established, against which code models and performance can be measured. In this study, the standards are described as – the performance of the code is considered adequate for each separate phenomenon or process if:

- The code state of the art is known, documented, and acceptable
- The code is generally acceptable for simulating each separate phenomenon of the application
- The code predicts the high-fidelity data from separate effect tests with acceptable accuracy
- The code is acceptable for the scale of the specific target applications

The performance of the code is considered adequate for integral scenarios if:

- The field equations represent the key processes and phenomena
- The code is generally acceptable for simulating the application scenarios with various phenomena and processes
- The code predicts the behavior of important phenomena as observed in appropriate integral effects tests with acceptable accuracy.
- The code represents the interactions of between phenomena, process, and system components

To provide a quantitative measure for the term "acceptability," criteria are needed for assessing the code accuracy and quality of validation database. In this study, the code accuracy is measured by simulation errors in predicting selected QoIs, while the quality of validation database is measured according to the attributes of scaling analysis. Table 4-3 shows the accuracy standards measured by error (L_1 relative error norm) in predicting selected QoI.

Table 4-3: Simulation code accuracy standards measured by error in predicting selected QoIs.

Accuracy Statement	**L_1 Relative Error Norm in predicting selected QoIs**
Excellent Accuracy	$\leq 10\%$
Reasonable Accuracy	$\leq 20\%$
Insufficient Accuracy	$> 20\%$

The scaling analysis can be defined as "a technique for characterizing degree of similarity between processes and phenomena to occur in a real prototypic system and those to occur in its reduced and experimental analog" [139]. A systematic scaling analysis is composed by "top-down" and "bottom-up" analysis. As discussed previously, the scaling analysis is tightly coupled with the model assessment process. Starting from the PIRT, the identifications of physical characteristics require complicate procedures such that the interconnections among system components, phenomena, and processes are well understood. In this study, the dimensionless groups are used to represent the physical characteristics that governs the system responses. Next, the relative importance of the dimensionless groups is investigated through various code assessment process until a relationship is developed between the selected phenomena and the dimensionless groups. A combination of various relationships is usually known as a "workable form". Next, the workable form will be validated by comparing against data until a final form is determined. Until this point, it can be claimed that all data involved in this process have been "compressed" into an equation-based workable form. As a result, the scaling process is to find the "encodings" for the database, and all the model assessment process is to test the applicability of this form. At the same time, scaling also aims to verify that the database is sufficient. In this study, the dimensionless number of experiments are compared to the applications, and the sufficiency can be verified if the test data envelop the application behavior. In another word, the dimensionless number in applications should be covered by the range of dimensionless groups in experimental database. However, the selected phenomena can be only peculiar to the experiments and is not expected in the applications. Such issue is usually known as the scaling distortion. As a result, a relevance analysis is needed for the database to ensure that all selected phenomena are also typical under prototypic conditions. Finally, the model predictions are made for the applications, and an uncertainty analysis is

performed with respect to each evaluated phenomenon. If the database is found to be sufficient, relevant and less distorted with respect to the application, the validation database is similar to the application scenarios, and it is applicable to the full-scale analysis. Table 4-4 shows general descriptions for the attributes and criteria of scaling analysis.

Table 4-4: Attributes and criteria of systematic scaling analysis.

Attribute	**Criterion**
Sufficiency	The physical characteristics of database bound those for applications.
Relevancy	The physical characteristics of database are relevant to those for applications
Distortion	The uncertainty of model predictions can be bounded when it is propagated through scales

Considering the previous discussion on code accuracy and database similarity, general performance standards can be identified. Such standard acceptance criteria will be applied to all accuracy assessment, where the data are either from fundamental tests, separate-effect tests, or integral-effect tests, and are defined as follows:

Excellent – For the high-rank phenomena, the L_1 relative error norm between the predicted and the measured quantity of interests is less than 10% for both external flooding and high wind. The code can also be confidently used in similar applications with errors being bounded by an uncertainty band.

Reasonable – For the high-rank phenomena, the L_1 relative error norm between the predicted and the measured quantity of interests is less than 20% for external flooding and 30% for high-wind- related tests. The code can also be confidently used in similar applications with errors being bounded by an uncertainty band, which is greater than those with excellent agreements.

Insufficient – For the high-rank phenomena, the L_1 relative error norm between the predicted and the measured quantity of interests is larger than the requirements for acceptable agreements. The errors of most predictions lie outside the uncertainty bounds, and major

modifications are needed before the code can be used in similar applications with sufficient confidence.

Also, a code can still be adequate even if their agreements are insufficient given that (1) the phenomena are demonstrated to not have the dominant impacts on the scenarios indicated by their PIRT ranks, and (2) there is an appropriate method for quantifying and predicting the simulation uncertainty

Presently, the minimum standard for acceptable code in terms of performance is the "reasonable" agreements, and this study focuses on the assessment of each separate phenomenon due to the restriction of database. Besides, the standard criteria are equally assigned for all PIRT high-rank phenomena in this study. In a word, SPH methods are acceptable for a separate phenomenon if they can predict the phenomenon with reasonable or excellent accuracy. In next chapters, accuracy assessments will be performed, and finally, an adequacy decision for each separate phenomenon will be made based on an accumulation of all processes.

4.2. Develop Assessment Base and Evaluation Model

In this chapter, both elements 2 and 3 are discussed. Element 2 aims to acquire appropriate experimental data relevant to the scenario being considered and ensuring the suitability of experimental scaling. There are five steps: (5) specify objectives for assessment base; (6) perform scaling analysis and identify similarity criterion; (7) identify existing data and/or perform IETs and SETs to complete database; (8) evaluate effects of IET distortions and SET scaleup capability; (9) determine experimental uncertainties as appropriate.

For external-flooding applications, the analysis purpose is to assess if the performance of SPH in simulating the scenario of "floods damage the building structures, enter the room, and cause diesel generator (DG) malfunctioning" is acceptable. Detailed event progressions are listed as follows:

(1) If there is an external flooding after a dam breaching, water overflows the NPP site, and the AC power fails immediately.
(2) If the water reaches the DG building, and the peak forces acting on the building exceed a limit, water starts to overflow the room.

The corresponding Figure of Merits (FoMs) include the response time and the structural loads on SSCs by floods. In this study, the response time is the time for the external floods to reach the DG building and to potentially fail the DGs. The response time is crucial for the nuclear safety analysis, including estimating the decay heat, initiating emergency procedures and identifying responses of SSCs, etc. The structural loads on stationary SSCs are important for determining the structural integrity and operational status of SSCs during external-flooding scenarios. For example, if the load exceeds the regulatory limits, certain SSCs need to be shut down; if the force exceeds the operational limits, certain SSCs can be malfunctioning and even failed. Besides, the structural loads on moving objects are important for the flooding-induced debris analysis, including the debris transport and its impacts to SSCs. Figure 4-2 shows a postulated scenario of external-flooding induced by dam failures. A simulation has been performed with an SPH software package named NEUTRINO, and some key configurations for this simulation are listed in Table 4-5. Therefore, the analysis purpose is to determine the accuracy of predicted FoMs by SPH and NEUTRINO.

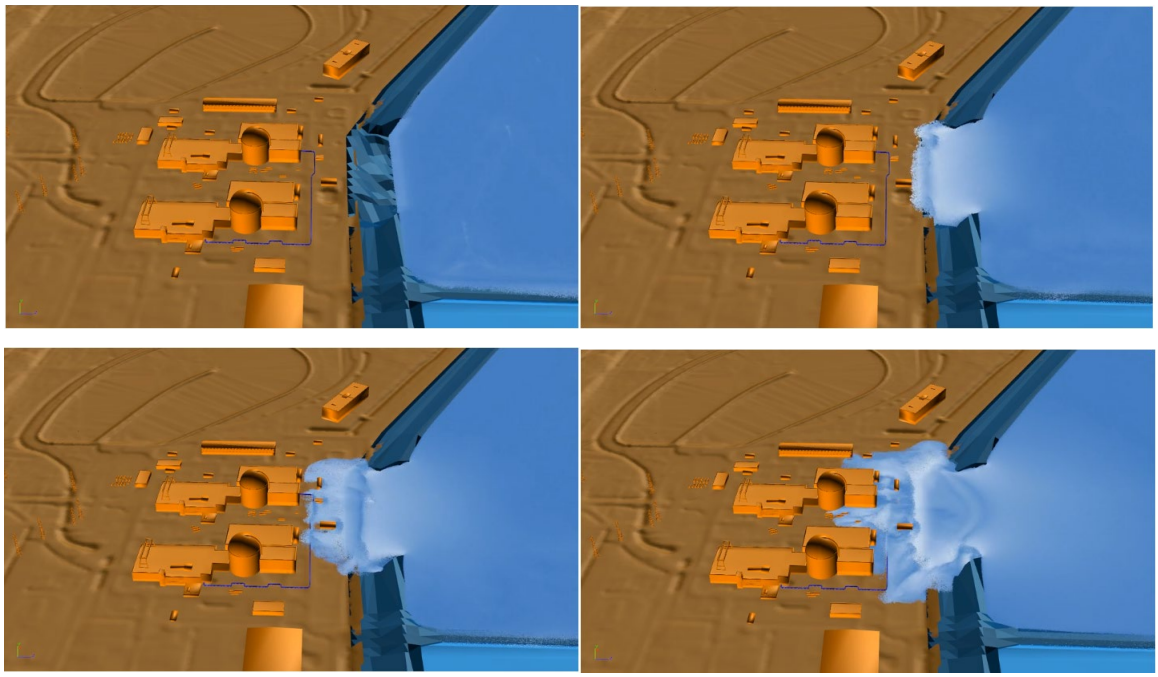


Figure 4-2: Demonstration for a postulated external flooding scenario induced by dam failures. A full-scale simulation is prepared with NEUTRINO software package [78].

Table 4-5: Scenario configurations, NEUTRINO setups, and predicted FoMs.

Scenario Configurations	
Water level at the dam breaching	5m
Breaching width	10m
Distance to buildings	20m
NEUTRINO Setups	
Particle size	2m
CFL number	0.6
Predicted FoMs	
Response time	20min
Peak force	400N

For high-wind applications, the purpose is to assess if the performance of SPH in simulating the scenario of "the long-last wind damages the offsite facilities including exhaust stack and electric wire" is acceptable. Detailed event progressions are listed as follows:

(1) If there is an EF1[1] scale tornado, the exhaust stack of NPP and electric wire can be damaged.
(2) If the electric wire is damaged, the NPP site may experience loss of offsite power (LOOP).

The corresponding FoMs include characteristics of the local wind fields experienced at the NPP site and the structural loads on SSCs by winds. In this study, the investigation on wind characteristics mainly includes the wind turbulence, vortex shedding, and wind generations, which are important for the structural analysis of SSCs. Besides, the wind field analysis is also important for the statistical analysis of wind-generated-missiles trajectories, which are important for the safety analysis of SSCs and plant operation personnel. Similar to the external-flooding scenario, the structural load analysis aims to determine the structural integrity and operational status of SSCs. Presently, no SPH simulation has been set up for a full-scale high-wind scenario, and this study mainly investigates the potential applicability and accuracy of SPH in capturing separate FoMs.

Since this study does not aim to design new facilities, the data are collected from existing literature. To ensure the data applicability to the application scenarios, a scaling analysis is needed to ensure the relevancy and sufficiency for all collected data. Table 4-6 shows the list of related numerical benchmarks for all high-rank phenomena collected from various sources, while the Quantity of Interests (QoI) for each benchmark are also listed. In this study, not all benchmark has been demonstrated, but at least one benchmark will be discussed for each phenomenon.

[1] Enhanced Fujita (EF) Scale rates the intensity of tornadoes in U.S. and Canada based on the tornado's damage. Detailed information for EF1-scale tornado can be found in the glossary table at the end of this document.

Table 4-6: Numerical benchmarks for all high-rank phenomena. Benchmarks in blue and italic fonts are investigated but not discussed in this study.

External-Flooding Scenario			
ID	High-Rank Phenomena	Numerical Benchmarks	QoI
A	Response Time		
	Solitary wave	*Wave Piston Machine* [136] [140]	Wave height
	Wave propagation	Falling of water column [97]	Position of water front edge
	Vortex shedding	Flow around cylinder [97]	Velocity field
B	Structural Loads		
	Hydrodynamic force on stationary structures	Dam Break [141] [142]	Force magnitude
	Hydrodynamic force on moving structures	Moving solids in static fluids [143]; *Flow-Induced Motion of Floating Bodies with wave maker;* [144]	Moving trajectory
High-Wind Scenario			
A	Wind Characteristics		
	Turbulence	Lid-driven cavity flow;	Velocity field
	Vortex Shedding	Wind flow over block;	Velocity field
B	Structural Loads		
	Hydrodynamic force on stationary structures	*Flow around cylinder* [97]	Force magnitude

4.3. Assess Evaluation Model Adequacy

In EMDAP framework, the assessment of simulation codes is divided into two parts. The first part is the bottom-up evaluation of model and code. The objective is to evaluate the model(s) pedigree, applicability, fidelity to appropriate fundamental or SET data, and scalability. The second part is the top-down evaluation, and the objective is to examine the integrated performance of the models. In a word, EMDAP tries to perform scaling analysis and accuracy assessment at the

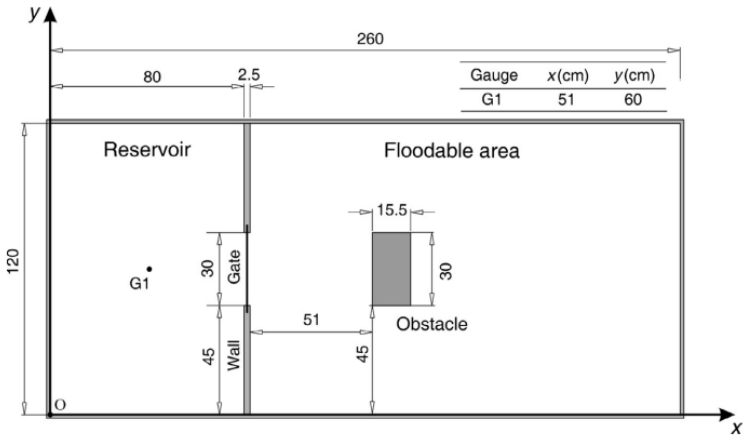
same time. In this study, since all data are collected from literatures, the scaling and accuracy assessment are presented separately by two chapters for clarity. This chapter mainly discuss the accuracy assessments of SPH methods. Two software packages, including LAMMPS-SPH and NEUTRINO, are applied. Since 3D NEUTRINO simulations can be too computationally expensive with fine particles, LAMMPS-SPH is used mainly for 2D simulations. Although NEUTRINO and LAMMPS have slight differences in solving the particle pressure, their results are similar [138].

4.3.1. External-Flooding Scenario

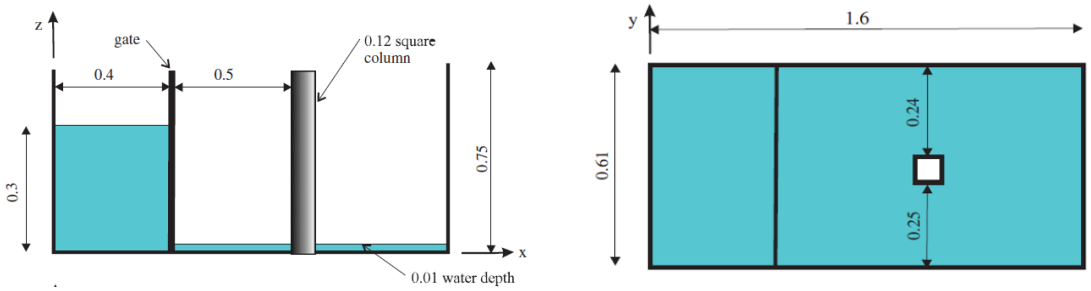
This section assesses the accuracy of SPH in predicting the force and the trajectory of moving objects for the external-flooding scenario. L_1 relative error norm is employed to measure the error between predicted results and the experimental data. If the error is less than the criterion (20%), it can be claimed that the accuracy performance of SPH methods is reasonably acceptable by for predicting a certain QoI. If the error is less than 10%, the accuracy performance is claimed to excellently acceptable. If the error is larger than 20%, the accuracy performance is insufficient.

4.3.1.1. Dam Break

In flooding scenarios, since the force exerted by the flow to the structures is an important quantity to measure, dam break is selected to test the capability of SPH methods in predicting force magnitude and flow propagation. The first simulation is designed based on the experiment by S.J. Cummins, et al. [141] The configuration of experimental setups is demonstrated in Figure 4-3.



(a) Schematic diagram of the dam geometry from Aureil, et al. [142]



(b) Schematic diagram of the dam geometry from Cummins, et al. 2012 [141]

Figure 4-3:Sketch plots of experimental setups from Aureil and Cummins work.

To reproduce the Cummins' experimental configurations in NEUTRINO, a gate is first put in position and held for 1 sec until all fluid particles are settled down. Then the gate is opened, and fluid collapses driven by gravity. Figure 4-4 shows the evolution of fluid over the surface. The dimensionless ratio for this simulation is 0.6.

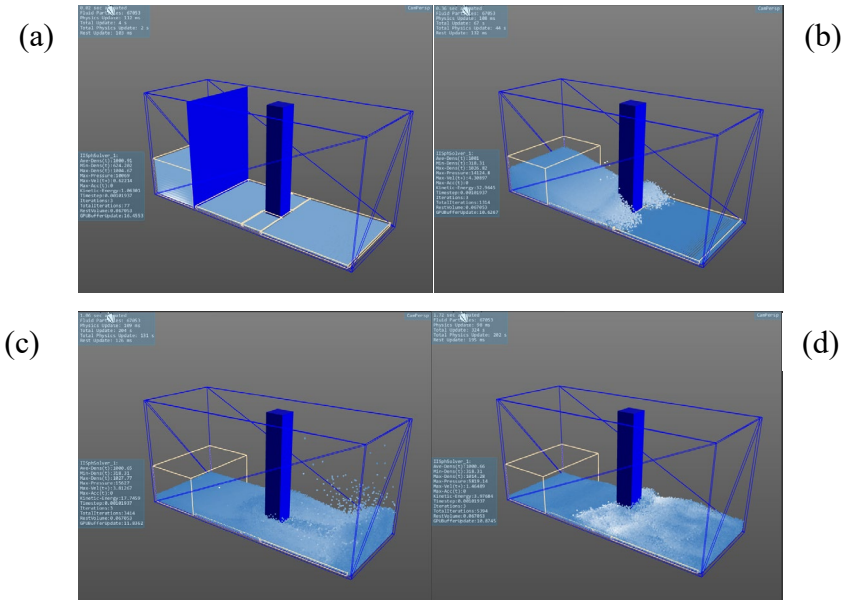


Figure 4-4: Evolution of the water collapse and interaction with the column simulated by NEUTRINO. (a) start of simulation; (b) primary wave hitting dam structure; (c) wave reflected; (d) Reflected wave hitting dam structure.

The transient comparison of simulation results again experimental data can be found in previous work [97]. The time step is affecting little on the results, as long as the CFL number is lower or equal to 0.5. Figure 4-5 shows the convergence plot of particle size in predicting different QoIs, including the maximum force of primary and reflected wave exerting on the structure, the time of those maximum forces. The percentage represents the relative error of simulation results against experimental data. When particle size reaches 0.01m, further refinement does not continue to improve the accuracy. And comparing to the force magnitude, contacting time of wave propagation is affected more by the particle size.

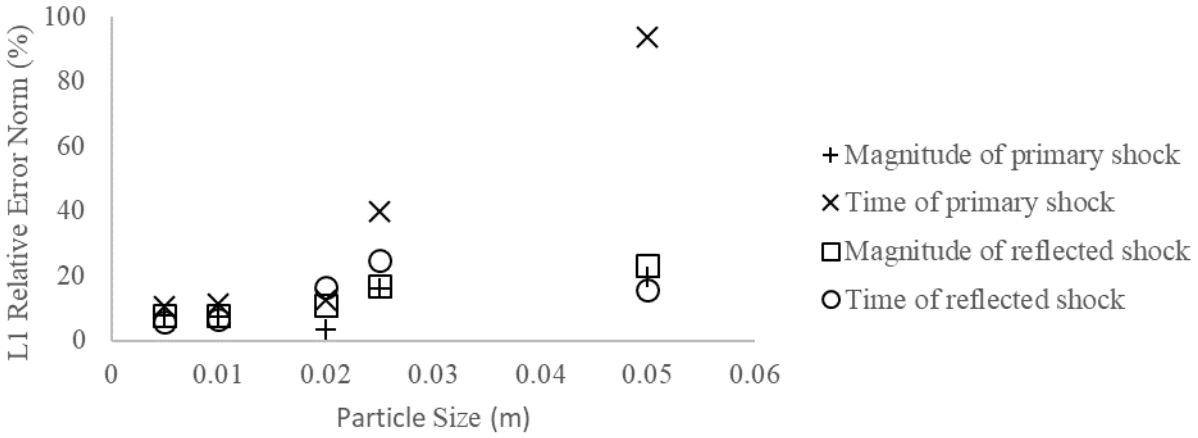


Figure 4-5: Particle size convergence plot for force magnitude and contacting time of primary and reflected wave.

Figure 4-6 shows the plot of relative error against the computational time. When the particles are refined, better results are obtained with more computational time needed. Note that when particle size is 0.02m, the simulation results are better (70% improvement in magnitude of primary shock; 60% in time of primary shock; 34% improvement in magnitude of reflected shock; 33% improvement in time of reflected shock) but consuming less time (5%) than the one with 0.025m particle size. Because the CFL condition is applied for every time step, the simulation with coarse particle size is not as stable as the one with finer size. Therefore, it requires smaller time step and longer time to simulate.

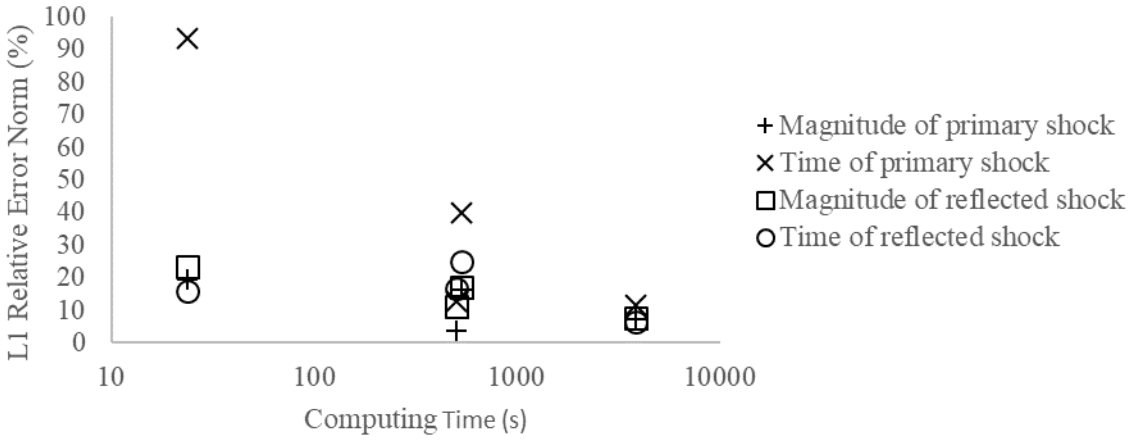


Figure 4-6: Plot of relative error against computing time for force magnitude and contacting time of primary and reflective wave.

Table 4-7 shows the list of simulation parameters that have the error of simulated QoIs fall into the bounded region. The yellow shaded row represents the parameter combination with lowest computational time, which is suggested as the optimal simulation parameters for numerical benchmarks. The L_1 Relative Error Norm is defined by

$$L_1 = |\frac{QoI_{preds} - QoI_{meas}}{QoI_{meas}}| \quad \text{Eq. 54}$$

where QoI_{preds} and QoI_{meas} represent the predicted QoI by simulations and measured QoI by experiments.

Table 4-7: Simulation parameters with 20% bounded absolute relative error for the case of dam break.

Particle Size (m)	Simulation Time/Physical Time	L_1 Relative Error Norm (%)	
		Max. Force Magnitude	Time of Peak force
0.05	11.96	19.58%	93.55%
0.025	265.8	16.16%	40%
0.02	252.4	3.58%	12.73%
0.01	1929	7.66%	10.53%
0.005	57364.5	7.26%	11.43%

In addition, a second simulation is performed according to experiments by F. Aureli [142], where forces with respect to various water height are measured, and the optimum particle size identified by the previous study (0.02m) is used for all simulations. Figure 4-7 shows the evolution of fluids

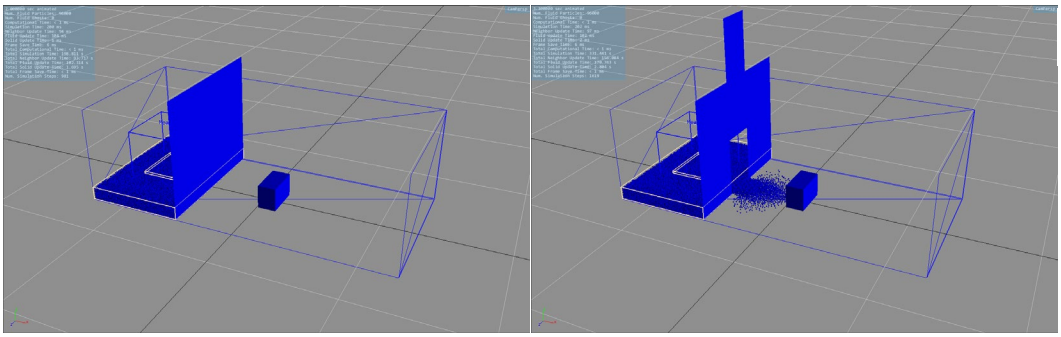
(a)  (b)

Figure 4-7: Evolution of the water collapse and interaction with the column based on Aureli's setups. (a) start of simulation; (b) primary wave hitting dam structure.

The Aureli's data can also be used for scaling analysis. In this study, it is found that the QoIs, including peak force magnitude and the time of peak force, are directly related to the initial depth of water h. To investigate the sufficiency and distortion of database for dam break, the initial depth h is non-dimensionalized as x^* by Eq. 55, where L is the distance between the gate and the stationary object. Figure 4-8 shows the comparison of the measured against SPH predicted peak

force with different dimensionless ratio x^* (Eq. 55). At the same times, linear functions are fitted to both datasets, based on which an error plot is made with respect to the dimensionless ratio x^*.

$$x^* = h/L \qquad \text{Eq. 55}$$

Figure 4-9 shows the plot of error for predicting peak force with respect to the dimensionless ratio x^* given the same particle size as 0.02m. When the x^* is larger than 0.2, the error can be properly bounded and reasonably accepted.

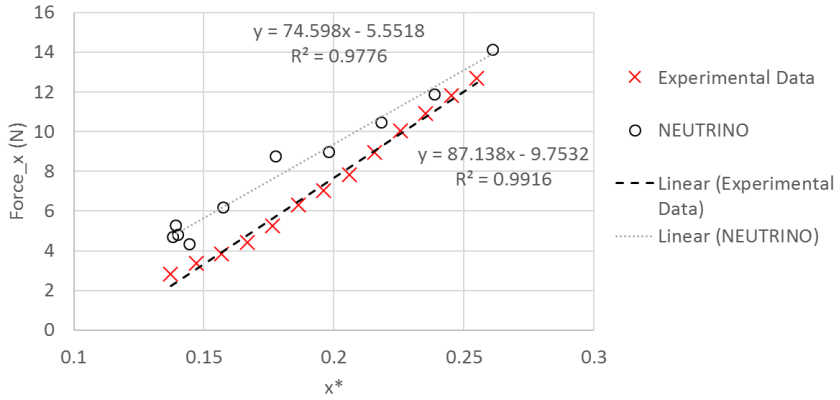


Figure 4-8: Comparison of the measured against SPH predicted peak force with different dimensionless ratio x^*.

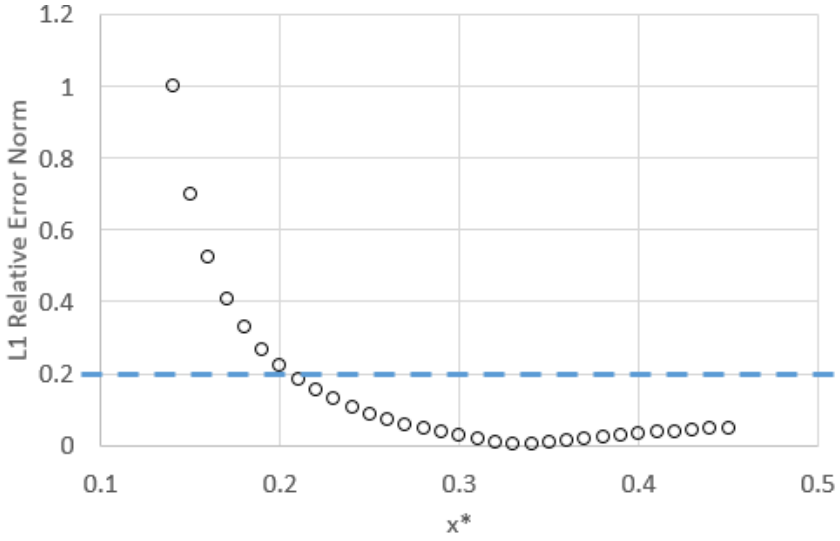


Figure 4-9: Plot of L_1 relative error norm for predicting the peak force versus the dimensionless ratio x^* with particle size as 0.02m.

Considering the effect of particle size and dimensionless ratio, the error in predicting the peak force by SPH can be propagated to the application scenario. Assuming that the dimensionless ratio is the major physical characteristic that measures the similarity, suggestions of simulation parameters can be made according to previous studies on particle size and dimensionless ratio. Table 4-13 shows the suggested simulation parameters for the application scenario based on the SET analysis. Since this study does not have data that has the same dimensionless ratio but different facility sizes, the particle size is suggested to be the same as or smaller than 2cm, such that the L_1 relative error norm for the predicted peak force is less than 20%.

Table 4-8: Simulation parameters that are scaled from numerical benchmarks to a postulated scenario with the same dimensionless group.

	Numerical Benchmarks	Scenario Simulation
Water height	0.1 ~ 0.13m	5m
Distance to the static structure	0.51m	20m
Dimensionless ratio x^*	0.2 ~ 0.26	0.25
Suggested particle size	0.02m	0.02m
Predicted peak force error (L_1 relative norm)	5%~20%	3.6%
Predicted time of peak force error (L_1 relative norm)	4%~18%	12.73%

4.3.1.2. Moving Solids in Static Fluid

In order to evaluate the accuracy of moving solids in flow, simulations with falling and floating block in the fluid are set based on the experiments by T.R. Wu, et al [143]. For the case of falling cube, the block, whose density is $2120\ kg/m^3$, is held initially at top of the tank and release after the fluid is stable. Figure 4-10 shows the dimensions of the water tank and cube for both the falling and floating experiments. Figure 4-12 shows the transient plot of block's vertical displacements from NEUTRINO simulation.

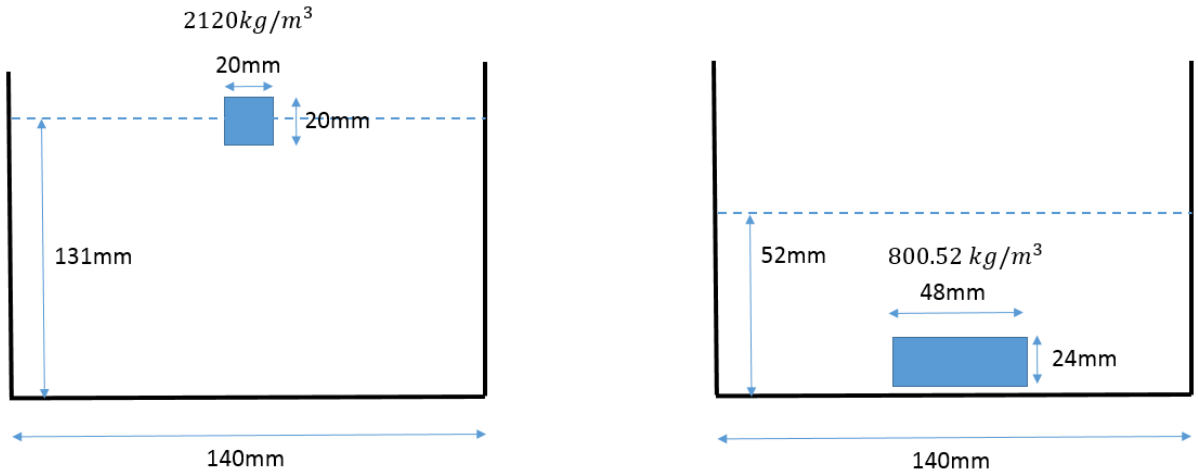


Figure 4-10: Dimensions of the water tank and cube in the falling experiment (left) and the floating experiment (right). The density of water is 1000 kg/m^3.

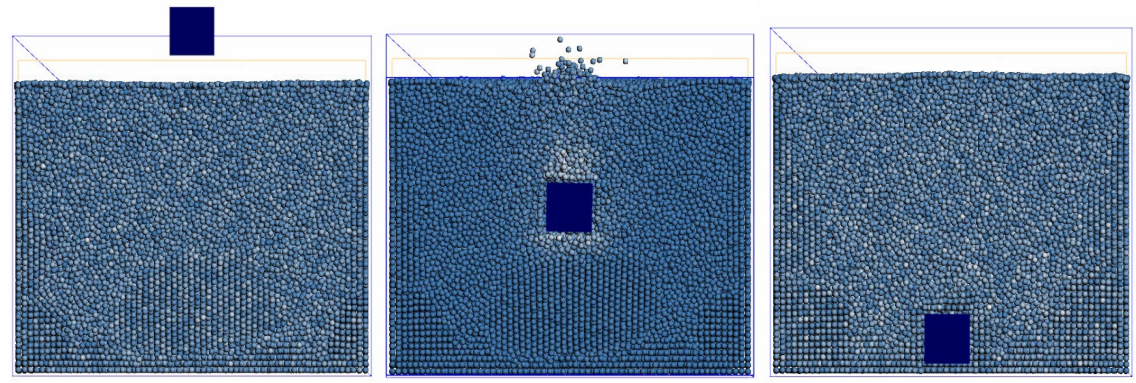


Figure 4-11: Transient plots of falling solids in fluid from NEUTRINO.

Figure 4-12 shows the plot of the cube vertical displacement against the time from both NEUTRINO simulation and experiment. Note that both the displacement and time have been non-dimensionalized, where t* and z* are defined in Eq. 56 and Eq. 57.

$$z^* = z/H \qquad \text{Eq. 56}$$

$$t^* = t \Big/ \sqrt{\frac{2H}{g}} \qquad \text{Eq. 57}$$

Three different particle sizes are tried with the simulation and a good agreement is found when the particles are 2.5mm. In this case, the falling time of the cube, defined as the non-dimensional time spent for the heavy cube to reach the bottom of the tank from its initial position, is selected as the QoI.

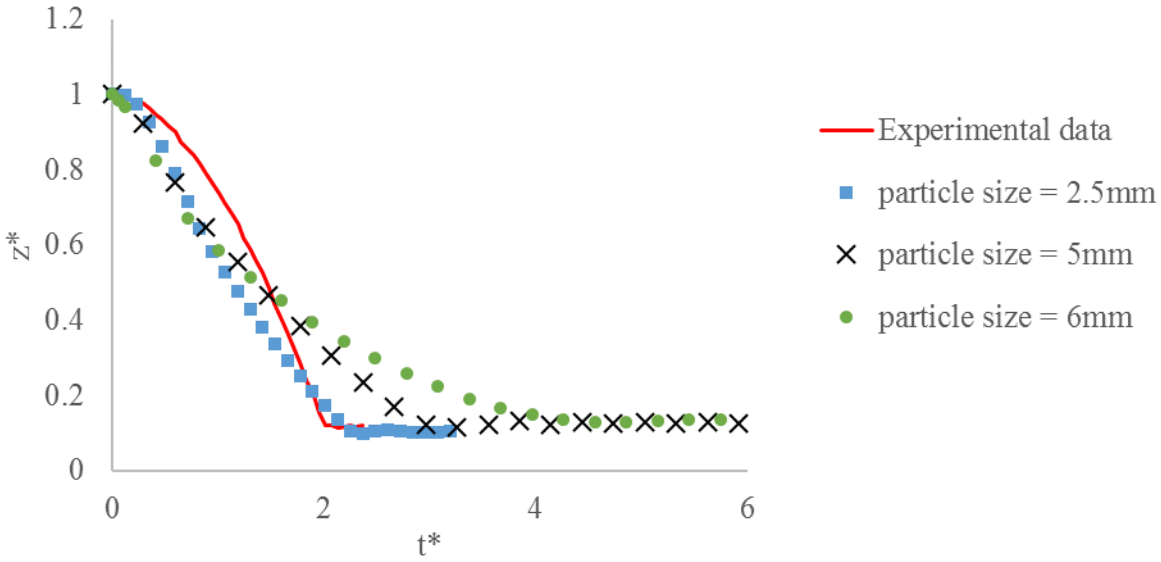


Figure 4-12: Comparison of vertical displacement of falling block simulated NEUTRINO with three particle sizes.

Table 4-9 shows the L_1 relative error norm by comparing the simulated cube falling non-dimensional time against the measured data. Error is decreasing as the particle being refined. And the heavy cube will fall slower as the particles get larger. If the particle size is larger than 6mm, the heavy cube will float on top of the fluid. Because the mass of cube is constant, and the insufficient number of particles will cause high simulation error, especially in pressure calculation, large particle size will result in unphysically high buoyance force.

Table 4-9: Particle sizes in NEUTRINO simulations for falling cube, and the absolute relative error of each simulation in predicting the cube falling time.

Particle Size	L_1 Relative Error Norm (%)
6mm	>100%
5mm	52.73%
2.5mm	5.52%

Figure 4-13 shows the transient plots of floating blocks from simulation, where the wooden block (800.52 kg/m^3) is held initially at bottom of the tank and released after the fluid is stabilized. Note in Figure 4-13, a gap is found beneath the cube, because the cube will stick to the bottom if there is no fluid beneath it. In this study, SPH typically needs 2-3 layers of fluid particles. The floating time, defined as the non-dimensional time spent for the light cube to reach highest position from its initial position, is selected as the QoI. Similar to the case of falling cube, the first QoI is aimed to characterize the capability of predicting buoyance force with SPH and Eq. 21.

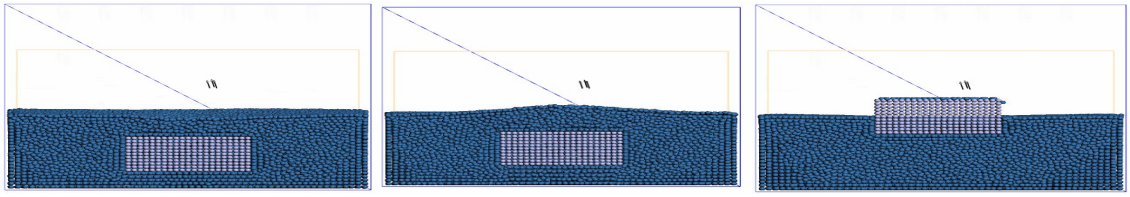


Figure 4-13: Transient plots of floating solids in fluid from NEUTRINO.

Figure 4-14 shows the time transient of block's non-dimensional vertical displacements from NEUTRINO with three different particle sizes. An insufficient agreement is found between the simulation and measurements. For comparison purposes, another simulation with the same configurations, including geometry, material, initial & boundary conditions, is performed with LAMMPS-SPH [145] software. An excellent agreement is found between the LAMMPS-SPH prediction and measurements. Furthermore, it is found that LAMMPS-SPH mainly deviates from NEUTRINO in dealing with pressure, where LAMMPS-SPH relies on equation of state and NEUTRINO uses Poisson iteration. Table 4-10 shows the L_1 relative error norm by comparing the simulation result against the measured data for selected QoI. Error is decreasing as the particle being refined, and a very good agreement is found with very refined particles SPH simulation. As

a result, when the particle size is large, NEUTRINO software tends to overpredict the hydrodynamic forces acting on the moving objects.

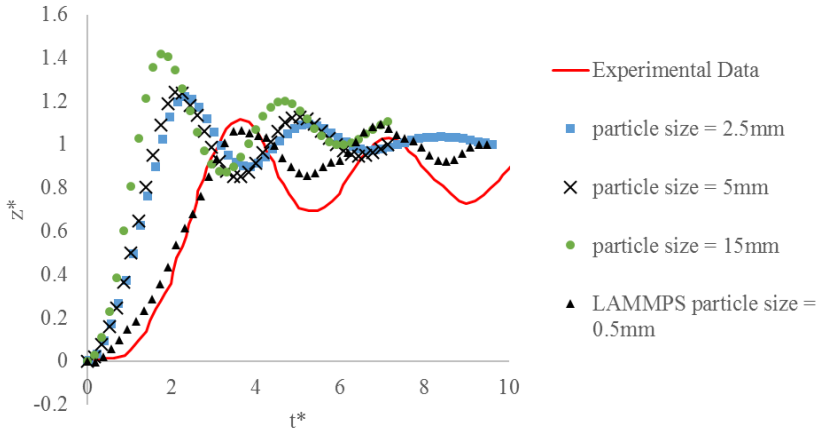


Figure 4-14: Comparison of vertical displacement of floating block simulated by NEUTRINO.

Table 4-10: Particle sizes in NEUTRINO simulations for floating cube, and the L_1 relative error norm of each simulation in predicting the cube floating time and oscillation period.

Particle Size	L_1 Relative Error Norm (%)	
	Floating time	Oscillation period
6mm	54.64%	16.42%
5mm	45.57%	14.30%
2.5mm	39.67%	13.28%
0.5mm (LAMMPS)	4.41%	6.71%

Both simulations suggest that particle size is crucial to the accuracy of SPH in predicting the falling and floating time, which is affected by the force balance between gravity and the buoyance force. And the accurate prediction of buoyance force depends on the particle intensity around the cube. Therefore, for the moving object calculation, the cube density ratio (ρ^* defined in Eq. 58) and the ratio between cube volume and average particle volume (V^* defined in Eq. 59), are selected as the dominant physics. $\bar{V}_{dp}$ is the average particle volume defined by Eq. 60 [133], and d is the initial particle diameter.

$$\rho^* = \rho_{cube}/\rho_{fluid} \quad \text{Eq. 58}$$

$$V^* = V_{cube}/\bar{V}_{dp} \quad \text{Eq. 59}$$

$$\bar{V}_{dp} = d^3 \quad \text{Eq. 60}$$

Suggestion on particle size will be made with these parameters being kept constant. Table 4-11 shows the suggestion on particle size selection for corresponding scenarios. For the case of falling cube, the simulation with 2.5mm particles is selected for scaling, while the simulation with 0.5mm particles is selected for the floating cube. However, since there is only one dataset for this phenomenon, the dimensionless group cannot be verified, and the dimensionless parameters for application scenarios are assumed to be the same as the numerical benchmarks.

Table 4-11: Simulation parameters that are scaled from numerical benchmarks to a postulated scenario with the same dimensionless group.

	Numerical Benchmarks	Application Scenario
Falling object		
Cube volume	8000mm^3	8m^3
Density	2120kg/m^3	2120kg/ m^3
Density ratio ρ^*	2.1	2.1
Volume ratio V^*	512	512
Suggested particle size	2.5mm	0.25m
Predicted error L_1relative norm)	5.52%	~5.52%
Floating object		
Cube volume	56448mm^3	451.584m^3
Density	800.52kg/m^3	800.52kg/ m^3
Density ratio ρ^*	0.8	0.8
Volume ratio V	451584	451584
Suggested particle size	0.5mm + Equation of State	0.1m
Predicted error (L_1 relative norm)	4.41%	~4.41%

In above sections, simulations of dam break and moving solids have been executed and the results are compared against experimental measurement. For both cases, SPH has shown good performances in predicting the selected QoIs, including the force magnitude, water propagation speed to dry surface, falling and floating time, and the errors are properly bounded. By requiring the predicted QoIs in scenario simulations to have less than 20% absolute relative error, simple scaling has been performed for both cases.

4.3.2. High-Wind Scenario

This section assesses the accuracy of SPH in predicting the velocity field for the high-wind scenario. $L1$ relative error norm is employed to measure the error between predicted results and the experimental data. If the error is less than the criterion (30%), it can be claimed that "the performance of SPH is acceptable by RISMC applications for predicting velocity fields in the high-wind scenario of the long-last wind damages the offsite facilities including exhaust stack and electric wire".

4.3.2.1. Flow over Obstacle

In high wind simulation, flow over a static obstacle is a very important consideration. In order to simulate a high Reynolds number, $k-\epsilon$ model is applied in mesh-based simulation methods and the velocity field is compared to SPH. The particle size is selected to be 0.01m. Experimental data from Danish Maritime Institute are also demonstrated for validation [146]. In this case, STARCCM is used as the mesh-based simulation tool. Figure 4-15 shows the dimension of the channel.

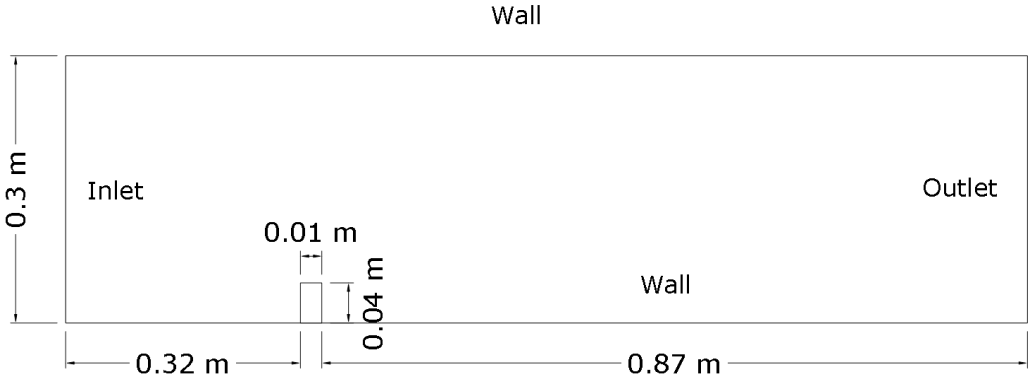


Figure 4-15: Setup and dimensions of channel flow with block.

The velocity profile along a vertical straight line is extracted and plotted before and behind the block. The profiles at x=0.38m (behind block) are plotted in Figure 4-16. It is found that SPH accurately represents the average velocity with correct inflow BC. The vortices formed behind the block are captured without the turbulence model. However, velocity near the top wall (around y=0.3m) is not correctly captured, because the boundary layer is hard to capture for turbulent flow

with SPH methods. Besides, the CFL number (Eq. 61) needs to be small or a particle vacancy will form behind the block (Figure 4-17). The maximum CFL number is around 0.1 ($\Delta t \approx 10^{-3}$) without creating a vacancy. However, the vortex is only accurately represented when CFL is lower than 0.08 (($\Delta t \approx 10^{-4}$) and further refinements will have little effect on the velocity profiles. Another technique is to apply Particle Shifting Algorithm, which could perfectly avoid the particle vacancy even with CFL number as high as 0.4. However, as shown in Figure 4-16, such technique over-predicts the vortex behind the block and maximum negative velocity is 100% more than experimental data with the particle size equal to 0.01m.

$$CFL = \frac{u\Delta t}{\Delta x} \qquad \text{Eq. 61}$$

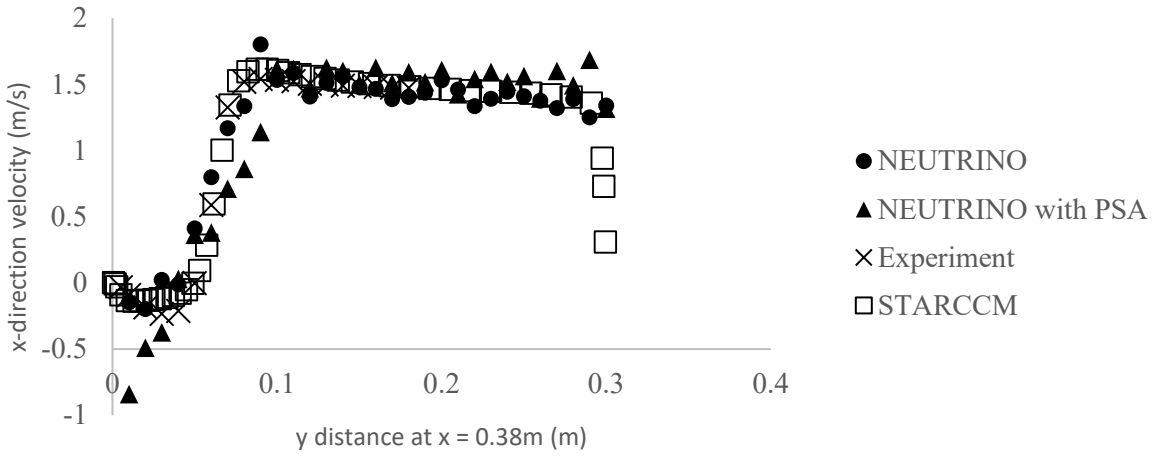


Figure 4-16: Comparison of velocity profile with and without particle shifting algorithm along vertical probe behind the block at x=0.38m.

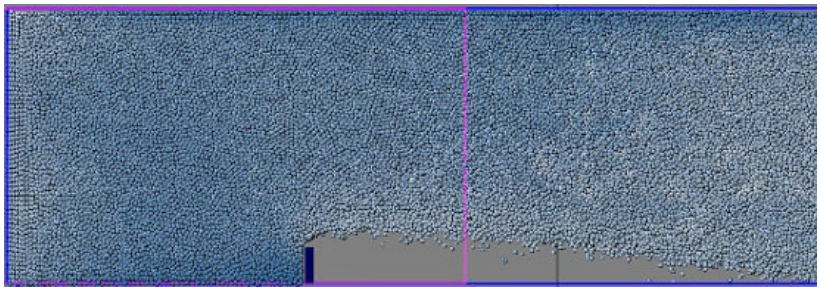


Figure 4-17: Velocity profile along vertical probe behind the block at x=0.38m.

In order to characterize the effect of particle shifting algorithm, time step and particle size, simulations with different CFL numbers, and particle size are run with and without the particle shifting algorithm. Figure 4-18 shows the plot of absolute relative error, which compares maximum negative velocity picked from SPH against experimental data, with computational times for 1.5s simulation. Note that when the particle size is changed from 0.01m to 0.025m, the particle shifting algorithm has shorter computational time (60s) with the same accuracy in predicting the maximum negative velocity (30% relative error). Besides, because particles need to move less at a time when the time step is shorter, the particle shifting algorithm should move particles weakly. Parameters of diffusion constant and maximum displacement length coefficient require careful tuning as the time step is modified [147]. Adaptive time step size needs to be carefully used when the particle shifting algorithm is applied.

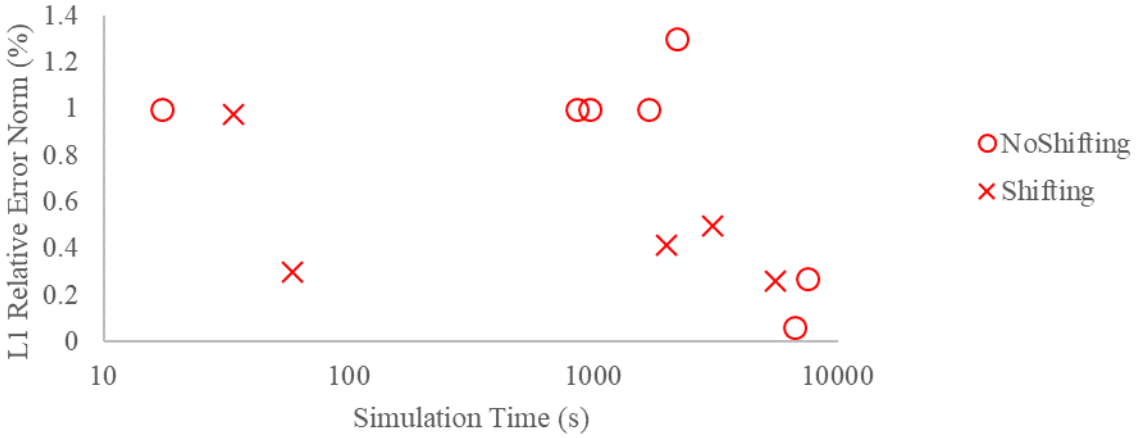


Figure 4-18: Plot of relative error for the minimum velocity for SPH with and without the particle shifting algorithm at x = 0.38m.

Table 4-12 shows the list of simulation parameters that have the error of simulated QoIs fall into bounded region. The yellow shaded row represents the parameter combination with lowest computational time, which is suggested as the optimal simulation parameters for numerical benchmarks.

Table 4-12: Simulation parameters with 30% bounded L_1 relative norm for the case of wind over block. Yellow shaded column represents the combination of parameters and models with the lowest computational cost.

particle shifting algorithm?	Particle Size	CFL	Simulation Time/Physical time	Absolute Relative Error
Yes	0.02	0.4	39.36	0.29
Yes	0.01	0.1	3695.49	0.26
No	0.01	0.07	4965.73	0.27
No	0.02	0.07	4447.07	0.06

For the third and fourth experiments, the relative error does not converge by refining the particle size and the simulation with 1cm particle tends to be more unstable than the one with 2cm. Moreover, the issue of mass-loss induced vacancy is significant. Without additional models, fixing this issue requires 10 times more computational costs, whose error remains fluctuating. Although, the particle shifting algorithm is able to fill the vacancy with much lower expenses, the algorithm violates the momentum conservation and the error fluctuation is significant. As a result, the applications of NEUTRINO for predicting the velocity of flow-over-obstacle phenomenon is not appropriate and the particle shifting algorithm is not capable of providing sufficient accuracy that can support RISMC results. If it is assumed that Reynolds number (Eq. 62) can fully characterize the flow feature and the QoI (velocity field) depends on this number only, suggestion on parameter and model selection can be made.

$$Re_f = \frac{v_{ave} \cdot h_{obstacle}}{v_{air}} \qquad \text{Eq. 62}$$

where v_{ave} is the average velocity across vertical direction of the channel, $h_{obstacle}$ is the height of obstacle, v_{air} is the kinematic viscosity of air ($1.48 \times 10^{-5} m/s^2$). However, due to the error fluctuation, the suggestion can only be made for scenarios with the same Reynolds number rather than the application scenario with $Re_f \sim 10^7$. Table 4-13 shows the suggestion of simulation parameters from numerical benchmarks to a scenario with the same Reynolds number. However,

comparing to the high-wind scenario with $Re_f \approx 10^6$, a technique is needed to extrapolate the predicted SPH error from reduced-validation to full-scale scenarios.

Table 4-13: Simulation parameters that are scaled from numerical benchmarks to a postulated scenario with the same dimensionless group.

	Numerical Benchmarks	Scenario with the same characteristics
Building height	4cm	40cm
Kinematic Viscosity	$1.48 \times 10^{-5} m/s^2$	$1.48 \times 10^{-3} m/s^2$
Wind speed	1.18m/s	11.8m/s
Reynolds number (Re_f)	3189	3189
Suggested particle size	2cm	2cm
Suggested numerical model	particle shifting algorithm	particle shifting algorithm
Suggested CFL number	0.4	0.4

4.3.2.2. Cavity Flow

In the high-wind scenario, it is important to investigate phenomenon including turbulence and vortex. In this study, a square box with $1m$ length and $1m/s$ driven velocity is set. The viscosity is adjusted such that cavity flow with various Reynolds number (Re_τ), defined in Eq. 63, can be investigated., where v_{lid} is the constant velocity of the driven lid and μ_{fluid} is the dynamic viscosity of fluid. To determine the error of SPH method, a mesh-based simulation is also performed with OpenFOAM [148].

$$Re_\tau = \rho_{fluid} v_{lid} l / \mu_{fluid} \quad \text{Eq. 63}$$

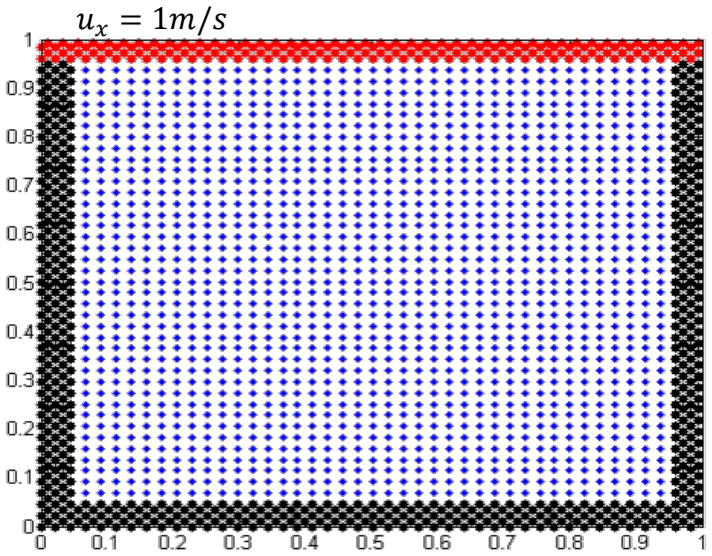


Figure 4-19: The particle distribution of SPH simulation for the lid-driven cavity flow. The red particles represent lid that is driven by a constant speed, black particles represent the stationary boundaries, while the blue particle represent the fluid.

The simulation information for OpenFOAM solver setups is listed in Table 4-14. To ensure the accuracy of high-fidelity simulations, validation is performed for the OpenFOAM simulation. Figure 4-20 shows the comparison of OpenFOAM simulation results against experimental measurements with $Re_\tau = 10000$, where the experimental data are obtained from Bouffanais's work [149]. It is found that the OpenFOAM agrees well with the experiment, and it can be treated as the source of high-fidelity data for lid-driven cavity flow under the conditions that Reynolds number Re_τ is less or equal to 10000 and the same settings (Table 4-14) are used.

Table 4-14: OpenFOAM solver information.

Solver Name	PISO
ΔT	$0.05s$
T_{end}	$100s$
Discretization Scheme	Finite Volume
Turbulence Model	LES
Mesh	120×120

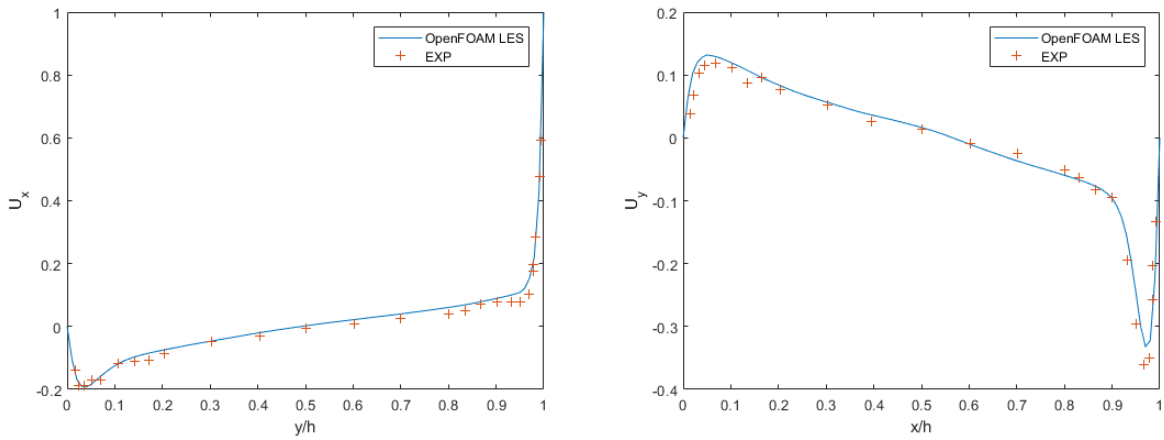


Figure 4-20: Comparison of velocity profiles from OpenFOAM and experimental measurements along the vertical centerline (left) and horizontal centerline (right) when $Re_\tau = 10000$.

In this section, simulations with Reynolds number ranging from 100 to 10000 are performed to demonstrate the capability of SPH in capturing the vortex. Since the accuracy of OpenFOAM simulation has been validated for the highest Reynolds number, the same settings are used for generating high-fidelity data in other simulations (Table 4-14). Moreover, NEUTRINO is found to have severe mass loss and accuracy issues. As a result, the LAMMPS-SPH is used as the SPH simulation tool for this benchmark. The effects of SPH particle sizes on simulation errors are investigated by comparing the L2 relative error norm. Figure 4-21 shows the centerline velocity of lid-driven simulations with Re_τ = 100 and 1000. Figure 4-22 plots the L2 relative error norm versus the Reynolds number Re_τ for SPH simulations. It is found that relationship between SPH error and Re_τ is highly nonlinear. For the same particle size, high Reynolds number will introduce more simulation errors until a critical Re_τ. After this value, the error tends to decrease due to the numerical dissipation. In this sense, numerical dissipations act as sub-scale filters that will dissipate in the same manner as LES low-pass filters. This approach, known as implicit LES [150], takes full advantage of grid resolution, and eliminates the computational cost of calculating a sub-filter scale model term. However, it is usually difficult to determine the shape of the LES filter that is associated with numerical schemes, and the truncation error will affect the numerical dissipation. For SPH simulation, the effect of implicit LES offers possibility of fast and high-accuracy simulations in large scale scenarios. However, When the particle size is 0.02m, the SPH simulation

satisfies the accuracy requirement only when $Re_\tau = 1000$. Though refining the particle to 0.0025m can improve the accuracy for satisfying requirements, the computational expense becomes 100 times more. Moreover, when the Reynolds number reaches 10000, particle vacancy is found (Figure 4-23), which indicates the violation of weakly-compressible assumption and the problem of mass loss.

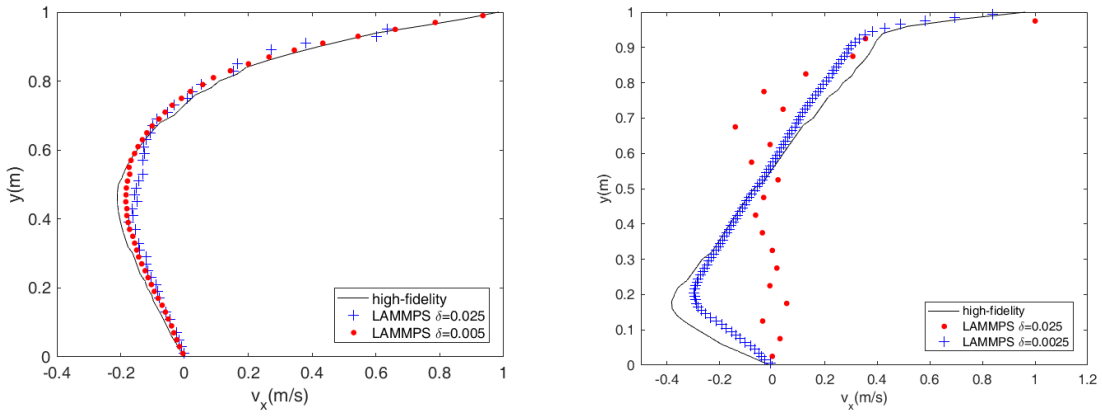


Figure 4-21: Velocity profile of different particle size on the vertical centerline with Re_τ=100 (left) and Re_τ =1000 (right).

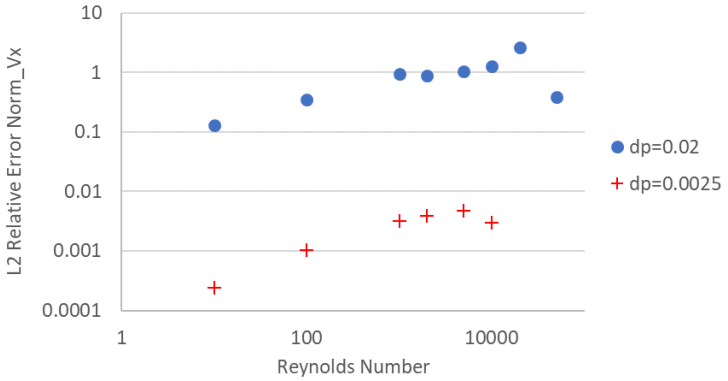


Figure 4-22: Plot of L_1 relative error norm for x-direction velocity field versus the Reynolds number.

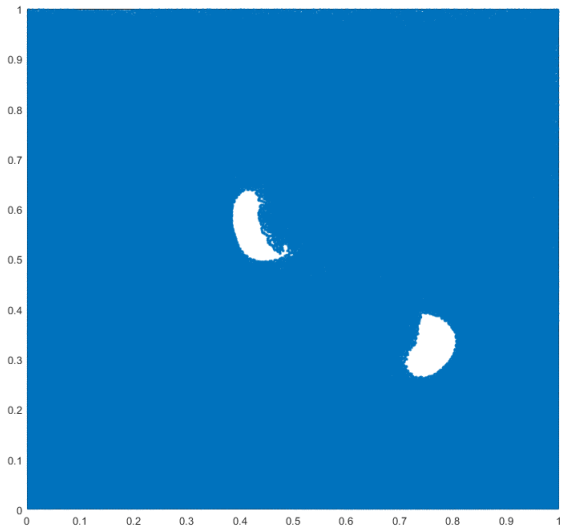


Figure 4-23: Particle distribution of SPH simulation with Re=10000.

Therefore, the accuracy of SPH in simulating the velocity field of lid-driven cavity satisfies the requirement only when particle size is less or equal to 1cm. Also, as Reynolds number is increased, issues of mass loss become severe. Figure 4-11shows the suggestion of simulation parameters from numerical benchmarks to a scenario with the same Reynolds number. However, comparing to the high-wind scenario with 10^7 Reynolds number, a technique is needed to extrapolate the predicted SPH error from reduced-validation to full-scale scenarios.

Table 4-15: Simulation parameters that are scaled from numerical benchmarks to a postulated scenario with the same dimensionless group.

	Numerical Benchmarks	Scenario with the same characteristics
Square Size	1m	10m
Kinematic Viscosity	$10^{-4}m/s^2$	$10^{-3}m/s^2$
Cavity Speed	1m/s	1m/s
Reynolds number (Re_τ)	10000	10000
Suggested particle size	$1cm$	$1cm$

4.3.3. Summary of Assessment Results

For the external-flooding scenario, the FoMs are the response time and structural loads on SSCs. For the high-wind scenarios, the FoMs are the characteristics of the local wind fields experienced at the NPP site and the structural loads on SSCs by winds. Next, according to the PIRT process, all high-rank phenomena are identified, and numerical benchmarks are set correspondingly. Next, the accuracy of SPH methods in simulating each scenario is assessed by comparing the SPH predicted QoIs against the measured data. Table 4-16 shows the assessment results of all numerical benchmarks for both scenarios, and their performances are rated by the performance measurement standards.

Table 4-16: Comparison of measured and SPH predicted QoIs for high-rank phenomena of external-flooding and high-wind scenarios.

Numerical Benchmarks	**QoI**	**Accuracy**
External-Flooding Scenario		
Dam break	Force magnitude; Water propagation speed	Excellent
Falling cube	Falling time	Excellent
Floating cube	Floating time; Oscillation period	Excellent
High-Wind Scenario		
Wind flow over block	Max. negative velocity behind the block	Insufficient
Lid-driven cavity flow	Velocity field	Reasonable

4.4. Scaling Analysis

To determine the relevancy and sufficiency of database for the intended applications, a scaling analysis is needed. A systematic scaling analysis is composed by “top-down” and “bottom-up” analysis. As discussed previously, the scaling analysis is tightly coupled with the model assessment process. Starting from the PIRT, the identifications of physical characteristics require complicate procedures such that the interconnections among system components, phenomena, and

processes are well understood. In this study, the dimensionless groups are used to represent the physical characteristics that govern the system responses. Next, the relative importance of the dimensionless groups is investigated through various code assessment process until a relationship is developed between the selected phenomena and the dimensionless groups. A combination of various relationships is usually known as a "workable form". Next, the workable form will be validated by comparing against data until a final form is determined. Until this point, it can be claimed that all data involved in this process have been "compressed" into an equation-based workable form. As a result, the scaling process is to find the "encodings" for the database, and all the model assessment process is to test the applicability of this form. At the same time, scaling also aims to verify that the database is sufficient. In this study, the dimensionless number of experiments are compared to the applications, and the sufficiency can be verified if the test data envelop the application behavior. In another word, the dimensionless number of a sufficient validation database should cover those number in application scenarios. Otherwise, the selected phenomena can be only peculiar to the experiments and is not expected in the applications. Such issue is usually known as the scaling distortion. As a result, a relevance analysis is needed for the database to ensure that all selected phenomena are also typical under prototypic conditions. Finally, the model predictions are made for the applications, and an uncertainty analysis is performed with respect to each evaluated phenomenon. Table 4-4 shows general descriptions for the attributes and criteria of scaling analysis.

Table 4-17 shows the summary of numerical benchmarks and their physical characteristics for both external-flooding and high-wind scenarios. Compared to the list of benchmarks in Table 4-6, some cases are not discussed due to a lack of validation data. It is known that the phenomena of turbulence and vortex shedding are governed by the Reynolds number that are differently defined. For the rest phenomena, the dimensionless groups are characterized based on author's best knowledge, and verification is still needed. For phenomena of hydrodynamic force on stationary structures, the QoI (peak force) is found to be linearly related to the initial height of water column according to Figure 4-8. In this study, the initial column height is non-dimensionalized by Eq. 55 Since the values of dimensionless group in the postulated full-scale scenario are covered by the range of experimental database, simulation errors are expected to be less than 10% with particle size equals to 0.02m. A similar analogy also applies to the phenomenon

of wave propagation, where the time of peak force is the QoI. Differently, the simulation errors are expected to be less than 20% with particle size quals to 0.02m. However, more data from facilities with different configurations are still needed to verify the scalability of dimensionless group. For the phenomenon of hydrodynamic force on moving structures, the density ratio ρ^* (Eq. 58) and volume ratio V^* (Eq. 59) are used, and they are assumed to be same in application and experiment scenarios. However, such a theory is not verified, and more data from different facilities and configurations are needed. Furthermore, this study assumes that the values of dimensionless group in the postulated full-scale scenario are the same as those from experimental database.

Table 4-17: Summary of numerical benchmarks designed for each phenomenon in external-flooding and high-wind scenario. For phenomena "wave propagation" and "hydrodynamics force on stationary structures", the range of dimensionless number in validation database covers the application value. For phenomena "hydrodynamic force on moving structures", the dimensionless number in application is assumed to be the same as that in validation database. For phenomena "turbulence" and "vortex shedding", the validation database does not cover the characteristics of application.

External-Flooding Scenario					
ID	Phenomenon Description	Numerical Benchmarks	Dimensionless Group		
			Symbolic Representation	Validation Database	Application
A	Response Time				
	Wave propagation	Dam Break	x^* (Eq. 55)	0.2~0.26	0.25
B	Structural Loads				
	Hydrodynamic force on stationary structures	Dam break;	x^* (Eq. 55)	0.2~0.26	0.25
	Hydrodynamic force on moving structures	Falling Objects	ρ^* (Eq. 58) V^* (Eq. 59)	$\rho^* = 0.8$ $V^* = 512$	$\rho^* = 0.8$ $V^* = 512$
		Floating Objects		$\rho^* = 2.1$ $V^* = 4.5\text{E}5$	$\rho^* = 2.1$ $V^* = 4.5\text{E}5$
High-Wind Scenario					
A	Velocity field				
	Turbulence	Lid-driven cavity flow	Re_τ (Eq. 63)	1000~10000	10^7
	Vortex Shedding	Wind flow over block;	Re_f (Eq. 62)	3189	10^7

Based on the criteria in Table 4-4 and the dimensionless group identified in Table 4-17, outputs of scaling analysis can be determined. Table 4-18 shows qualitative judgements for the sufficiency,

relevancy, and distortion obtained from the scaling analysis. Due to the limitation of database, most of the distortion analysis is not applicable (N/A). For the turbulence and vortex shedding, previous studies identify large error fluctuations and strong correlations to model parameters. Therefore, the error is not likely to be bounded for the applications.

4.5. Adequacy Decision

This chapter aims to assess the adequacy of SPH methods for simulating the designated external-flooding and high-wind scenarios. The decision is made by combining results from reviews, theories, assessments, PIRT, and scaling analysis in this study. Both Figure 1-16 and Figure 4-1 have illustrated the scheme for integrating various components towards a final adequacy decision. Since the database is mainly collected from literature, this study focuses on the accuracy assessment of separate phenomena, while the integral performance has not been properly measured. Although a scaling analysis has been performed, the dimensionless groups and their relationships to QoIs are not verified. Table 4-18 summarizes the results from various components. Meanwhile, suggestions for particle-size selections are included according to the results of accuracy assessments and scaling analysis.

Table 4-18: Validation result for SPH methods in simulating the external-flooding and high-wind scenario. For each phenomenon, the particle sizes are sample, and the one with acceptable accuracy and least computational expenses are suggested.

External-Flooding Scenario							
ID	Phenomenon Description	Performance	Accuracy (L1 error)	Scaling Analysis			Suggested Particle Size
				Relevancy	Sufficiency	Distortion	
A	Response Time						
	Wave propagation	Reasonable	Reasonable (12.7%)	Yes	Yes	Bounded	0.02m
B	Structural Loads						
	Hydrodynamic force on stationary structures	Excellent	Excellent (3.6%)	Yes	Yes	Bounded	0.02m
	Hydrodynamic force on moving structures	Falling – Insufficient	Excellent (5.52%)	Yes	No	N/A	Falling – 0.25m
		Floating – Insufficient	Excellent (4.41%)				Floating – 0.1m
High-Wind Scenario							
A	Velocity field						
	Turbulence	Insufficient	Reasonable (22.3%)	Yes	No	Unbounded	0.01m
	Vortex Shedding	Insufficient	Insufficient	Yes	No	Unbounded	N/A

For the scenario of "floods damage the building structures, enter the room, and cause diesel generator malfunctioning", SPH methods can predict the hydrodynamic force on both stationary and moving structures with acceptable accuracy. For the phenomena with stationary structures, it is suggested that the particle size of NEUTRINO simulation should be 0.02m. At the same time, according to the scaling analysis by dimensionless group x^*, the database is sufficient for predicting a similar phenomenon in full-scale scenarios. Such argument also applies to the phenomenon of wave propagation, and the suggested particle size is 0.02m for a reasonable performance. For the phenomena with moving structures, the suggested particle size for a postulated scenario is 0.25m for falling cubes and 0.1m for floating cube. Since there is only one dataset for this phenomenon, distortion analysis is not applicable (N/A), the database is not sufficient. A performance measurement cannot be made regarding the SPH's integral performance

for the designated flooding scenarios, and the adequacy decision should be “inadequate” with respect to the available database.

For the scenario of “the long-last wind damages the offsite facilities including exhaust stack and electric wire”, the performance of SPH methods in simulating turbulence and vortex shedding phenomena is insufficient. Moreover, the database is not sufficient, and the distortion is large. The issue of particle vacancy problems is severe when the Reynolds number grows. Though the model of particle shifting algorithm can fill the vacancy and greatly reduce the costs, the error fluctuation remains problematic and the simulation becomes less stable. The applicability of database for vortex shedding phenomenon is Not Applicable (N/A) since only one set of data is available. Considering the computational expenses of SPH for DNS-scale simulations. It is suggested that additional models, including turbulence model, wall function, and adaptive particle size, etc., should be developed. At the same time, more data from the large-scale flooding scenarios, wind tunnel, debris transport, and so on, should be gathered. Moreover, a performance measurement cannot be made regarding the SPH’s integral performance for the designated flooding scenarios, and the adequacy decision should be “inadequate” with respect to the SPH methods and database.

In general, with the collected database, the adequacy of SPH methods in simulating external-flooding and high-wind scenarios cannot be decided. For the external-flooding scenarios, the general performance of SPH methods in predicting separate phenomena is reasonable. However, the performance for high-wind scenarios is insufficient.

4.6. Issues and Challenges

First, one challenge of SPH validation is to define validation goals and requirements for all involved phenomena, processes, and components from the perspective of risk analysis. In this chapter, for clarification, the validation goals of each numerical benchmark are defined by accuracy and calculated by validation or error metrics. Furthermore, they are assumed to be the same for all involved phenomena. However, for RISMC applications, the requirement on model accuracy is subject to change given uncertain scenarios. For scenarios where facilities have enormous safety margins, the tolerance for simulation uncertainties can be much larger than

scenarios that have the fragile equipment. In addition, for scenarios that have risks with severe consequences, the requirement on model credibility will be more stringent than scenarios that have slight effects. Second, for complex scenarios with multiple phenomena, processes, and components, the requirements on model accuracy in simulating each separate effect should be inferred rigorously according to the requirements for integral behaviors. In CSAU and EMDAP framework, none of such connection has been explicitly discussed. Besides, the subjective components of model credibility in nuclear safety-related applications cannot be ignored when external hazards are analyzed, and the uncertainties are significant (>Level 2). To be specific, the subjective component is mainly introduced by two aspects: scaling analysis and physical pattern recognition. The first aspect mainly depends on the similarity (sufficiency, relevancy, and distortion) of the validation database to prototypic conditions. The second aspect depends on the understanding of all involving physics and their interactions. Both aspects require subjective assessments and their criteria for success depend on the subjective confidence level. With the same validation goal, scenarios with complex physics require higher confidence in the scaling analysis and physical pattern recognition than those with simple phenomena, such that the model can be used for full-scale applications. However, neither CSAU or EMDAP has formalized or explicitly discussed such situations.

Second, there is no framework or procedure in the previous validation that formalizes the decision-making process for validation adequacy. Although Figure 4-1 shows a schematic structure with all inputs, detailed information, like relative importance, dependency, acceptance criterion, etc., are not explicitly discussed. At the same time, due to a lack of data from integral facilities under prototypic conditions, adequacy of simulation models has to be decided according to their performances in predicting data from separate phenomena and reduce-scale facilities. Considering the complexity of scenarios and the multitude of evidence, human assessments can be heuristic and obscure. As a result, to remain convincing, the adequacy decision-making process needs to be formalized and represented transparently. At the same time, the formalization needs to be consistent and robust, such that the decision-making process can be defendable and satisfactory across a wide range of plausible situations. Besides, for scenarios that evidence has not been exhaustively collected, the formalized decision-making process should be theoretically improvable when new knowledge and data become available.

Third, SPH methods are found to be insufficient in predicting various high-rank phenomena. Although an "excellent" performance is desirable for every phenomenon, it can hardly be achieved for SPH methods based on the previous assessments. When an insufficient performance is found for high-rank phenomena, SPH methods can still be considered adequate despite the deficiency if (1) the phenomenon can subsequently be demonstrated to not have the dominant influence on the course of the application or (2) an appropriate method can be demonstrated for quantifying the calculational uncertainty resulting from the deficiency, and the uncertainty is acceptable. The database in this study is not sufficient for verifying the PIRT, efforts are devoted to the uncertainty quantification of SPH methods. Regular UQ methods require separations of total uncertainty with respect to sources and properties and assumes that each source is scale-invariant. However, all sources of uncertainties, including discretization errors, simulation errors, model form errors, and input errors, are tightly coupled, and characterizing each of them by independent error distributions is difficult. Besides, it is found from section 4.3.2.2 that the simulation error of SPH methods depends on the characteristics of scales. Moreover, the uncertainties of situations, where the regular UQ methods apply, are much smaller than those of external-hazards scenarios. Therefore, new UQ methodologies are needed for estimating total simulation uncertainties under large uncertainties. At the same time, the new methodologies are expected to propagate (interpolate and extrapolate) the uncertainties from reduce-scales to full-scale conditions.

4.7. Summary Remarks

This chapter describes the validation of SPH methods for the external-flooding and high-wind scenario according to the CSAU/EMDAP framework. It is found that for the designated external-flooding scenario, SPH methods can predict the hydrodynamic force acting on both stationary and moving structures with excellent accuracy, where the L_1 relative error norm is less than 10%. At the same time, for the phenomenon of wave propagation, SPH simulation accuracy turns to be reasonable (L_1<20%). For each phenomenon, suggestions on particle-size selections are also made according to the results of convergence study. Next, a scaling analysis is performed for identifying the sufficiency, relevancy, and distortion of validation database from the perspective of the full-scale applications. In addition, dimensionless groups are suggested for

evaluating the full-scale behaviors based on the existing database. However, due to insufficient database, the dimensionless groups and their relationships to the system quantities of response have not been verified, and the performance of integral code predictions cannot be assessed. As a result, for the designated external-flooding scenario, the adequacy decision for SPH methods should be "inadequate" with respect to the available database.

Meanwhile, it is found that for high-wind scenarios, the performance of SPH methods in simulating turbulence and vortex shedding phenomena is insufficient. Moreover, the database does not sufficiently cover the behaviors of full-scale applications, and the scaling distortion is expected to be large. As a result, for the designated high-wind scenario the adequacy decision for SPH methods should also be "inadequate" with respect to the SPH methods and available database.

Since this study does not aim to improve the SPH methods with new models or design experimental facilities for gathering more data, three issues of established validation framework are identified, including the formulation and assessment of validation goal, obscure decision-making process for the code adequacy, and insufficient code performance. By assuming that the scaling methodology is consistent and improvable (Assumption A5), next chapter will discuss potential resolutions against identified issues and to improve the established framework by incorporating new methodologies into the EMDAP validation process.

5. VALIDATION FRAMEWORK FOR RISMC METHODOLOGY

It is found through the previous discussion that the uncertainties during the SPH validation are usually large, and it is hard to represent those uncertainties precisely by parametric distributions. As a result, this study identifies three major issues for the EMDAP-based SPH validation and proposes resolutions against each issue. Three resolutions have been proposed for each identified issue, meanwhile, case studies are prepared for demonstrating their procedures and applicability. At the same time, discussions regarding limitations and issues are made from both fundamental and technical perspectives. Finally, the resolutions will be incorporated into the EMDAP framework. Meanwhile, potential impacts for the SPH validation are discussed with the improved EMDAP framework.

5.1. Sufficient Accuracy

Considering the subjective nature of validation, the purpose of accuracy assessment is replaced by credibility assessment. To ensure the consistency of credibility requirements in validation process, a new concept of sufficient accuracy is proposed to align the safety analysis with the model credibility requirement. At the same time, such concept also guides the propagation of credibility requirement from prototypic systems to various separate-effect systems in a consistent manner. Next, a case study is performed to demonstrate the application of sufficient accuracy. Finally, findings and limitations are discussed according to the case studies.

5.1.1. Technical Development

To validate selected RISMC simulation tools, a validation framework is needed for the uncertainty quantification and scaling. In the EMDAP framework, the M&S is “frozen” for the specific scenario after the validation. However, model forms and parameters of RISMC tools are subject to change since computational efficiency is another important consideration in addition to accuracy. For example, on one hand, for scenarios where facilities have enormous safety margins, the tolerance for simulation uncertainties can be much larger than scenarios that have the fragile equipment. Under such conditions, coarse-grid simulations or low-fidelity models, which consume

few computational resources, can be used with adequate confidence; On the other hand, for scenarios with higher risks (due to severe consequences and high likelihood of occurrence), the requirements on model credibility will be more stringent than scenarios that have slight effects. And only fine-grid simulations and high-fidelity models can be confidently employed. In consequence, since the validation goal needs to be determined by the application of interest (Assumption A2), this study proposes the concept of sufficient accuracy for aligning the model credibility requirements with safety requirements. Moreover, since RISMC analysis aims to evaluate the reactor's safety margin during external hazards, the validation goal can only be qualitatively defined (Assumption A3). At the same time, the uncertainty involved is so large that the subjective nature of validation is inevitable. As a result, the primitive variable is "meaning" rather than "numbers". And previous approaches, where the M&S performance standards are characterized by one single error number, become less meaningful. Indeed, reducing the "quantitative error" implies better accuracy, but the work becomes more complicated and less scrutable. Even though the simulation error appears to be acceptable or far less than the criterion for each separate phenomenon, the credibility of that M&S for the prototypic system could remain suspicious. As a result, this study proposes the concept of sufficient accuracy that qualitatively represents the model credibility requirements by objective accuracy and subjective belief.

To be specific, the "accuracy" of a simulation result is represented by the maturity level, and the maturity level is defined as the quantitative assessment of predictive accuracy for all important QoIs at the prototypic condition; while the "sufficient" is defined as the subjective belief on the M&S maturity, and it corresponds to the subjective nature of validation under large uncertainties. To assess if the model's credibility satisfies the requirement, a reward or utility value is assigned to each maturity level, and an expected reward/utility is calculated for the assessed model according to Eq. 64.

$$Expected\_Utility(model) = \sum_{j=1}^{N} P\left(Maturity_j | Evidence\right) \cdot U\left(Maturity_j\right) \qquad \text{Eq. 64}$$

where N represents the number of levels of maturity; $P\left(Maturity_j | Evidence\right)$ represents the belief on the statement that the model is at maturity level based on evidence; $U\left(Maturity_j\right)$

represents the utility/reward value assigned to each maturity level j. At the same time, a utility/reward criterion is calculated according to the required sufficient accuracy. Next, the expected utility is compared against the required utility/reward. If the expected utility is higher than the required value, the model credibility is claimed to satisfy the sufficient-accuracy requirement. Otherwise, it is claimed that the model is not optimal with respect to the sufficient-accuracy requirement. With the involvement of human artifacts, it becomes crucial to document all materials in a transparent manner. Table 5-1 shows an example of maturity level for the prototypic systems and the sub-scale systems when the uncertainties are large. In this study, the maturity for model credibility is defined by the quantitative assessment of predictive accuracy for all important QoIs in the prototypic system.

Table 5-1: Example of maturity level and their descriptions for situations with large uncertainties due to a lack of prototypic data. The attribute is designated according to the scaling analysis.

Maturity Level	**Descriptions**		
	Prototypic Systems	**Sub-Scale System**	
High	The model's predictive accuracy for all important QoIs in the prototypic system is high	Relevance	*Facilities that generates QoI data are highly relevant to the applications at conditions/geometries;*
		Scaling	*The QoI data come from systems with highly similar phenomena to the prototypic systems;*
		Uncertainty	*The uncertainty of QoI data has been well characterized;*
		Validation Result	*Model predictions for QoIs highly aligns with the data;*
Low	The model's predictive accuracy for all important QoIs in the prototypic system is low	Relevance	*Facilities that generates QoI data are hardly relevant to the applications at conditions/geometries;*
		Scaling	*The QoI data come from systems with different phenomena to the prototypic systems*
		Uncertainty	*The uncertainty of QoI data has not been characterized;*
		Validation Result	*Model predictions for QoIs poorly aligns with the data.*

Guided by the concept of sufficient accuracy, two technical terms are applied for practical applications: acceptance domain and bounded error. The acceptance domain represents a range of model credibility that can be considered acceptable. For example, if the rewards/utility is used for credibility assessment, any model with credibility higher than the requirement is considered acceptable. Besides, since the sufficient accuracy represents the credibility requirement of the models in simulating prototypic systems [48], it is necessary to have another term that represents the validation goals for sub-scale and separate-phenomenon models. Moreover, since the interactions of prototypic-system models and its sub-scale models are mostly nonlinear, the validation goal of sub-scale and prototypic-system models can be largely different. Similar to the term of acceptance domain, the validation goal of sub-scale models should be a range that "bounds" their errors. In this study, the inverse uncertainty quantification technique (inverse UQ) is used for inferring the validation goal of sub-scale and sub-phenomenon models. Detailed descriptions for

these terms are included in the glossary table at the end of this report. However, such inverse UQ requires data from prototypic systems. Considering practical challenges in obtaining prototypic data, the maturity level is usually divided into four attributes, and each of them is assessed separately at the sub-scale systems. Next, these separate assessments need to be integrated for informing the decision-maker to judge the model adequacy. In this study, the first three attributes (Relevance, Scaling, Uncertainty) are classified as Data Applicability. Detailed descriptions for each attribute can be found in N. Dinh's work [151] [152], while some case studies can be found in P. Athe's [153] work. Such technique requires many subjective assessments, and a properly formalized process is needed to ensure its quality.

Comparing to the accuracy concept in previous validations, the sufficient accuracy has three major differences. First, it introduces the subjective beliefs in addition to the accuracy standards and aims to establish that the belief is high enough to regard the M&S as a tool of certain maturity rather than estimating the precise accuracy. Though such concept appears to be trivial when conservative treatment is available, it avoids the excess conservatism while maintains convincing. This conservatism suggests that sufficient accuracy has to be transparent and convincing. At the same time, the analysis (of both logical and evidential) towards the sufficient accuracy should be complete. Second, it clearly separates the aspects of well-posed error/discrepancy calculation and the aspects of human confidence that are subject to large variations. The former aspect can be classified as a "causal relation" that is deterministic and obtained by direct comparisons. The latter one is "intangible" that is probabilistic depending on the scenario and human artifacts. As a result, this requires a properly document for human artifacts. Also, a robust and consistent process is needed for making confident decision regarding the M&S accuracy. Third, the concept of sufficient accuracy focuses on the area where large uncertainties exist. Instead of quantitatively defining the model credibility, the sufficient accuracy aims to bound the uncertainty by a qualitative statement. Therefore, as uncertainties being reduced, the process guided by sufficient accuracy should be improved. In conclusion, this study proposes a risk-informed validation framework that is transparent, consistent, and complete (at least improvable). To ensure the logical completeness and consistency, the framework is suggested to be built upon the existing studies, which is EMDAP in this work.

From a methodological point of view, the concept of sufficient accuracy shares many commons with the Risk-Informed or Risk-Oriented concept. Therefore, sufficient accuracy essentially represents a concept of risk-informed validation. Though the Risk-Informed/Risk-Oriented concept does not explicitly discuss the belief, it does stress the importance of agreement and confidence within the expert group: "ultimate safeguard is in wide and coherent expert participation and synergism…the basic premise is that once the whole community of experts in a given problem area is convinced, the problem may be considered solved". [154]

Table 5-2: Comparison of Sufficient Accuracy concept against Risk-Informed/Risk-Oriented concept.

	Sufficient Accuracy	**Risk-Informed/Risk-Oriented**
Large Uncertainty	Yes	Yes
Qualitative Assessment	Yes	Yes
Qualitative Levels	Maturity Level	Process Likelihood
Belief	Yes	Implicitly Yes

5.1.2. Case Study

To demonstrate the concept of sufficient accuracy, case studies are designed to demonstrate how sufficient accuracy is able to connect validation goal with safety analysis. Presently, a synthetic scenario is designed first, then two case studies are prepared: the first study aims to demonstrate the relationship between the subjective belief and the safety analysis; the second study aims to demonstrate the propagations of sufficient accuracy from prototypic systems to separate-effect models.

5.1.2.1. Scenario Description

It is assumed that there is a Pressurized-Water-Reactor-like system near a body of water, and there is a dam structure for preventing flooding. During an external-flooding hazard, an accident scenario of external-hazard-induced Station Blackout (SBO) is investigated. Figure 5-1 shows a schematic plot of the postulated scenario. At the same time, the response time is selected

as the QoI for the corresponding risk analysis. The response time is defined as the time it takes for one event consequence to occur starting from an initiating state. The progression of the accident is defined as six consequent events, and each of them has an initiating state and a consequence. In this study, the sub-events (2) and (4) are investigated, and physical models are used for calculating the corresponding response time. For simplification, sub-event (2) is named as landscape overflow, while sub-event (4) is named as vent overflow.

(3) If there is a heavy rainfall after a storm surge, water overflows the barrier and floods the NPP site.
(4) If the water level in NPP site reaches a critical value, the AC power fails immediately.
(5) If the water level is higher than the window of DG room air vents, water starts to overflow the room when.
(6) If the water level in DG room reaches a critical value, the DG power fails immediately.
(7) If both AC and DG power fail, the accident of SBO occurs

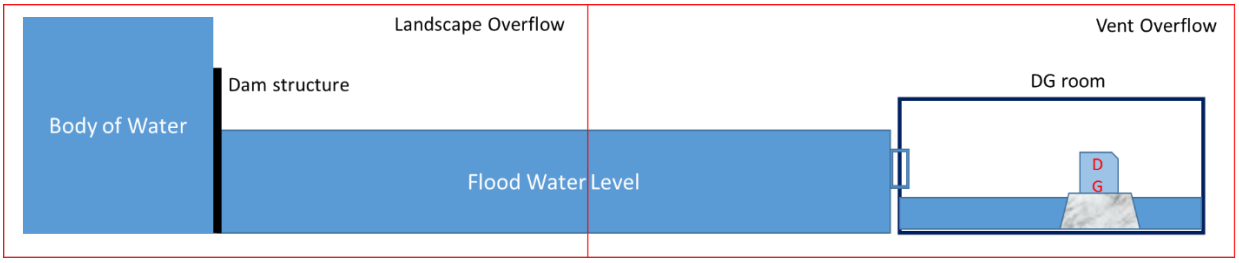


Figure 5-1: Demonstration of landscape (left) and vent (right) overflow scenario.

Given the assumption that the flood will distribute uniformly across the site, two Bernoulli-based model is applied for calculating the volumetric flow rate: for landscape overflow, a basic head-discharge coefficient equation (Eq. 65) is applied; for vent overflow, a simplified discharge model is used (Eq. 66). Furthermore, it is assumed that the discharge coefficient in both equations is the only error source for predicting the flow rate.

$$Q_{land} = c_d c_v \frac{2}{3}\sqrt{\frac{2}{3} g b_c}\, h_1^{1.5} \quad \text{Eq. 65}$$

$$Q_{vent} = c_d W \sqrt{2g(\Delta H - W/2)} \quad \text{Eq. 66}$$

where c_d is the discharge coefficient, c_v is the approach velocity coefficient, b_c is the width of the rectangular control session, h_1 is the upstream water depth, g is the gravity, ΔH is the water level above the DG room vent, W is the width of the DG room vent. To represent the transient process for the overflow, a simple function is made as Eq. 67, where the water depth will increase quickly to the maximum H_0 and gradually decrease until the time limit of 60min. Figure 5-2 shows the transient plot for the water depth h_1 given that the maximum depth H_0 is 2m. To account for the effects of uncertain scenarios, H_0 is assumed to be uniformly distributed in the range of 2.4m to 3.5m. According to Eq. 65, the landscape overflow rate to the reactor site (Q_{land}) can be determined.

$$h_1(t) = \begin{cases} H_0 \sin\left(\frac{\pi}{2}t\right), & t < 1min \\ H_0 - 0.002(t-1), & 60min > t \geq 1min \end{cases} \quad \text{Eq. 67}$$

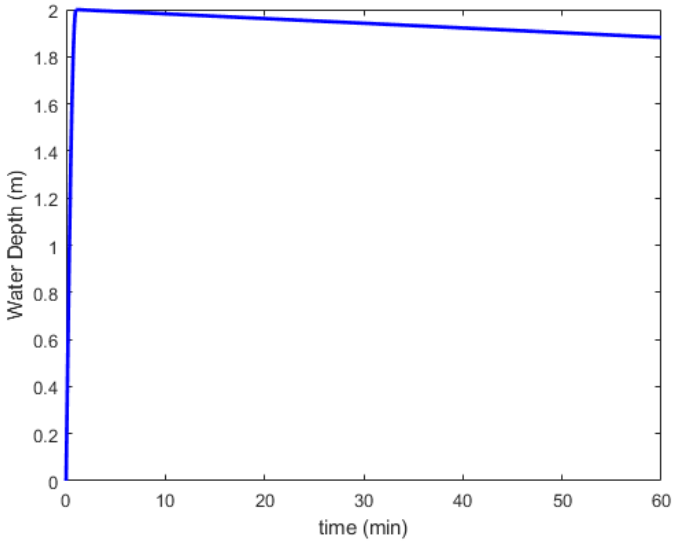


Figure 5-2: Transient plot of water depth for the synthetic overflow scenario with the maximum depth $H_0 = 2m$.

Next, the transient water level in reactor site and ΔH can be calculated by substituting Eq. 65 into Eq. 68, and the vent overflow rate into the DG room (Q_{vent}) can be determined according to Eq. 66. Finally, the transient water level in the DG room can be calculated Eq. 68. Assuming that the

critical water level is $2m$ for both the vent and landscape overflow, the response time of each event can be found as the time when water level first reaches the critical level.

$$h(t) = \int_0^t \frac{Q}{S} dt \qquad \text{Eq. 68}$$

5.1.2.2. Vent Overflow Sufficient Accuracy

In this case study, the critical belief on the model's maturity level for the discharge-coefficient model is estimated according to the safety analysis of the vent overflow. Before the demonstration, it is assumed validation data come from prototypic systems such that the predictive accuracy is quantitatively assessed by the prediction error. However, under the situation that the uncertainties of the validation process are large, such that the decision-maker cannot precisely represent the uncertainty with well-defined parametric distributions. As a result, qualitative analysis and the concept of sufficient accuracy is applied to guide the designation of validation goal. First, the model's accuracy is defined according to the error ε_{Cd} of model prediction (Eq. 65).

$$C_{d_{predict}} = C_{d_{real}} + \varepsilon_{C_d} \qquad \text{Eq. 69}$$

Next, according to the range of error ε_{Cd}, two maturity levels are designated for representing the credibility of model, and their definitions are shown in Table 5-3. To characterize the effect of subjective belief on the model's maturity level, this case study aims to determine the critical belief on maturity level, where the prototypic system can remain safe.

Table 5-3: Definition of sufficient accuracy and their probability density distribution for model prediction error.

Sufficient Accuracy	Probability Density
High	$\rho_{high}(\varepsilon_{C_d}) = U(0,0.1)$
Low	$\rho_{low}(\varepsilon_{C_d}) = U(0,1)$

In addition to the maturity level, subjective belief is also included for the designation of validation goal. In this study, for situations with partial belief on the model's maturity, the error probability distribution is calculated by Eq. 70, where $z \in [0,1]$ represents the subjective belief on the model's high maturity level. Figure 5-3 demonstrates the error distributions for high and low maturity level with 100% belief, and the one when belief on the model's high maturity is 50% (Eq. 70).

$$\rho_{mix}(\varepsilon_{Cd}) = z\rho_{high}(\varepsilon_{Cd}) + (1-z)\rho(\varepsilon_{Cd}) \qquad \text{Eq. 70}$$

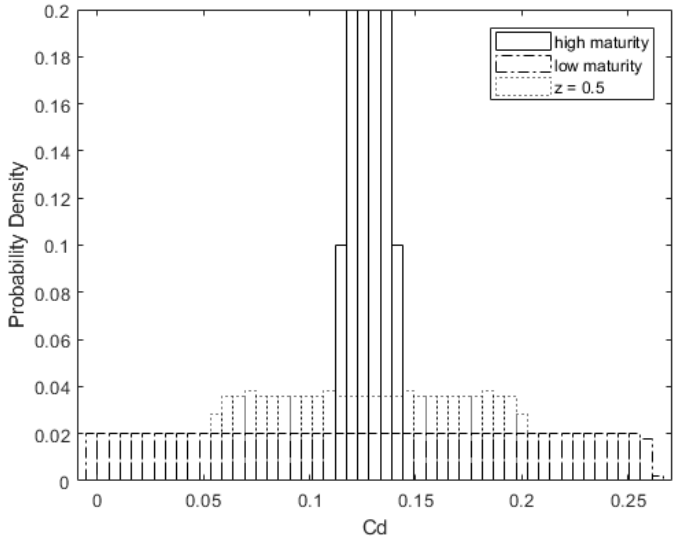


Figure 5-3: Demonstration of error distribution for high maturity, low maturity, and 50% belief on high maturity.

To demonstrate the relationship between validation goal, especially the subjective belief, and the safety analysis, a group of discharge coefficient C_d is first determined by sampling the subjective belief z on the model's high maturity level. Next, safety analysis is performed for each discharge coefficient, and a group of response time can be calculated. Since it is further assumed that the response time for the vent flow event should be higher than 30min, the critical subjective belief can be determined when the minimum response time equals to 30min. In this study, the critical subjective belief is defined as the situations where "model predictions will always satisfy the safety goal as long as the subjective belief on model's high maturity is equal or higher than the critical values". Figure 5-4 shows a plot of predicted response time against the belief on the

model's high maturity level, and the dashed line represents the safety requirement (30min). The intersection of two lines is the critical belief, and the safety goal can always be satisfied given the belief is higher than the critical value. It is found that when $\Delta H = 2.87m$, the safety goal can always be satisfied when the belief on the model's high maturity level is equal or higher than 0.6.

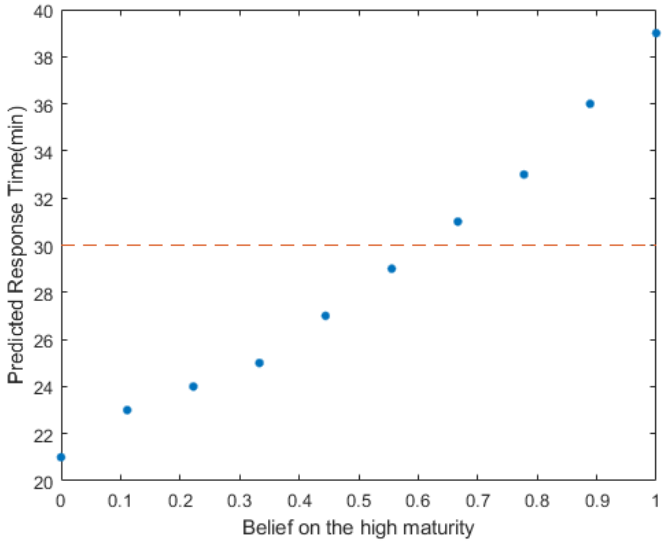


Figure 5-4: Plot of predicted response time against the belief on the model's high maturity level when $\Delta H = 2.87m$. The dashed line represents the safety requirement where the response time should be higher than 30min. The point that two lines intercept is defined as critical belief, where the safety goal can always be satisfied given the belief is higher than the critical value.

With the same strategy, the critical belief on the high-maturity model can be estimated with respect to each initial flood depth ΔH. Figure 5-5 plots the critical belief with respect to the initial flood depth, and the acceptance domain, located left of the function line, can be visualized. In addition, there are two regions in red boxes, where the subjective beliefs on model maturity levels do not affect the satisfaction of safety goal. Generally, these regions represent the situations where safety margins of investigated facilities or components are much higher (or lower) than the model uncertainties.

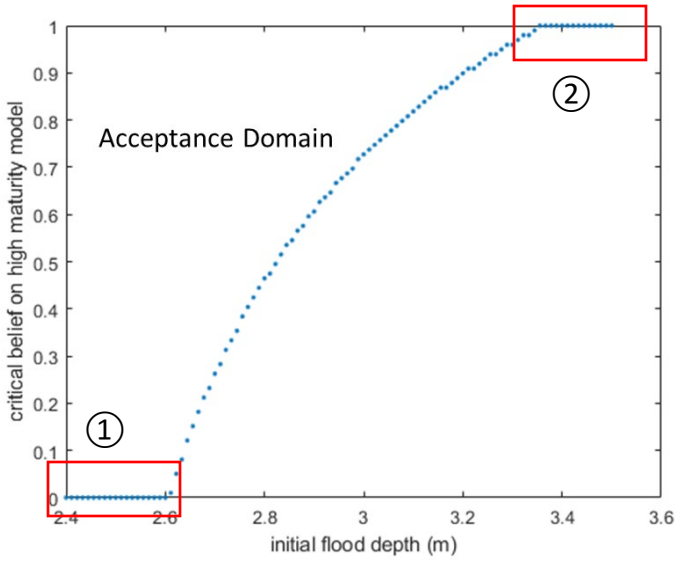


Figure 5-5: Estimated critical belief on high-maturity model versus the initial flood depth, the region on the left of function line is the acceptance domain. The regions in red boxes suggest the situations where subjective beliefs on model maturity levels do not affect the satisfaction of safety goal. The bottom box indicates the situations where safety margins of facilities or components are much more than the model uncertainties, while the upper box indicates that the margins are less than the model uncertainty.

In RISMC scenarios, assuming that the load distribution is calculated by the simulation uncertainties and the capacity distribution is predefined, region ① represents the scenarios, where load and capacity distributions have no overlap (). Meanwhile, for scenarios of region ②, the load distribution is so wide that it fully covers the capability distribution.

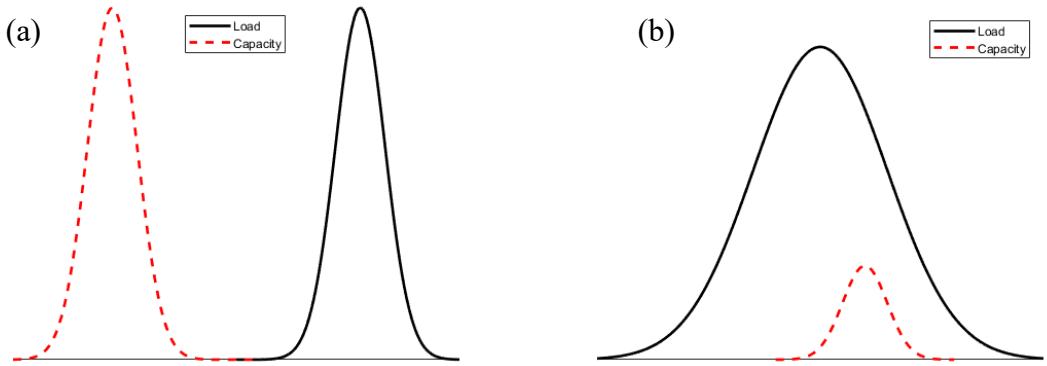


Figure 5-6: Demonstration of risk-informed load and capability distributions for (a) scenarios with no overlap between load and capacity distribution (b) scenarios where load distribution is so wide that it fully covers the capacity distribution. Scenario (a) corresponds to the region ①, and the safety goal is always satisfied given all ranges of subjective beliefs on the model's maturity. Scenario (b) corresponds to the region ②, and the safety goal can never be satisfied by whatever subjective beliefs.

5.1.2.3. Landscape Overflow Sufficient Accuracy

In this case study, the validation goal for the volumetric-flow-rate model is estimated according to the safety analysis of the landscape overflow. At the same time, the validation goal is propagated from the prototypic model to the sub-scale models. Again, it is assumed that validation data come from prototypic systems such that the sufficient accuracy is defined only by the validation result, which is represented by the error of model prediction (Eq. 71).

$$Q_{predict} = Q_{real} + \varepsilon_Q \qquad \text{Eq. 71}$$

Two maturity levels are used for representing the accuracy, and their definitions are shown in Table 5-4. Since the sufficient accuracy is represented by maturity level and belief, this case study aims to determine the required belief on maturity level according to the results of safety analysis.

Table 5-4: Definition of maturity level and their probability density distribution for model prediction error.

Maturity level	Probability Density
High	$\rho_{high}(\varepsilon_Q) = N(0,0.1)$
Low	$\rho_{low}(\varepsilon_Q) = N(0,1)$

In addition to the maturity level, subjective belief is also included for the designation of validation goal. In this study, the error probability distribution is calculated by Eq. 72 when the belief z on the model's high maturity level is changed. Figure 5-7 demonstrates the probability density distribution of model prediction error for 100% high and low maturity level, and the one when belief on the model's high maturity is 50% (Eq. 72).

$$\rho_{mix}(\varepsilon_{Cd}) = z\rho_{high}(\varepsilon_{Cd}) + (1-z)\rho_{low}(\varepsilon_{Cd}) \quad \text{Eq. 72}$$

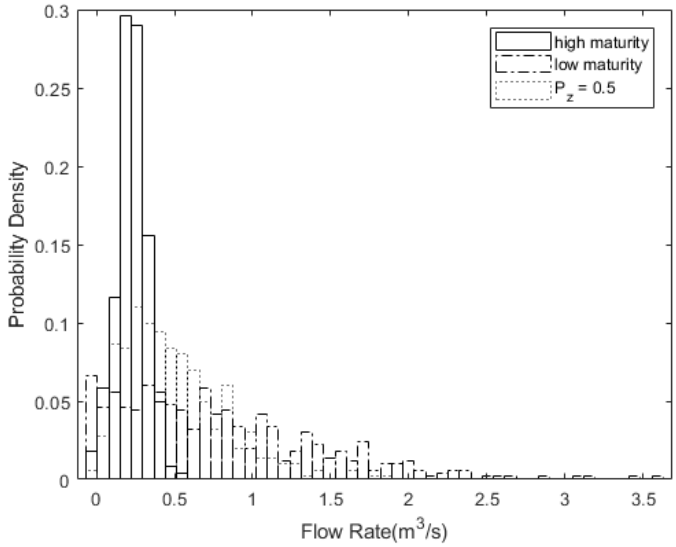


Figure 5-7: Demonstration of error distribution for high maturity, low maturity, and 50% belief on high maturity.

To demonstrate the relationship between subjective belief and the safety analysis, a group of flow rate Q is first determined by sampling the subjective belief z on the model's high maturity level (defined in Table 5-4). Next, safety analysis is performed with the group of flow rate, and it is assumed that the response time for the vent flow event should be higher than 30min. Besides, to

account for the effect of uncertain scenario in RISMC, it is assumed that the value of flow rate Q is uncertain and uniformly distributed from $0\ m^3/s$ to $2.2\ m^3/s$. Figure 5-8 shows a plot of predicted response time against the belief on the model's high maturity level, and the dashed line represents the safety requirement (30min). It is found that when flow rate $Q = 0.5m^3/s$, the safety goal can always be satisfied when the belief on the model's high maturity level is equal or higher than 0.6.

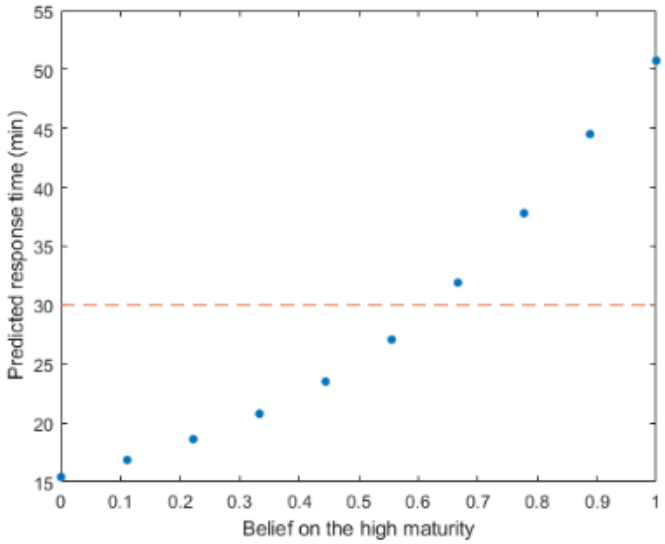


Figure 5-8: Plot of predicted response time against the belief on the model's high maturity level when $Q = 0.5m^3/s$. The dashed line represents the safety requirement where the response time should be higher than 30min. The point that two lines intercept is defined as critical belief, where the safety goal can always be satisfied given the belief is higher than the critical value

With the same strategy for vent overflow model, the critical belief on the high-maturity model can be estimated with respect to each flow rate Q. Figure 5-9 plots the critical belief for the corresponding flow rate, and the acceptance domain, located left of the function line, can be visualized. Compared to the error that is drawn from a uniform distribution, the estimated critical beliefs are fluctuating and scattering around a given flow rate.

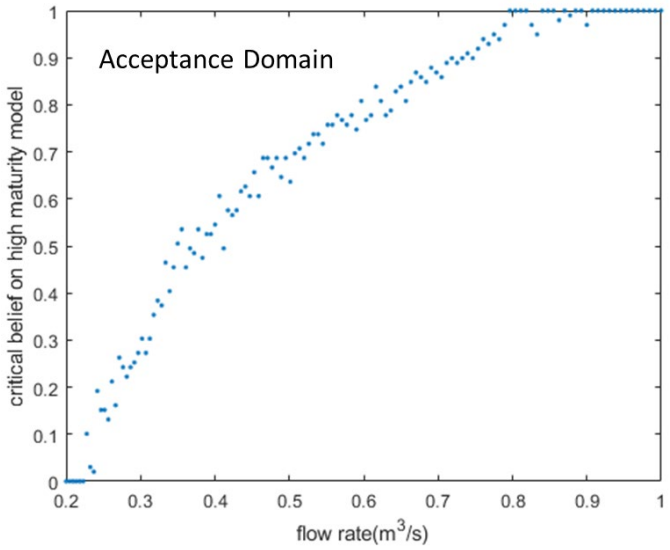


Figure 5-9: Estimated critical belief on high-maturity model versus the flow rate, the region on the left of function line is the acceptance domain.

Besides, to demonstrate the propagation of validation goal from prototypic model to sub-scale model, a sub-scale model of Eq. 73 is used for calculating the discharge coefficient in Eq. 65. In this case study, a validation goal is first defined for the flow rate predicted by Eq. 65. Next, the validation goal for parameters α and β will be inferred according to the goal of integral models. According to the concept of sufficient accuracy, the validation goal is defined as a combination of subjective belief and maturity level. The definitions for maturity level are listed in Table 5-4. Since the correlation between α, β and Q are nonlinear, the maturity level definition for α and β will not follows the definitions in Table 5-4. Therefore, the concept of bounded errors is used for representing the error range of sub-scale or separate-effect models. In this study, the propagation of validation goal is achieved by inverse uncertainty quantification, and the goal is to find the error range of parameter α and β according to the subjective belief on the maturity of flow rate prediction. A statistical toolbox of Bayesian Inference named DRAM [155] is used to inversely inferred the distribution of parameters α and β with the Bayes' theorem.

$$C_d = \left(\frac{H_1}{L} - \alpha\right)^{\beta}$$

$$\alpha = \alpha_{real} + \varepsilon_{\alpha} \qquad \text{Eq. 73}$$

$$\beta = \beta_{real} + \varepsilon_{\beta}$$

At the same time, the effect of subjective belief is also considered by sampling parameter z, and the error distributions for the predicted flow rate with respect to each belief can be described by Eq. 70. In this case, the belief on the prototypic model's maturity level is sampled uniformly from 0 to 1, and one hundred Bayesian Inferences are performed for each belief level. The distributions of parameters α and β are extracted from each inference, and the error bounds are identified by finding the maximum and minimum of all parameter distributions. Furthermore, the inferred parameters distributions are combined and forward propagated to determine the range of flow rate. Figure 5-10 shows the error bounds that are inferred by different subjective beliefs on the model's maturity in predicting the flow rate. The error bound is expanding as the belief on model's high maturity becomes less. Also, it is found that the Bayesian Inference will not converge to stable parameter distribution when the belief z is lower than 0.65. Although the concept of sufficient accuracy allows for validation goal with beliefs from 0 to 1, when sub-scale models are incorporated, a number of beliefs for the integral-model predictions become unacceptable. And the range is "constraint" because some errors of crude models are not explanatory by the sophisticated models. This finding is usually known as the "physics-constraint" concept, where the range of predictions are constrained by assumptions of models.

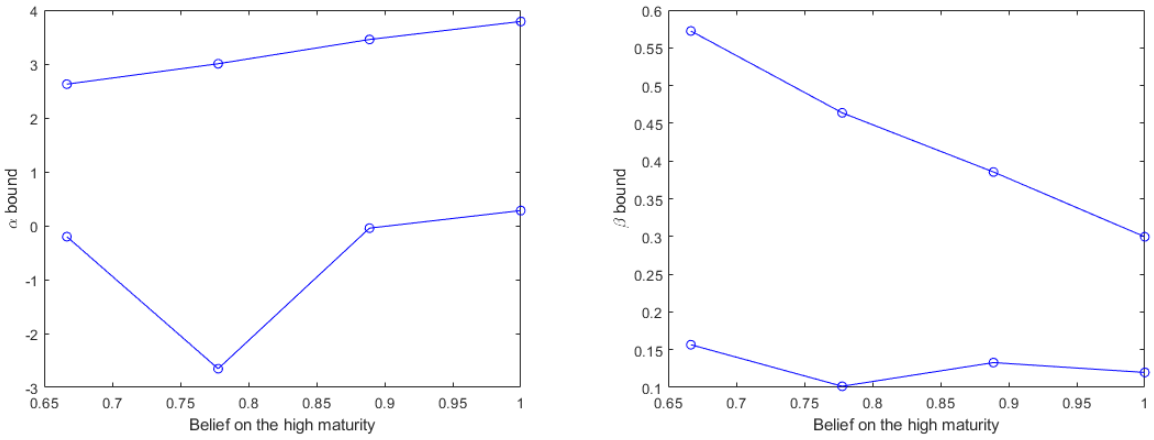


Figure 5-10: Demonstration of error bounds for parameter α and β.

To demonstrate this finding, the inferred parameters are substituted back to Eq. 73 and Eq. 65 for determining the error range of flow rate prediction. And Figure 5-11 compares the range of

physics-constraint flow rate against that by direct sampling the error distribution from maturity definitions (Table 5-4). It is found that the inferred error bounds are much narrower than those from direct sampling since the sub-scale model (Eq. 73) eliminates samplings that violate the physical model form. In another word, the degree of freedom from direct sampling is higher than the one from model-form-inferred sampling. As a result, the development of sub-scale model (Eq. 73) can reduce the model uncertainty and increase the credibility. However, it only applies when the accuracy of prototypic models is higher than a critical value since the sub-scale model is not able to explain large errors in predicting the prototypic systems. In this case, the critical belief is found to be 0.65.

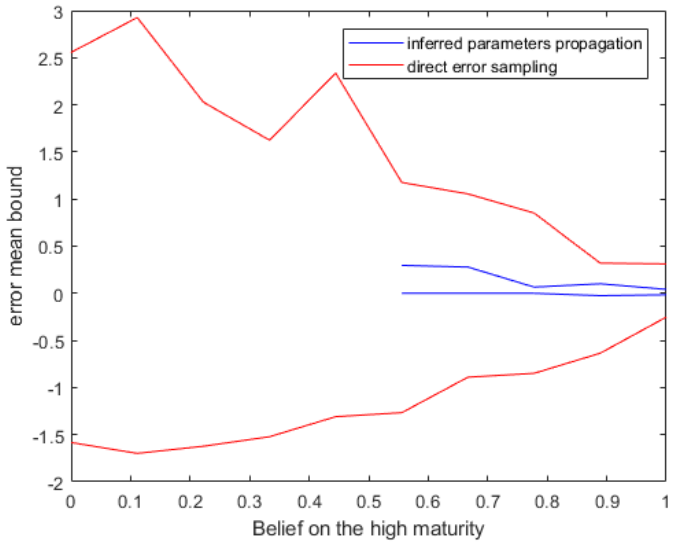


Figure 5-11: Comparison of errors bounds for predicting the flow rate Q with inferred parameters against the error bounds with direct sampling.

Figure 5-12 shows the relationship of subjective belief and response time for cases using inferred parameters and those using direct sampled errors. It is found that the error bounds form inferred parameter distributions are much narrower than those from direct sampling according to the error distribution of given maturity levels. Such finding is expected since most errors are screened out by physics, and the physics-constraint models can be more precise than those fitted by crude statistical models. Besides, it is found that the critical belief values from inferred parameters are higher than those from direct sampling. And the inferred parameters become unavailable when the

belief is lower than 0.6. Such finding suggests that scenarios with complex physics require higher confidence in validation than those with simple physics.

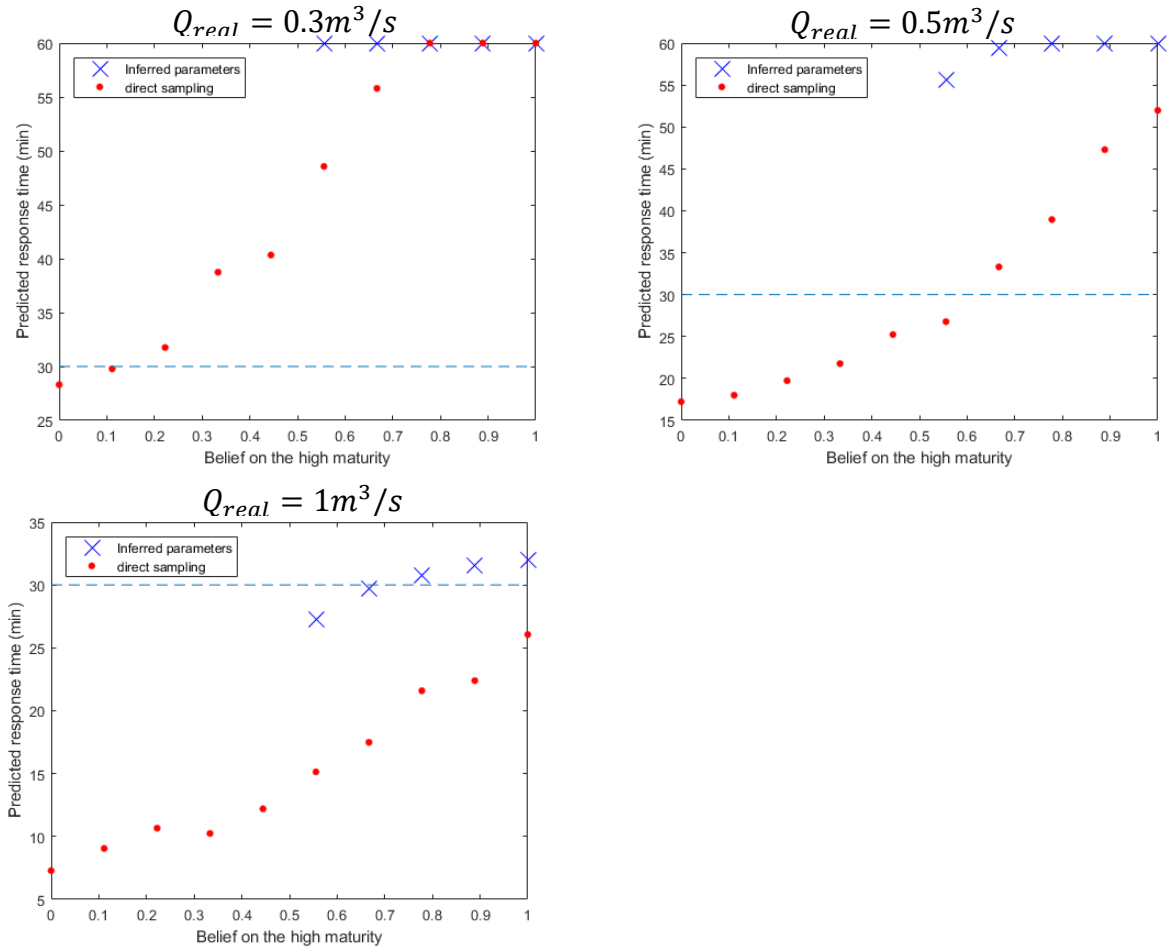


Figure 5-12: Relationship of subjective belief and response time with respect to three uncertain scenarios. Comparisons are also made for each scenario by using the inferred parameters and the direct sampled prediction error.

5.1.3. Findings and Issues

In this section, a concept of sufficient accuracy is developed for guiding the designation of model credibility requirements according to the safety analysis for prototypic systems. The concept of sufficient accuracy combines both probabilistic and deterministic analysis, and the validation goal is represented by subjective belief and maturity levels. Comparing to the regular approach for designating validation goals, the sufficient accuracy explicitly connects accuracy requirements

with safety analysis. At the same time, the sufficient accuracy admits the qualitative nature of model validation, and it aims to document involving human artifacts to remain convincing. Moreover, the sufficient accuracy aims to avoid the excess conservatism by seeking the critical belief to regard the model as acceptable. To demonstrate the concept and application of sufficient accuracy, two case studies are performed, and the critical beliefs on the model maturity are found according to the safety requirements and uncertain scenarios. The case study also shows the process of propagating the validation goal from prototypic models to sub-scale model. It is found that: with the same sufficient accuracy, the critical beliefs assessed by sub-scale models are higher than those by prototypic models, because the predictions from sub-scale models have much narrower error range than those from prototypic models. It is also found that the sub-scale models restrict the distribution of prediction errors, and it fails to be self-explanatory when the confidence on model's credibility is low.

Presently, the application of the sufficient accuracy concept is limited to the validation processes with large uncertainties (Level 2 and 3), but not deep uncertainties (Assumption A4). For situations with level 1 uncertainty, the error can be precisely quantified with distributions, and no qualitative analysis is needed. For situations with deep uncertainties, no reliable approach has been found in this study, and more investigations are still needed. Meanwhile, the concept of sufficient accuracy is more applicable when the validation process is required to be both efficient and convincing. This study investigates the relationship of subjective beliefs and assumes that the maturity levels are objective and prescribed. However, in many scenarios, the designation of model maturity levels can be subjective, and its effects have not been looked into. Besides, this study uses Bayesian Inference to propagate the error range from prototypic systems to sub-scale models, which only applies to scenarios with Level 1 uncertainty. Also, the inverse UQ techniques depends heavily on data coverage and data quality, their effects on the propagation of sufficient accuracy have not been included.

5.2. Predictive Capability Maturity Quantification by Bayesian Network

To formalize and quantify the decision-making process in the validation process, a decision-theoretic framework named Predictive Capability Maturity Quantification by Bayesian

Network (PCMQBN) is developed. The goal is to represent the validation decision-making process in a transparent, consistent, and improvable manner. After the technical descriptions, case studies are performed to demonstrate the application of PCMQBN. Finally, findings and limitations are discussed according to the case studies and literatures.

5.2.1. Technical Development

With the rapid growth in computer power, modeling and simulation have been widely used to support the decision regarding design, operation and safety of the nuclear power plant. As a result, validation has become crucial for assessing the credibility of M&S. Due to the subjective nature of validation process, various validation frameworks, including EMDAP, PCMM, and AIAA guide, have been proposed for formalizing the process. And the goal is to assure the process quality and make convincing arguments. However, in EMDAP framework, the decision-making for assessing the model credibility has not been formalized. As a result, it becomes hard to perform a credibility assessment independently or to reach an agreement regarding the model's credibility collectively. Moreover, such an issue could become severe when the uncertainties become larger. Although the validation adequacy is required to be assessed, focuses have been put on reconciling diverging expert opinions. It is suggested in this study that the decision-making of M&S credibility under uncertainties should be further formalized with a sufficient resolution, structured knowledge base, and sophisticated inference engine. The goal is to support decision-making of complex validation problems by reasoning through bodies of knowledge, which includes objective data, causal relationships, and subjective opinion. Also, a properly formalized process can help characterize and document relevant materials, including objective evidence, expert opinion, subjective beliefs, and so on. Besides, to further ensure the consistency of validation process, it is suggested that the objective of the validation framework is to make the most robust estimation for model credibility rather than searching for the best predictive model.

In this study, the validation decision-making is formalized as an argument process that aims to argue the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model (Assumption B1). Such definition is learnt from a similar definition by W.L. Oberkampf [43]. Differently, the new definition adopts the concept of

safety case [156], which emphasizes the subjective nature of validation and the role of reasoning. Several safety-case-based validations, known as "validation case", have been performed for evaluation models [153] [157], and Goal Structure Notation (GSN) is used as the graphical representation for the validation argument. Presently, evidence, structured argument, and decision criteria are three major components that form the reasoning toward model credibility. Following paragraphs will introduce the development of each component.

There are various ways of characterizing evidence. Sun [158] categorizes evidence as direct evidence, backing evidence and counter evidence based on their association with the confidence in the safety case. Similar to Athe's study [153], the direct evidence for validation in this study is categorized into two sub-attributes: Validation results and Data applicability. The evidence associated with validation result is characterized by comparing the simulation result with the expected value, while the evidence of data application is characterized by R/S/U grading system [152]. The R/S/U grading system has three sub-attributes: (a) relevance, (b) (physics) scaling, (c) data uncertainty, and the goal is to assess the quality of experiment and data. The relevance (R) is determined based on the geometric similarity and material scaling; the (physics) scaling (S) measures the degree of similarity between phenomena in the application and experiment on the basis of physics scaling; the uncertainty (U) measures the data uncertainty due to instrumentation errors and limited resolution of measurement instruments. Different from the previous study, the evidence is further characterized according to concept of sufficient accuracy, where maturity level and subjective belief are used at the same time for representing the assessment results of each attribute. Table 5-5 shows an example of descriptor for validation result and R/S/U attributes. It is found from literature [153] [151] that in addition to the numerical error, another attribute named "Coverage" is used for assessing the attribute of validation result. Presently, the quantification of the coverage attribute is still being investigated, and this study will only consider the evidence of numerical error.

Table 5-5: Grade table for validation experiment.

Validation results (VR_i) $m_j = \{VL, L, M, H\}$		R/S/U $m_j = \{VL, L, M, H\}$	
Descriptor	Meaning	Descriptor	Meaning
VL	$Norm\_Error \in [0.7,1)$	VL	$R/S/U = 1$
L	$Norm\_Error \in [0.4,0.7)$	L	$R/S/U = 2$
M	$Norm\_Error \in [0.1,0.4)$	M	$R/S/U = 3$
H	$Norm\_Error \in [0,1]$	H	$R/S/U = 4$

Comparing to the maturity definition in Table 5-1, attributes of PCMQBN compose the general description of maturity level in the validation goal. The reason is that the sufficient-accuracy-guided validation goal is designated for the prototypic systems, while PCMQBN aims to assess the model credibility according to evidence from small-scale or separate-effect facilities. Presently, the data applicability and numerical R/S/U grades are designated by scaling analysis and expert opinions[2].

To integrate all related evidence and represent the process by mathematical languages, the Bayesian decision theory (Assumption B2, B3, and B4) is applied to determine the model credibility based on the characterized evidence. The Bayesian decision theory uses the Bayesian conditionalization to represent the belief updates by obtaining and integrating new evidence. Eq. 74 shows the calculation of model credibility, named as Code Adequacy (CA) by integrating evidence of Validation Result (VR) and Data Applicability (DA).

$$p\left(CA_j \middle| DA, VR\right) = \sum_i P\left(CA_j | DA_i\right) P(DA_i) + \sum_i P(CA_j | VR_i) P(VR_i) \qquad \text{Eq. 74}$$

where CA_j represents the maturity level of the code adequacy, $P\left(CA_j | DA, mVR\right)$ represents the subjective belief of obtaining the model maturity CA_j based on evidence from Data Applicability (DA) and Validation Result (VR). $P\left(CA_j | DA_i\right)$ and $P(CA_j | VR_i)$ represents the update rule for code adequacy given the evidence from data applicability and validation result respectively.

[2] Detailed descriptions of PCMQBN's attributes can also be found in the glossary table at the end of this document

$P(DA_i)$ and $P(VR_i)$ represent the subjective belief on the maturity level i for data applicability and validation result. Considering the complexity of validation and the amount of evidence, this study uses a graphical model of Bayesian Network to represent the code adequacy calculation. Bayesian Network is first introduced by Pearl [159] as a directed acrylic graph (DAG) for probabilistic reasoning. Presently, the Bayesian Network has been extensively used to model belief in biological science, medicine, forensic science, law, risk analysis, and so on [160] [161]. Recently, researchers also apply Bayesian Network to model validation [162], where the network structure is mapped from PRA analysis. In this study, the network structure is constructed by the concept of validation cubic. The concept of validation cubic is first proposed by N. Dinh's [151], and the objective is to integrate and analyze the uncertainty of a computer model based on validation data from experiments of different scales. Validation cubic is a three-dimension diagram that plots the reactor prototypicality parameter (RPP) against the system and physics decomposition. RPP represents the global accuracy of the computer model, the system decomposition represents the experiments and scenarios designed for the full-scale application. The physics decomposition represents the model developed for every phenomenon in the phenomenon pyramid. To account for the relative importance of various evidence, a weight factor is used and calculated as Eq. 75 in the validation cubic model [163] [152].

$$\psi_{E_i} \sim \left(1 - EMU_J\right)^m \left[RPP_{K,J}\right]^n \qquad \text{Eq. 75}$$

where ψ_{E_i} is the weight factor for an evidence E_i; EMU is Experimental Measurement Uncertainty that measures the VUQ quality of a certain experiment. m and n are constants that represent the significance of experiment J and the physics K. RPP, Reactor Prototypicality Parameter, "represents the similarity of validation experiments to the application condition" [163] [152]. It can be defined as the ratio of the governing scaling parameter Sc_{Mod_K} for "model" of physical process K calculated for test/experimental conditions, $\left[Sc_{Mod_K}\right]_{EXP}$ and reactor application's conditions $\left[Sc_{Mod_K}\right]_{APP}$(Eq. 76):

$$RPP = \left[Sc_{Mod_K}\right]_{EXP} / \left[Sc_{Mod_K}\right]_{APP} \qquad \text{Eq. 76}$$

For the study of PCMQBN, the EMU is determined by the attribute of data applicability. The constant m, which represents the experimental significance, is assumed to be mainly affected by the repeatability of validation experiments for a certain physics. In another word, if multiple experiments are designed for a single physics model, then the value of m will decrease as number of experiments increases. The constant n is assumed to be determined by the importance of models or physics, which is based on sensitivity study or experts' opinion. After determining the weight factor ψ_{E_i} for each evidence E_i, they are normalized to $\tilde{\psi}_{E_i}$ according to Eq. 77 and used for updating the belief.

$$\tilde{\psi}_{E_i} = \psi_{E_i} / \sum_{i=1}^{n} \psi_{E_i} \qquad \text{Eq. 77}$$

Figure 5-13 shows examples of Bayesian network with equal and unequal weight factor. Though they have the same end nodes, the resulting adequacies turn to be different. Currently, the conditional probability of each maturity level is assumed to be the same across all parent nodes such that the posterior belief in child node is a linear assembly of all prior beliefs and their corresponding weights in parent nodes.

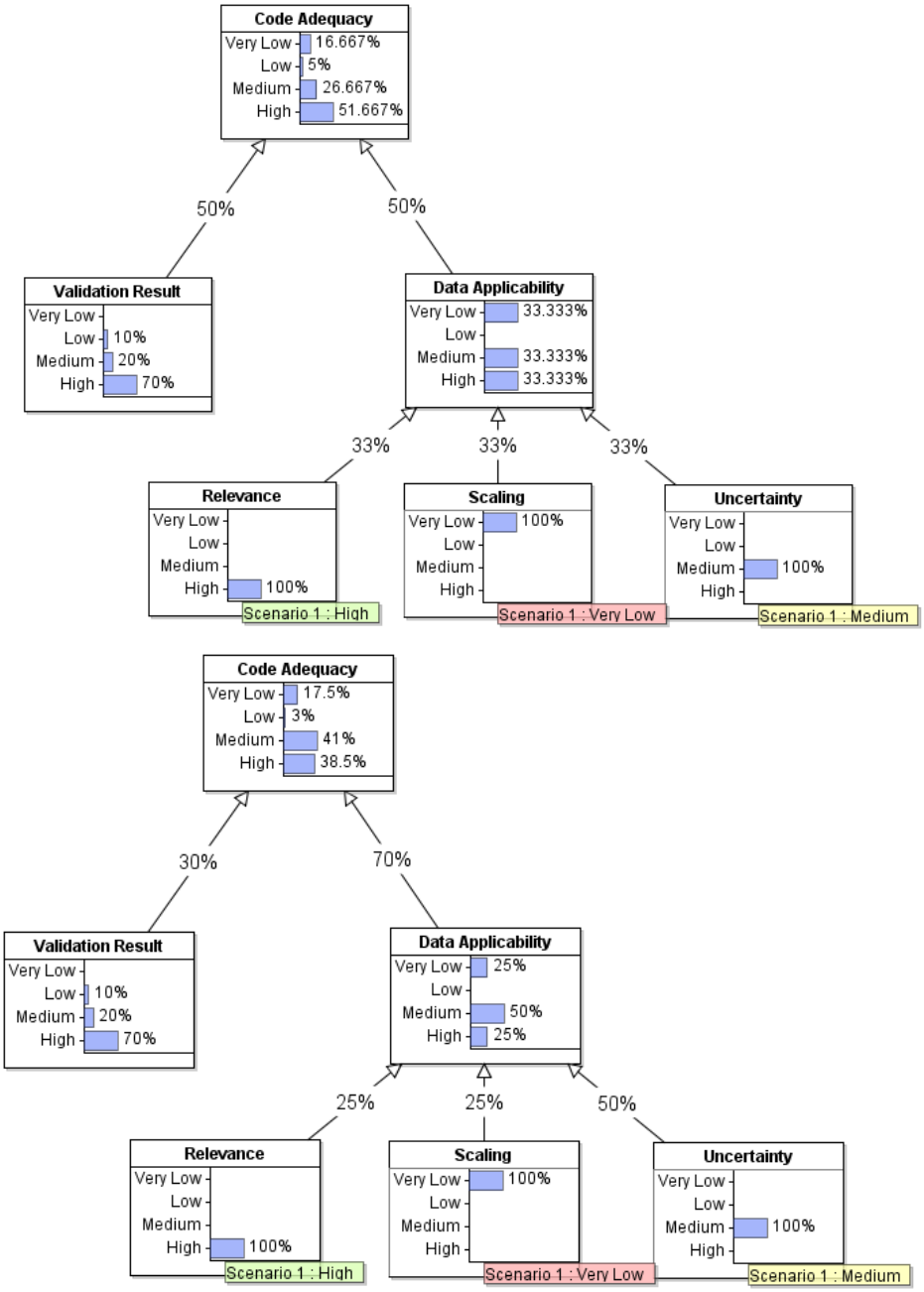


Figure 5-13: Example of Bayesian network for a "separate" code adequacy assessment with equal and unequal weights. The plot is prepared with AgenaRisk [164]

To assess if a model credibility satisfies the requirement and to compare the credibility of various models, a reward or utility value is assigned to each maturity level, and an expected reward/utility is calculated for the assessed model according to Eq. 78.

$$EU(a) = \sum_{j=1}^{N} P(CA_j|DA, VR) \cdot U(CA_j) \quad \text{Eq. 78}$$

If a utility is assigned, $U(mCA_j)$ represents the agent's preference on the corresponding maturity level. Otherwise, the reward simply represents the benefits of obtaining a certain maturity level. It is required that $U(mCA_j)$ is a bounded function such that Eq. 70 is always finite. Eq. 70 is essentially the same as Eq. 64, while the model credibility is named as code adequacy. Since the extraction of utility functions requires complicate elicitation process, this study calculates the Expected Monetary Value (EMV) with pre-defined monetary rewards. Table 5-6 shows two examples of assigning rewards value to maturity level. In the following case study, the value #1 will be used for assessing if the model has satisfied the requirement and which model has the best predictive capability.

Table 5-6: An example of monetary rewards assigned to each maturity level

Maturity Level	**Value #1**	**Value #2**
Very Low (VL)	$ 1.0	$ -100
Low (L)	$ 1.5	$ -50
Medium(M)	$ 2.0	$ 50
High (H)	$ 2.5	$ 100

5.2.2. Case Study

To demonstrate the application of PCMQBN, case studies are designed to demonstrate how PCMQBN can be used for simulation models. The scenario is the same as the one for the sufficient accuracy. The first case study demonstrates how PCMQBN decide the model credibility according to posterior information, where the measured QoIs are used for the model calibration; while the second case study demonstrates the same process, but the model predictions are made prior to the acquisition of application QoI. Decisions regarding the selection of best model are made according to validation goals.

5.2.2.1. PCMQBN for Landscape Overflow Model

In order to assess the credibility of head-discharge equation (Eq. 65) in scenario landscape overflow, which is composed of two separate equations (Eq. 73 and Eq. 79). Three sets of data are extracted for the credibility assessment of the landscape overflow model. Figure 5-14 shows the schematic plot of phenomenon decomposition and their corresponding models. The Integrated Effect Test (IET) measures the flow rate with respect to the total energy head H_1, and an IET-based model is obtained by calibrating a, b and u simultaneously. There are two Separate Effect Tests (SETs) that measures the discharge and approach velocity coefficient with respect to the total energy head respectively. A SET-based model is then obtained by calibrating a, b in Eq. 73 and u in Eq. 79 separately. All model parameters are calibrated by a Bayesian Inference software package named PYMC3 [165]. In this work, the validation data of IET are obtained from the empirical correlation, whose parameters are determined by the table of rating equation parameters and ranges of application (Table 8-4 in Ref. [166]). The validation data for two SETs are extracted from reference documents (Figure 1.12 and Figure 4.10 in [167]).

$$C_v = \left(\frac{H_1}{h_1}\right)^u \qquad \text{Eq. 79}$$

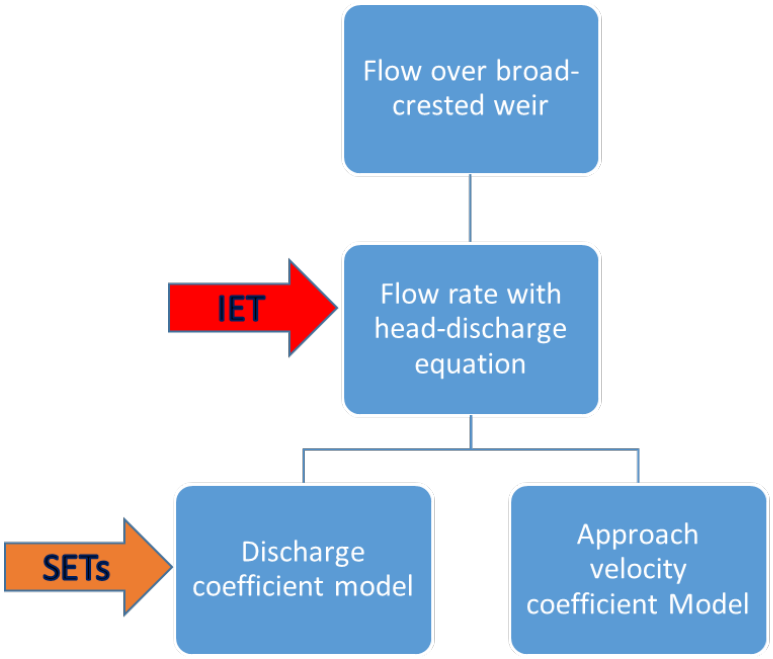


Figure 5-14: Demonstration of phenomenon decomposition for flow over the broad-crested weir. The IET measures the flow directly with respect to upstream water depth, while two SETs measure the discharge coefficient and the approach velocity coefficient respectively.

Figure 5-15 shows the plot of two model predictions and the experimental measurements. Table 5-7 shows the maturity ranking of the model's validation result, which is calibrated based on the SETs and IET respectively. The R/S/U grades of data applicability for three tests are also listed. There turns to be a large model inconsistency between the model and its components. When the model is calibrated based on SET data, the result matches the experimental C_d and C_v with an average normalized L_1 error ~ 0 (High maturity). However, the predicted flow rate turns to deviate from the measurement with an average L_1 error larger than 1 (Very Low maturity). Correspondingly, when the model is calibrated based on IET data, model prediction on flow rate match the measurement with the error ~ 0 (High maturity), while the predicted discharge coefficient deviates from the measurement with the error > 1 (Very Low maturity).

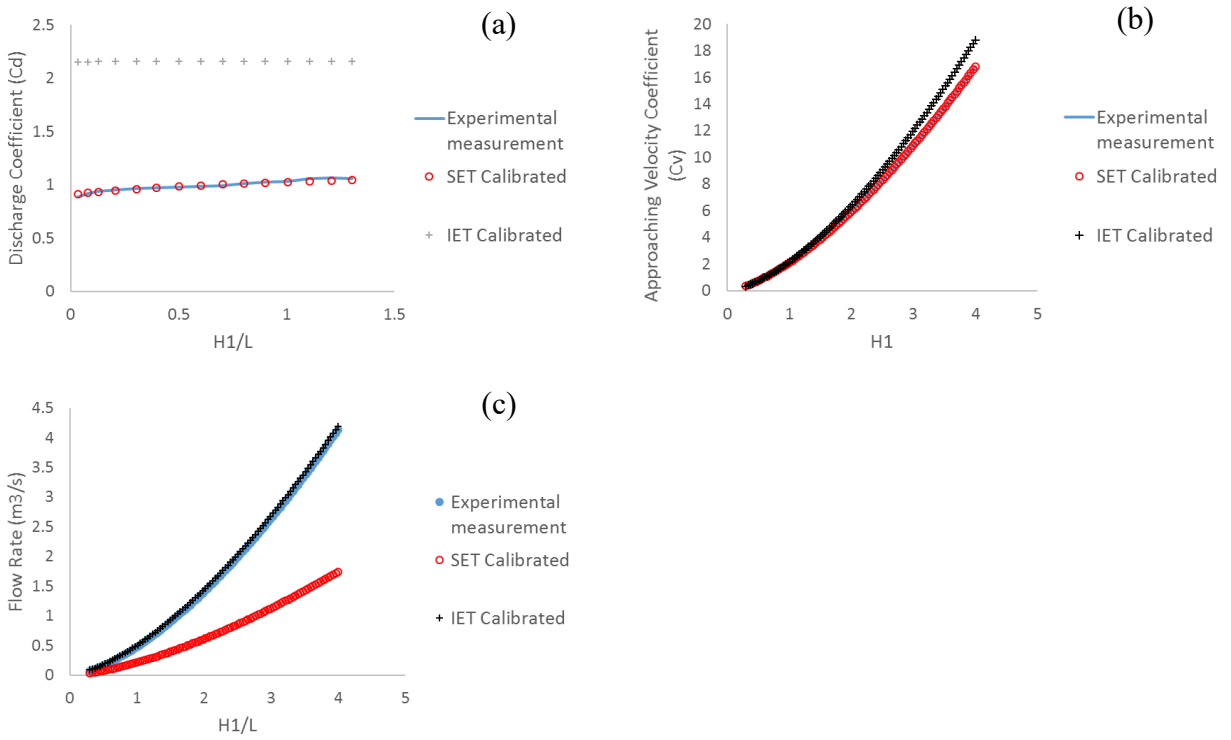


Figure 5-15: Plot of model predictions for SET and IET calibrated parameters versus experimental data. (a) discharge coefficient; (b) approach velocity coefficient; (c) flow rate.

Table 5-7: Value of model parameters calibrated based on SET and IET.

		RSU Grade			Validation Result's Maturity	
		R	S	U	SET-based	IET-based
SETs	C_d	L	H	VL	H	VL
	C_v	L	H	VL	H	H
IET	Q	H	H	VL	VL	H

Based on the maturity ranking of the validation result and R/S/U grades for three experiments, PCMQBN structures are constructed for two calibrated landscape overflow model. Furthermore, utilities are assigned to each maturity level based on Value #1 of Table 5-6. The EMV of every code adequacy can then be determined based on the probability and the utility of each level. In this study, cases with both equal and non-equal weight factors are tested. For the non-equal case, the weight factor of IET is assumed to be twice as much as the one for any SET, Table 5-8 shows the code adequacy and the corresponding EMVs. If an equal weight factor (33%)

is assigned to all experiments, the code adequacy of SET and IET calibrated model turns to be the same. If non-equal weight factors are assigned, the model calibrated based on IET data turns to have higher EMVs, which indicates that the IET-based model is more preferable.

Table 5-8: EMVs of the model adequacy for different combinations of model and weight factor.

		VL	L	M	H	EMV ($)
Equal	SET-based	33.3%	11.1%	0	55.6%	1.9
	IET-based					
Unequal.	SET-based	41.7%	8.3%	0	50%	1.8
	IET-based	29.2%	8.3%	0%	62.5%	2.0

Besides, if the validation goal for the same phenomenon (section 5.1.2.3) is employed, it is found that all models will be acceptable when the scenario flow rate $Q = 0.3m^3/s$; only the IET-based model is acceptable when the weight factors are not equal and the scenario flow rate $Q = 0.5m^3/s$; no model is acceptable when the scenario flow rate $Q = 1m^3/s$.

5.2.2.2. PCMQBN for Core Flow Pressure Loss Model

In the scenario description of section 5.1.2.1, a system-thermal-hydraulics simulation is needed for predicting the response time of event progression (6). Although the validation of system codes is not the objective of this study, the numerous physical models and the complex phenomena do provide a better demonstration for PCMQBN, especially in assessing model credibility and selecting models. A similar validation process is also applied to the single-phase flow pressure loss model (Eq. 80) of system thermal-hydraulics. The QoI is designated as the pressure drop across the reactor core due to wall friction and grid spacers. In calculating the pressure loss, two coefficients, including the wall friction f_{fric} and form loss coefficient K_{form}, are needed.

$$\Delta P_{TS} = \Delta P_{wall\ friction} + \Delta P_{grid\ spacer} \quad \text{Eq. 80}$$

$$\Delta P_{wall\ friction} = f_{fric} \times \left(\frac{L}{D_e}\right) \times 0.5 \times \rho v^2 \quad \text{Eq. 81}$$

$$\Delta P_{grid\ spacer} = N_{spacer} \times K_{form} \times 0.5 \times \rho v^2 \quad \text{Eq. 82}$$

where L is the length of test session; D_e is the equivalent hydraulic diameter of the flow channel; ρ is the density of coolant through the test session; N_{spacer} is the number of spacer grid; v is the velocity of coolant calculated by Eq. 84; G is coolant mass flow rate; A_{flow} is the unobstructed coolant flow area.

$$v = \frac{G}{A_{flow} \cdot \rho} \quad \text{Eq. 83}$$

Experimental measurements from PISA and KALLA facility are extracted to determine the maturity of validation result. Note that both facilities measure the pressure drop across the spacer grid (SET) and the whole test session (IET). The weight factor ratio of SET to IET is determined to be 1:2. Different from the previous study, the measured quantity is not used for calibrating the candidate models. By comparing conditions of two validation experiments to a postulated application, their R/S/U grades and weight factors can be determined as in Table 5-9.

Table 5-9: R/S/U grades and weight factors for data from PISA and KALLA facility.

		RSU Grade			Weight factor	
		R	S	U		
PISA	$\Delta P_{grid\ spacer}$	H	H	L	22.8%	9.6%
	ΔP_{TS}				45.6%	19.2%
KALLA	$\Delta P_{grid\ spacer}$	H	M	L	10.5%	4.4%
	ΔP_{TS}				21.1%	8.9%

In this study, Blasius (Eq. 84) and Colebrook-White (C-W) correlation (Eq. 85) are applied to calculate the friction coefficient, while Rehme (Eq. 87), Shiralkar (Eq. 88) and a constant-value model (Eq. 86) are applied to determine the form loss coefficient.

$$f_{fric} = 0.316/Re^{0.25} \quad \text{Eq. 84}$$

$$f_{fric} = \max(\frac{64}{Re}, f_{turb}) \quad \text{Eq. 85}$$

$$f_{turb} = 0.0055[1 + \left(2 \times 10^4 \frac{r}{d_{hyd}} + \frac{10^6}{Re}\right)^{1/3}]$$

$$K_{form} = 1.2 \quad \text{Eq. 86}$$

$$K_{form} = C_v \varepsilon^2$$

$$C_v = \min(3.5 + \frac{73.14}{Re^{0.264}} + 2.79 \times \frac{10^{10}}{Re^{2.79}}, \frac{2}{\varepsilon^2}) \quad \text{Eq. 87}$$

$$K_{form} = K_L + K_f$$

$$K_L = \left[\frac{\sqrt{k'(1-\sigma)} + (1-\sigma)}{\sigma}\right]^2 \quad \text{Eq. 88}$$

$$K_f = \frac{f_{fric} L P_U}{4 A_U} [\frac{P_R}{P_U}\left(\frac{1}{\sigma}\right)^3 - 1]$$

Six combinations of models are validated with the PCMQBN framework and the same monetary rewards are assigned to each maturity level as the value #1 in Table 5-6 . The EMV of every code adequacy can then be determined based on the probability and the utility of each level. Table 5-10 shows the assessed code adequacy and the corresponding EMV for both models.

Table 5-10: EMVs of the code adequacy for six combinations of wall friction and form loss coefficient model.

Model Combination	Equal	Non-Equal
Blasius and Constant	2.178	2.219
Blasius and Rehme	2.183	2.223
Blasius and Shiralkar	2.176	2.217
C-W and Constant	2.132	2.141
C-W and Rehme	2.114	2.109
C-W and Shiralkar	2.162	2.199

It can be found that the combination of Blasius wall friction and Rehme form loss results in the highest EMV when equal and non-equal weight factors are assigned. Though the Shiralkar and C-W correlation are high-order and sophisticated, the performance is not necessarily better than the low-order model since additional information, including the rod roughness, hydraulic diameter, and type of spacer grid, are required. Therefore, the selection of models highly depends on the availability of information. When information and measurement are lack, high-order model may result in worse credibility.

5.2.3. Findings and Issues

In this section, a technique of PCMQBN is developed to formalize the validation decision-making process with mathematical languages. Considering the subjective nature of validation, its decision-making process is first defined as an argument process. Next, aiming at three components (evidence, structure, and acceptance criterion) of the validation arguments, PCMQBN further represents the argument process with Bayesian Decision theory: the evidence is characterized according to the concept of sufficient accuracy; the argument structure is constructed according to the validation cubic model; and the acceptance criterion is made according to the sufficient-accuracy-guided validation goal. When the scenarios have Level 1 uncertainty, the PCMQBN becomes a Bayesian inference process, where the maturity level of sub-attributes can be inferred. For scenarios with Level 2 or 3 uncertainties, PCMQBN is able to support the adequacy decision according to expert knowledge. Besides, two case studies are prepared for demonstrating the application of PCMQBN in estimating the model credibility. It is found that when the QoI is used for model calibration (posterior information), the experiments that measure the QoI should be more important than others. At the same time, for the same model, it is more preferable to have the model parameters being calibrated by the QoI directly rather than being calibrated separately. Besides, it is found that when only prior information is used for model prediction, the model credibility depends heavily on the amount and coverage of data. If important information cannot be accurately measured, high-order model may result in bad prediction.

In this study, the evidence characterizations, validation structures, acceptance criteria are prepared by author's knowledge. Models, including validation cubic and weight factor, are

tentative, and detailed analysis is still needed. Besides, for Bayesian application, strong assumptions must be made for the precision of probability and utility such that they can be represented precisely with a single additive value [53]. Considering the large uncertainty of the validation process, additional analysis is needed to make the result convincing. One convenient approach is called Robust Bayes [75], where multiple validation assessments are performed, and the class of results are ensembled to inform the final decision of model credibility (Assumption B6). If the class of decisions is approximately the same, it can be claimed that a robust result is obtained. Otherwise, the range can be taken as an expression of confidence from the analysis. Such process is also known as Bayesian sensitivity analysis. Figure 5-16 shows the schematic plot of the Bayesian sensitivity analysis and the entire loop of validation is performed iteratively until a consistent result or an expression of confidence is obtained. Besides, the sensitivity study is a unique feature enabled by the quantified framework. Once an agreement is reached on the validation structures, acceptance criteria, and evidence characterizations, validation of PCMQBN can be performed transparent and independently by all individuals from the expert community that serves as the "defense-in-depth" in assuring the validation quality.

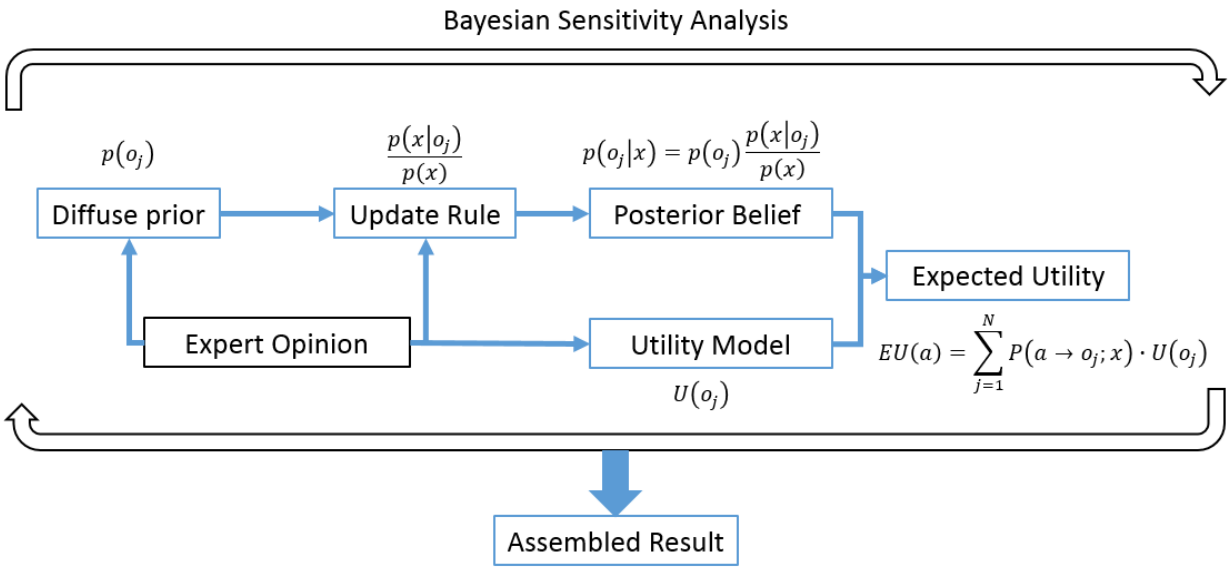


Figure 5-16: Scheme of validation decision and the performance of Bayesian Sensitivity Study.

In addition to the problem of precision, the attribute of coverage is not included. Besides, the effect of Process Quality Assurance (PQA) factors, classified as indirect evidence, is not

considered. Due to the limited scope of this study, EMVs are determined by postulated monetary rewards and used to guide the validation decision-makings. Expected utility is still needed to properly connect to the concept of sufficient accuracy.

5.3. Local Data-Driven Uncertainty Quantification

To quantify the uncertainty of SPH simulations, a local data-driven uncertainty quantification technique is developed by training a surrogate model for predicting the simulation error based on the local simulation information. At the same time, to identify the application range of the trained surrogate model, this section also investigates the predictive capability of the trained surrogate model by extrapolating the simulation uncertainty with respect to different similarity indexes. A case study is performed to demonstrate the capability of the proposed technique. At last, findings and limitations are discussed according to the case study and literatures.

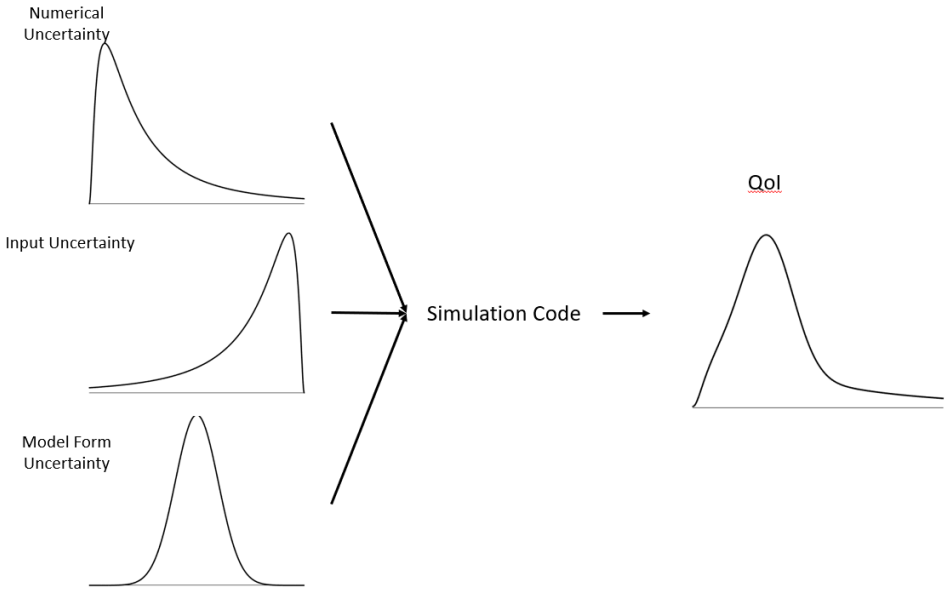
5.3.1. Technical Development

It has been discussed in section 1.3.2 that the regular uncertainty quantification approach relies on the separation of uncertainty with respect to its sources [54]. However, it is found from the SPH studies that all sources of uncertainties are tightly coupled and nonlinearly interacted. And such problem becomes severe when the simulation starts, and the distribution of particles is getting "disordered" [96] [133]. Therefore, the SPH simulation error is very likely to fluctuate and expand as the fluid becomes violent. To reduce the simulation error, the SPH community has been trying from two major aspects: by regulating particle distributions and by introducing additional models. For the former approach, the Fickian diffusion is used to provide shifting towards areas with lower concentration extending to free-surface flows [168]. However, it is found from this study that the particle shifting can be more unstable when the Reynolds number gets higher. For the latter approach, turbulence models including RANS [121] [124] and LES [120] [95] have been investigated for a long time. For LES approach, the performance is found to be very limited, especially in dealing with wall-boundary layers [118]. Although, refining the particles can improve the simulation when it reaches quasi-DNS scales, SPH can be far more computationally expensive than the mesh-based methods. On another hand, the RANS-based approach has proven to be

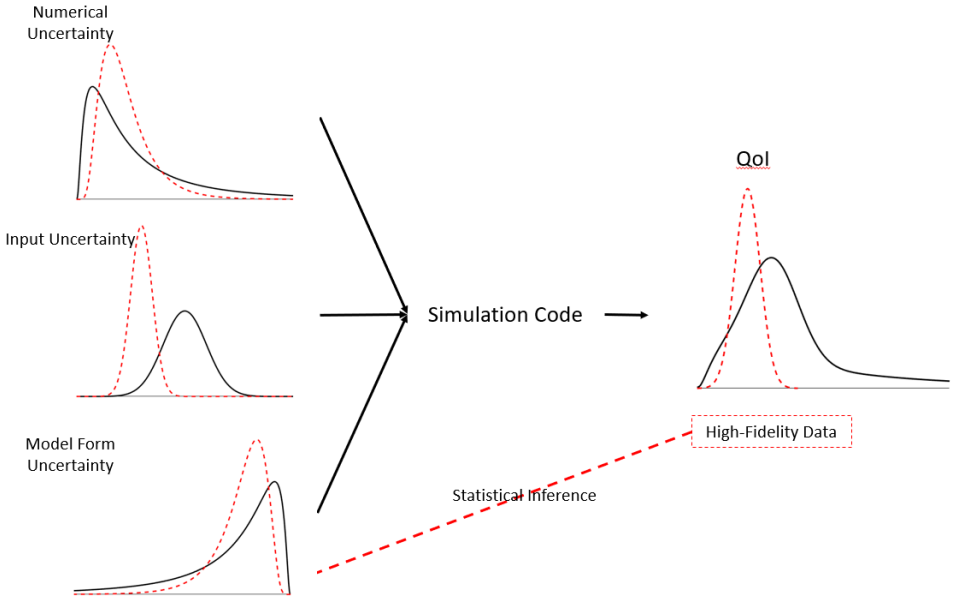
successful with respect to both stability and accuracy when it is combined with incompressible SPH solver. However, the computational cost of such approach is too high for achieving a comparable accuracy to the mesh-based simulations. It is reported from Leroy [96] that ISPH-USAW takes 32h to simulate a lid-driven simulation with 200×200 particles for 60 physical times, while for the same physical times, FV method only takes around half an hour with 200×200 mesh and LES model. In addition, even with sophisticated turbulence models, the disordered particle distribution could potentially induce severe error fluctuations. As a result, it is very difficult to improve the accuracy of SPH simulations by simply incorporating better turbulence models since the nature of particle-based CFD methods deviates largely from mesh-based methods. Therefore, focuses in this study have been put on characterizing uncertainties of SPH simulations for better understanding the SPH methods.

There are two types of flowcharts for quantifying the simulation uncertainty. Figure 5-17 illustrate the process of forward and inverse uncertainty quantification process. The forward approach identifies the uncertainty distributions with respect to all sources of uncertainties separately and propagate the distributions through equations; while the inverse approach assimilates data to infer the uncertainty distributions and propagate the inferred/calibrated distribution for the uncertainty analysis. In some literatures, the forward UQ is known as "data-free" methods [169], while the inverse UQ is mostly known as "data-driven" methods [170]. Though the calibrated model usually provides better fitting and precision against experimental measurements, its applicability and accuracy to experiments with different conditions and timescales remains to be a concern. Besides, there are two types of mathematical forms on representing the uncertainties, which includes parametric and non-parametric forms. For parametric approach, a fixed model form is assigned for representing uncertainties, and model parameters are assumed to be the sources of uncertainty for predicting the QoI. The nonparametric approach does not assume any constant model form, and it is able to investigate the uncertainty of any modeled term, including QoI or simulation field data. Clearly, the nonparametric approach reveals more insights than the one from model coefficients. And since the nonparametric approach makes less assumptions, it is claimed to be more reliable and robust than the parametric approach. However, the nonparametric approach theoretically has infinite degrees of freedom, and it is not achievable for practical purposes. Preventions against overfitting and underfitting requires a very

large amount of available, relevant, and adequately evaluated data (ARAED). Since it is difficult to characterize uncertainty with pre-specified uncertainty distributions for each source, this study aims to quantify the simulation uncertainty with data-driven techniques. Furthermore, considering the complex interactions of various uncertainty sources, robust nonparametric forms are used for characterizing the simulation uncertainties.



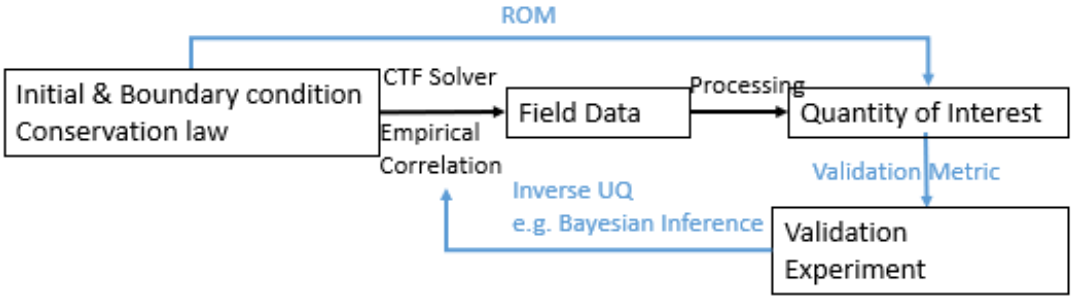
(a) Forward uncertainty quantification



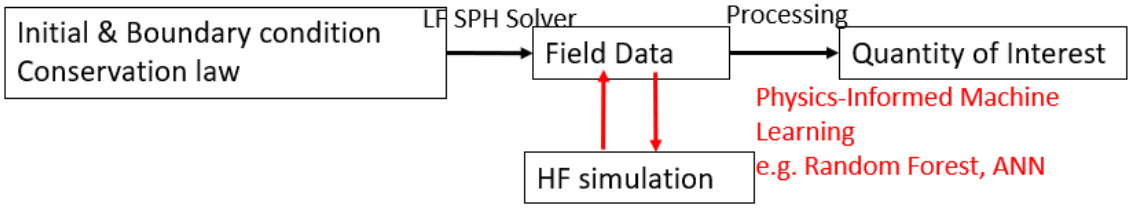
(b) Inverse uncertainty quantification (data-driven)

Figure 5-17: Illustration of forward and inverse uncertainty quantification.

Depending on the assumptions, the data-driven UQ can be further divided into two groups: the global data-driven UQ and the local data-driven UQ. Figure 5-18 shows their schematic workflow. For the global approach, closure models are assumed to be the major sources of uncertainty. At the same time, the simulation is assumed to be well-verified such that the numerical uncertainties can be neglected. As a result, the global approach mainly investigates the uncertainties of model form and model parameter. For the local approach, no explicit assumption is required, and it is able to characterize all sources of uncertainties at the same time. Besides, different from the global data-driven UQ, the objective of the local approach is to estimate the simulation uncertainty of field data. A detailed description and corresponding examples can be found in the reference document [171].



(a) Workflow of global data-driven UQ approach



(b) Workflow of local data-driven UQ approach

Figure 5-18: Illustration of workflow for global and local data-driven UQ approach.

Although the UQ aims to propagate/predict the simulation uncertainty, it is more straightforward and feasible to compare the corrected simulation field against the high-fidelity results. In this study, the quantified uncertainty will be substituted back to the simulation and the corrected result will be compared against the real value to test the applicability of proposed

methods, and the quality of uncertainty prediction is quantified by L_2 relative error norm (defined in Eq. 53). In fact, when Eq. 89 and Eq. 53 are used together, the L_2 relative error norm for the uncertainty prediction is the same as the norm for corrected velocity field. Furthermore, the method of Physics Evaluated Machine Learning (PEML) is used for determining the simulation uncertainties based on local physical features. PEML has first been proposed for characterizing the simulation error of each computing element that is induced by turbulence models [172] [173]. The objective is to construct a surrogate model between the simulation error and the physical features. Except for the PEML, there are four other types of machine learning frameworks, including Physics-Separated ML, Physics-Integrated ML, Physics-Recovered ML, and Physics-Discovered ML. Detailed descriptions for each types of machine learning framework can be found in Chang's work [174].

$$\varepsilon_{d,V_i}(\vec{\boldsymbol{v}}_i) = \vec{\boldsymbol{v}}_i(HF) - \vec{\boldsymbol{v}}_i(LF) \qquad \text{Eq. 89}$$

Since the local approach targets at coarse-grid SPH simulations, one challenge is that the high-fidelity data is obtained from mesh-based simulation. A mapping method is needed to transfer the field data (pressure, velocity, energy) from Eulerian method to Lagrangian method. In this study, a simple averaging method is used by assuming that each particle has a stable size the properties in the particle volumes are approximately the same. In general, the PEML approach has two steps: training and testing:

Training Step:

(1) Design and conduct high-fidelity CFD and SPH coarse-grid simulation according to the PIRT
(2) Map the local data from high-fidelity CFD to SPH element and calculate the simulation error of SPH coarse-grid simulations
(3) Identify and calculate the local physical features based on SPH coarse-grid simulation data
(4) Train a regression between the simulation error and the local physical features using machine learning algorithm

Testing Step:

(5) Characterize the similarity of testing and training physical features

(6) Predict the simulation error for testing cases using trained regression and test its capability

5.3.2. Case Study

To demonstrate the application of the local approach, an illustrative example is prepared, and the objective is to demonstrate how PEML predicting the uncertainty of SPH in simulating the velocity field. A lid-driven simulation is constructed with LAMMPS-SPH and OpenFOAM, and OpenFOAM simulation results are treated as the high-fidelity data. The simulation configurations for SPH and OpenFOAM are the same as those in section 4.3.2.2, and the Reynolds number Re_τ is the only parameter being changed (Table 5-11). All SPH simulations are run on one cluster node with 64 processers, and the computational time is also listed in Table 5-11. In addition, a constant time step of $\Delta t = 10^{-5}s$ is used for all scenarios, and all simulations run for 100s. Currently, only the dynamic viscosity is changed, therefore, the Reynolds number Re_τ is used to characterize each case. To test the applicability of PEML, it will be trained with selected cases, named as training sets, and applied to other cases, named as application sets.

Table 5-11: Summary of simulation ID and configurations. Dynamic viscosity is currently the only variables being changed.

ID	$\nu(m^2/s)$	Re_τ	Computational Time (sec)
dp1	10^{-3}	1000	4872
dp2	5×10^{-4}	2000	4470
dp3	2×10^{-4}	5000	5056
dp4	10^{-4}	10000	5132

Before determining the uncertainties of SPH simulation, the high-fidelity mesh-based simulation data need to be mapped to the SPH simulation. In SPH, the particle behaves like material points that have mass and volume. Therefore, the particle can be treated as macroscopic

entities that bear kinematic and thermodynamic quantifies, and mov over time. For each particle, mesh grids covered by the particle volume are first identified. Next, unweighted averages are taken for all fluid properties carried by the identified mesh grids, and such averages will be treated as the high-fidelity SPH results from mesh-based simulation. Also, for meshes that are not fully covered by the particles, their properties will still be considered. In Figure 5-19, all shaded mesh grids are treated as covered by the particles. During this averaging process, it is assumed that each particle has a stable size and the volumes are close to the mean value.

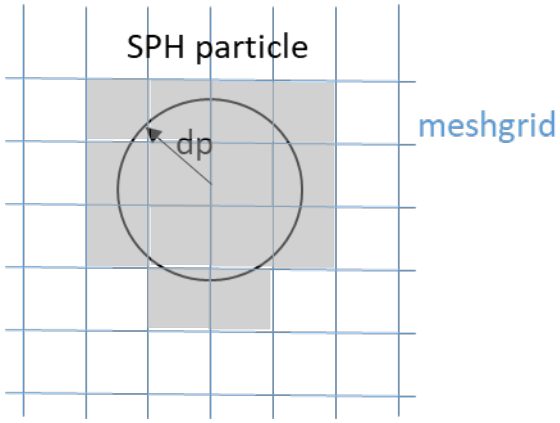


Figure 5-19: Identification of mesh grids that are covered by particle volumes.

Besides, since the particles are moving with time, a convergence study has to be performed to find the approximate time when the flow is fully developed. Figure 5-20 shows the time averaged L2 relative error norm by four time-windows (TW) versus the physical time of simulation. It is found that the averaged results by different time windows are approximately the same after $70s$. Since this investigated case has the highest Re_τ, other cases with lower Re_τ can also be treated as fully developed after $70s$. In this study, the instantaneous field at $t_{physical} = 100s$ is used for training and testing the PEML.

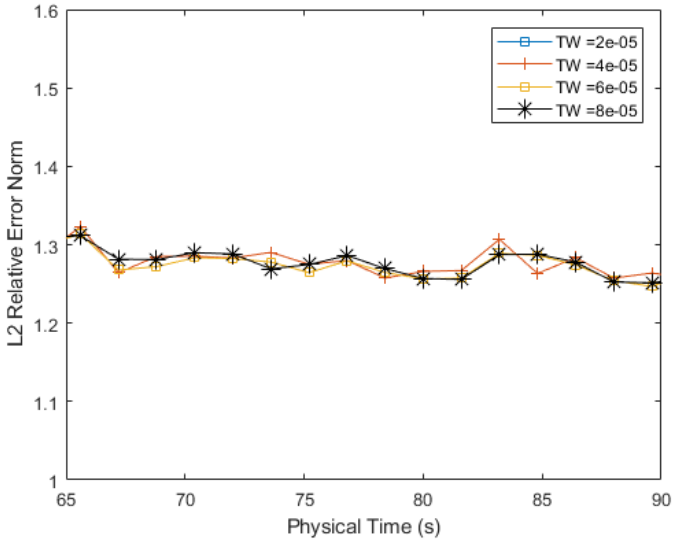


Figure 5-20: Time averaged L2 relative error norm by four time-windows versus physical time for at fully developed state with $Re_\tau = 10000$.

Figure 5-21 shows the comparison of centerline velocity from the original SPH, mapped SPH and the mesh-based simulation by OPENFOAM. After the mapping, the SPH uncertainties in predicting field data are calculated by Eq. 89.

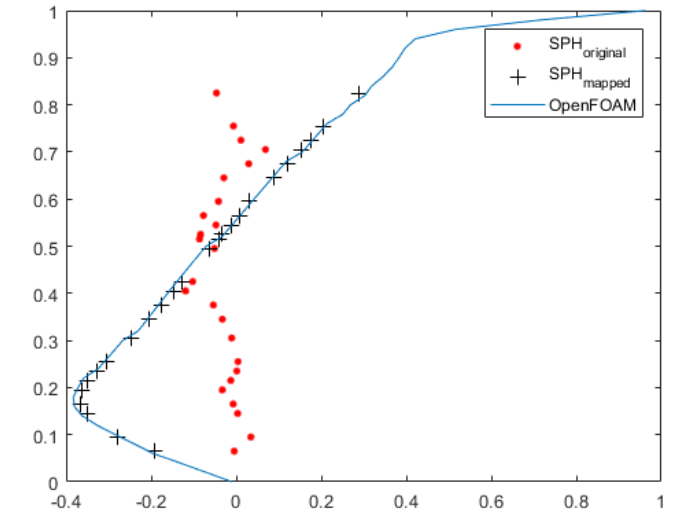


Figure 5-21: Comparison of V_x at the vertical centerline.

Next, based on the SPH equation for velocity calculation and the theoretical analysis for the truncation and discretization error, five physical features are selected as the input variables (Eq. 90) of PEML. For a two-dimensional simulation, each physical feature is propagated to two direction and there will be ten features in total.

$$\left(\vec{F}_p\right)_i; \left(\vec{F}_v\right)_i; \vec{\boldsymbol{v}}(\boldsymbol{r}_i)\sum_j w_h(\boldsymbol{r}_i-\boldsymbol{r}_j)V_j\,; \nabla\vec{\boldsymbol{v}}(\boldsymbol{r}_i)\sum_b w_h(\boldsymbol{r}_i-\boldsymbol{r}_j)\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)V_j\,; \boldsymbol{r}_{\boldsymbol{i}_{wall}} \quad \text{Eq. 90}$$

Next, both input features and target variables are fed into the machine-learning engine and the simulation error of velocity is predicted for each particle. To characterize the accuracy of PEML in predicting simulation uncertainties for SPH application, L_2 relative error norm is used for the predicted field error and the corrected field. In this case study, these two error norms have the same value. Table 5-12 shows a summary of all cases, including their training datasets, application datasets, error of corrected velocity field based on the PEML result, and error of original SPH velocity field.

Table 5-12: Summary of PEML results with Random Forest as the machine-learning engine.

	Re_τ (Training Group)	Re_τ (Target Group)	L_2 (Corrected Field)	L_2 (Original Field)
A1	1000	5000	[1.27,2.64]	[1.04,1.04]
A2	2000		[0.85,0.81]	
A3	10000		[1.09,0.84]	
A4	1000+10000		[0.78,0.81]	
A5	1000+2000		[0.77,0.75]	
A6	2000+10000		[0.83,0.75]	
A7	1000+2000+10000		[0.78,0.76]	
B1	1000	10000	[1.46,1.10]	[1.29,1.25]
B2	2000		[1.17.0.92]	
B3	5000		[0.96,0.81]	
B4	1000+2000		[1.07,0.95]	
B5	1000+5000		[1.00,0.89]	
B6	2000+5000		[0.90,0.82]	
B7	1000+2000+5000		[0.96,0.90]	

For both case studies, it can be found that the PEML trained by the first training set has the worst prediction capability since the training group has the most distant Re_τ to the target group. Also, it is reported from various literatures [175] [176] that when the Re_τ is larger than 1000, the flow field becomes unstable, and both the upstream second eddy and Taylor-Gortler-like vortices becomes unsteady. Therefore, the poor performances of case A1 and B1 are due to the transitions of dominant physics. All other PEMLs successfully reduce the error of field data when the predicted error is added back to the low-fidelity data. For case study #1, it is found that as more datasets are used for training, the error of corrected fields becomes smaller. Furthermore, it is found from Figure 5-23 that the simulation field corrected by interpolated PEML has better accuracy than that by extrapolated PEML. In this study, the interpolated PEML is defined as the case where the target Re_τ falls within the range of Re_τ of training groups (case A4, A6, A7), and the extrapolated PEML is defined as the case where the target Re_τ falls outside the range of Re_τ

of training groups(A1, A2, A3, A5). For case study #2, it is found from Figure 5-22 that as the Re of training set becomes closer to the target set, the error of corrected field gets smaller.

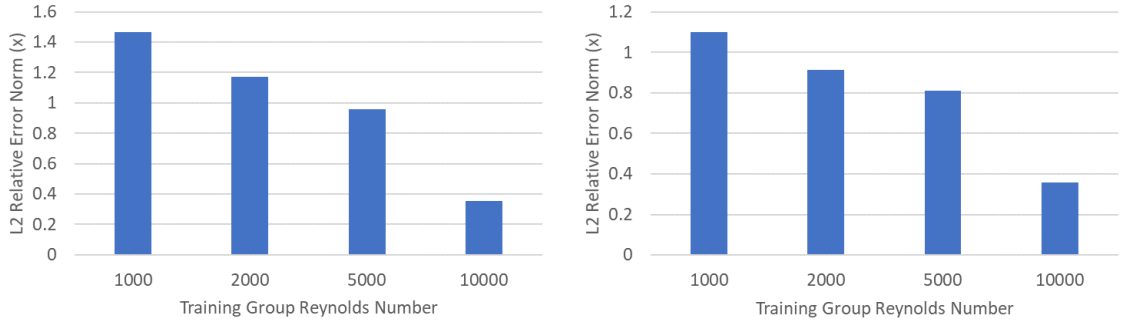


Figure 5-22: Histogram of PEML predicted uncertainty when the training group contains a single simulation and the target group has $Re_{\tau-target} = 10000$.

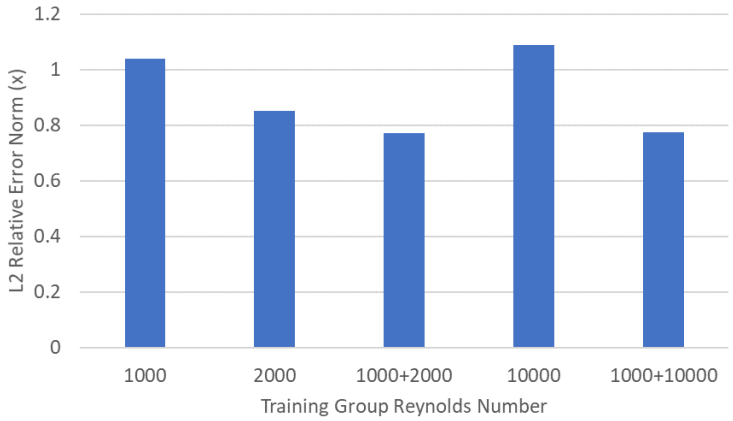


Figure 5-23: Histogram of PEML predicted uncertainty when the target group has $Re_{\tau-target} = 5000$.

In addition to the random forest [177] machine-learning engine, this study also tests the feedforward neural network engine [178] [179]. It turns out that random forest is faster and more robust than the neural network, however, the error reduction with respect to the data amount is more obvious for the neural network.

5.3.3. Findings and Issues

In this section, a local data-driven UQ method is developed based on PEML, and the objective is to quantify the uncertainty of SPH simulations. Random Forest is used as the machine-learning engine for constructing a surrogate between local physical features and simulation uncertainties. Five lid-driven cavity simulations with different Re_τ are set, the machine learning algorithm is trained and tested with various combinations of cases. The performance of PEML has been qualitatively analyzed: it is found that other than case A1 and B1, PEML is able to reduce the error of field data by adding the predicted back to the low-fidelity simulations. Besides, it is found that as more datasets are used for training, the performance of PEML becomes better, and the performance of interpolated PEML is better than the one of extrapolated PEML[3]. Moreover, this study also tests the capability of feedforward neural network as the PEML engine. It turns out that random forest is faster and more robust than the neural network, however, the error reduction with respect to the data amount is more obvious for the neural network.

It is found that the performances of case A1 and B1 are not as good as others, and PEML is not likely to characterize the uncertainties due to physical transitions. Besides, to guide the uncertainty extrapolation from training to target cases, a mathematical correlation is needed between the characteristics of training and target cases. However, this study has not found a strong correlation between the similarity index, represented by fluid ratio and K-L divergence, and the PEML performance. Further investigations are needed for characterizing the relationship between training and target cases. It is suggested that more metrics and index can be tested for determining the cases' similarity. Besides, considering the subjective nature of similarity judgements, a new concept is needed for extrapolating the simulation uncertainties. Presently, a new concept named "physical coverage" has been proposed. It first reduces the dimensionality of multiple physical features with t-Distributed Stochastic Neighbor Embedding method, and then visualize the physics coverage condition of target case with respect to the training case. Moreover, it is suggested from biomolecular discipline [180] that for high-fidelity measurements with averaged quantities over many elements and long periods of time, the model should be calibrated based on maximum-entropy principle. The objective is to avoid overfitting associated with the calibration of few model

[3] Definitions for interpolated and extrapolated PEML can be found in the glossary table at the end of this document

parameters. However, machine learning algorithms like Random Forest and feedforward neural network are based on maximum-likelihood principle. Therefore, it is necessary to test more maximum-entropy based algorithms, like Bayesian-maximum entropy integrative methods, etc.

5.4. RISMC Model Validation Framework

Since the objective of this study is to assess the credibility of NEUTRINO and to improve the EMDAP framework. New methodologies discussed in this chapter are incorporated into the EMDAP framework for (1) Identifying credibility requirements in validation process; (2) formalizing the decision-making process in the validation process (3) quantifying the uncertainty of SPH simulations in a transparent, consistent, and robust manner. Figure 5-24 shows a schematic flowchart for the improved EMDAP framework. The solid blocks (Elements #1 - #4) are inherited from the EMDAP framework, meanwhile, new components in bold fonts, including sufficient accuracy, data-driven validation, and PCMQBN are developed and included. The dashed blocks are defined with respect to their scopes and requirements.

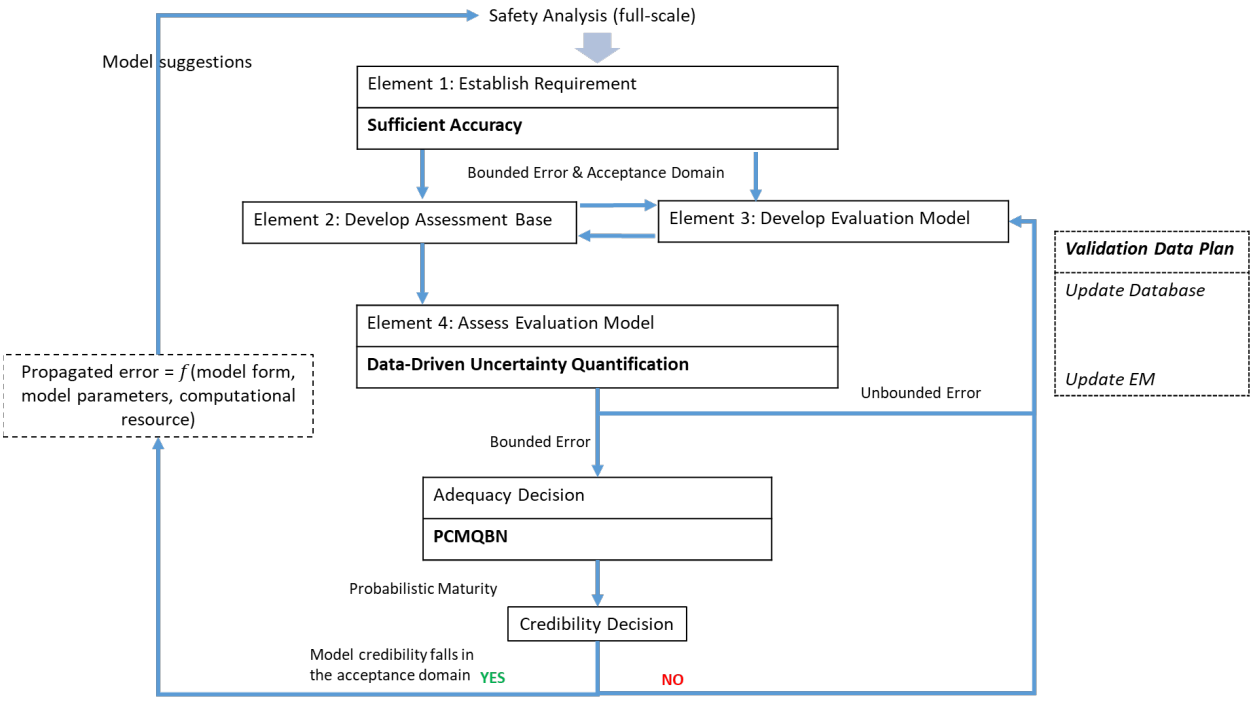


Figure 5-24: A schematic flowchart of the improved EMDAP framework. The solid blocks are mainly inherited from the well-established EMDAP framework, and three of them have been improved with new methodologies (in bold font). The definitions and case studies of bold methodologies have been demonstrated in this study. The scopes and requirements of dash blocks have been defined for the dashed blocks.

5.4.1. Blocks Improved by Formalized Methodologies

The concept of sufficient accuracy is applied to Element 1. The validation goal, represented by maturity level and corresponding subjective beliefs, is determined according to the safety goal, where a critical belief and a range of acceptance domain for model credibility are identified. The transparency and consistency of this process is ensured by formalized procedures and mathematical derivations. At the same time, by admitting the subjective nature of validation and introducing qualitative validation goal, the robustness of this process is ensured by achieving a designated level of "credibility" across the validations and assumptions that are consistent with known facts. The PCMQBN is applied to the element for adequacy decision. The validation decision-making process is first defined as an argument process. Next, aiming at three components (evidence, structure, and acceptance criterion) of the validation arguments, PCMQBN further

represents the argument process with Bayesian Decision theory: the evidence is characterized according to the concept of sufficient accuracy; the argument structure is constructed according to the validation cubic model; and the acceptance criterion is made according to the sufficient-accuracy-guided validation goal. Again, the transparency and consistency are ensured with formalized procedures and mathematical representations. The robustness is ensured by a formalized Bayesian sensitivity analysis, also known as Robust Bayes [75]. The Data-Driven Uncertainty Quantification is applied to Element 4. The uncertainty of SPH simulations are quantified with techniques of PEML, where a surrogate model is constructed between the local physical features and simulation uncertainties. Presently, Random Forest is used as the major algorithm for constructing the surrogate model. Meanwhile, the capabilities of constructed PEML are tested for both the interpolated and extrapolated scenarios[4]. The transparency is ensured by the opensource simulation engines and the open-access validation data management system. Meanwhile, the consistency is ensured by treating the surrogate model as a fast data extraction from high-fidelity database. The robustness is mainly ensured by the reliability and robustness of modern machine learning algorithms.

5.4.2. Scoping Developments for Validation Data Plan and Uncertainty Scaling

Validation Data Plan (VDP) is "a dynamic planning instrument to guide, and potentially optimize activities on data production and acquisition, data analysis and management, and data usage so that they enable effective support for development, assessment and application of simulation tools intended for challenge problem" [151]. In the flowchart of Figure 5-24, VDP refers to a decision model that (1) integrates information from all related validation activities, and (2) prioritizes data activities, based on cost-benefit analysis of possible activities. In this study, the cost-benefit analysis is performed based on the value of information theory[5], and a decision tree is used for calculating the expected value of sample information (EVSI). The EVSI refers to the expected increase in utility or monetary value that a decision-maker could obtain from gaining access to a sample of additional observations before making a decision [181]. A synthetic example is prepared for demonstrating the procedure of calculating EVSI according to the adequacy

[4] Definition of interpolation and extrapolation can be found in the glossary table at the end of this document
[5] The term "value of information" is defined in the glossary table at the end of this document

maturity from PCMQBN and a "postulated" new model. For the vent overflow model that has been discussed in section 5.2.2.1, it is assumed that a postulated model is designed such that the maturity level of validation result is improved as in Figure 5-25.

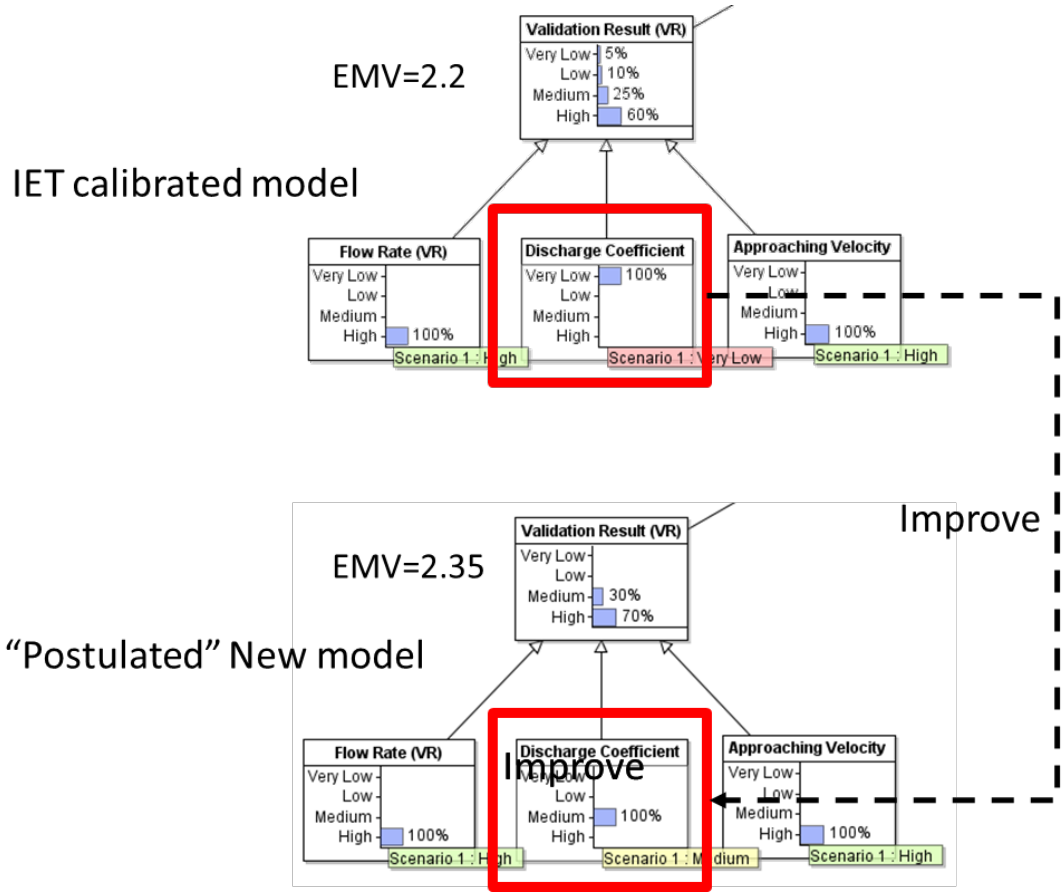


Figure 5-25: Demonstration of maturity and code adequacy improvement by a "postulated" model.

It is further assumed that the chance of getting such an improvement is 40%, and there is 60% chance that a model as bad as the SET-Calibrate model is obtained. A decision tree, shown in Figure 5-26, can be built for determining the expected increase in monetary value by having an improved model rather than having the TDMI-calibrated model. In this scenario, the expected value is found to be negative, and it indicates that the efforts devoted to the development of such new model are more than those being already obtained. Therefore, the VDP for such a model is not preferred.

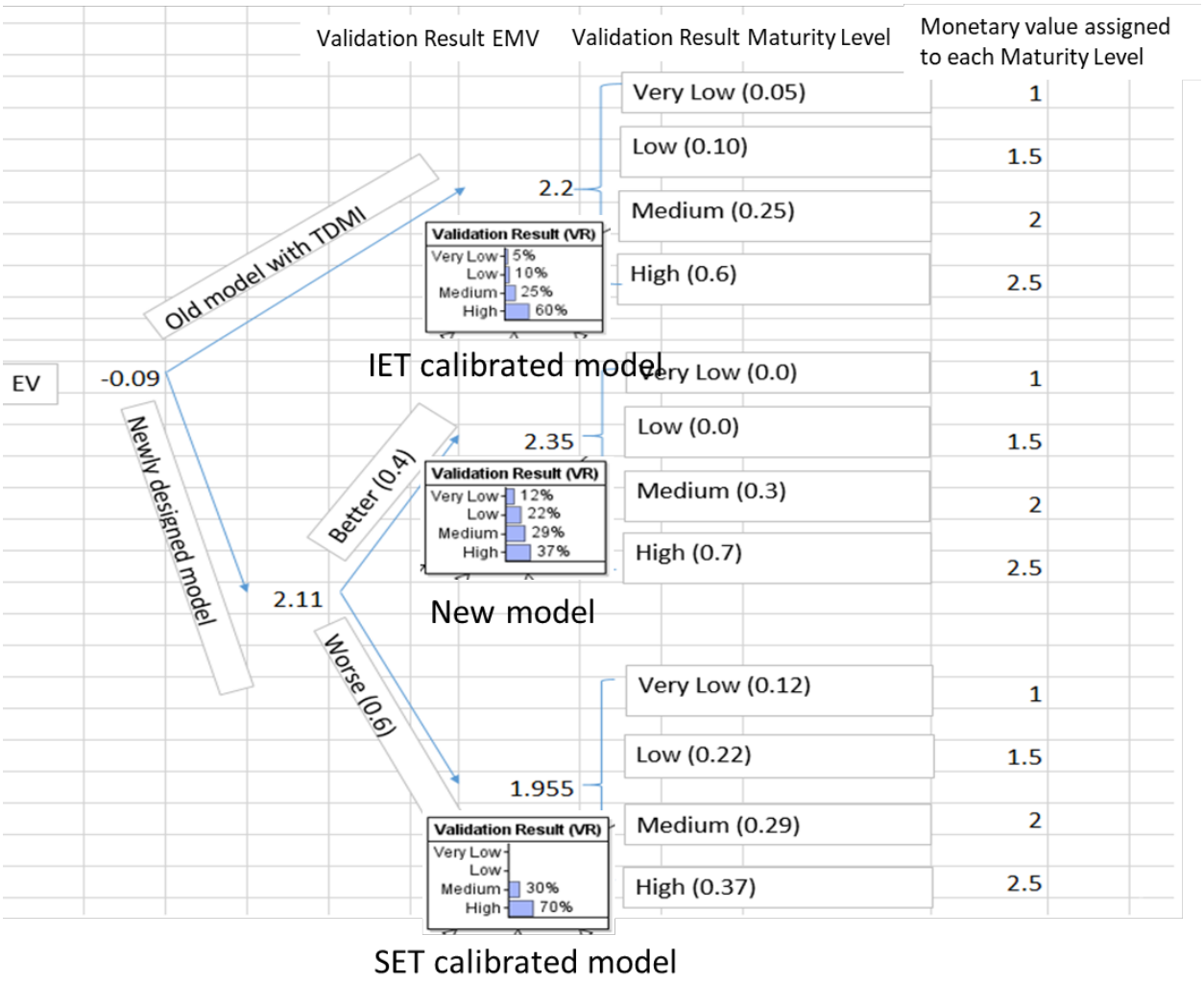


Figure 5-26: Decision tree for calculating the expected increase in the monetary values by having a postulated new model rather than having the IET-calibrated model.

Another block being developed is the error propagation, where the simulation uncertainty will be propagated from validation to application scenarios. Since the biggest challenge of such propagation comes from the scaling analysis, such error propagation is also named as uncertainty scaling. The objective is to (1) characterize the applicability of reduce-scale validation to full-scale application. (2) construct a mathematical relationship between the simulation error and the characterized applicability. In previous section, a qualitative relationship has been constructed between the PEML corrected simulation error and the dimensionless number. An initial study has been performed for constructing a mathematical relationship between the simulation uncertainty and the similarity index. Presently, for the case study of lid-driven cavity, similarity indexes refer

to the Reynolds ratio and symmetric Kullback-Leibler Divergence, and their goal is to identify the similarity of training and testing scenarios. The Reynold ratio is estimated by the ratio of Re_τ (Eq. 63), while the symmetric Kullback-Leibler Divergence is estimated by probability distributions of local physical features. In this study, the Reynold ratio $r^*$is defined in Eq. 91, while the symmetric Kullback-Leibler Divergence D_{KL} is defined in Eq. 93.

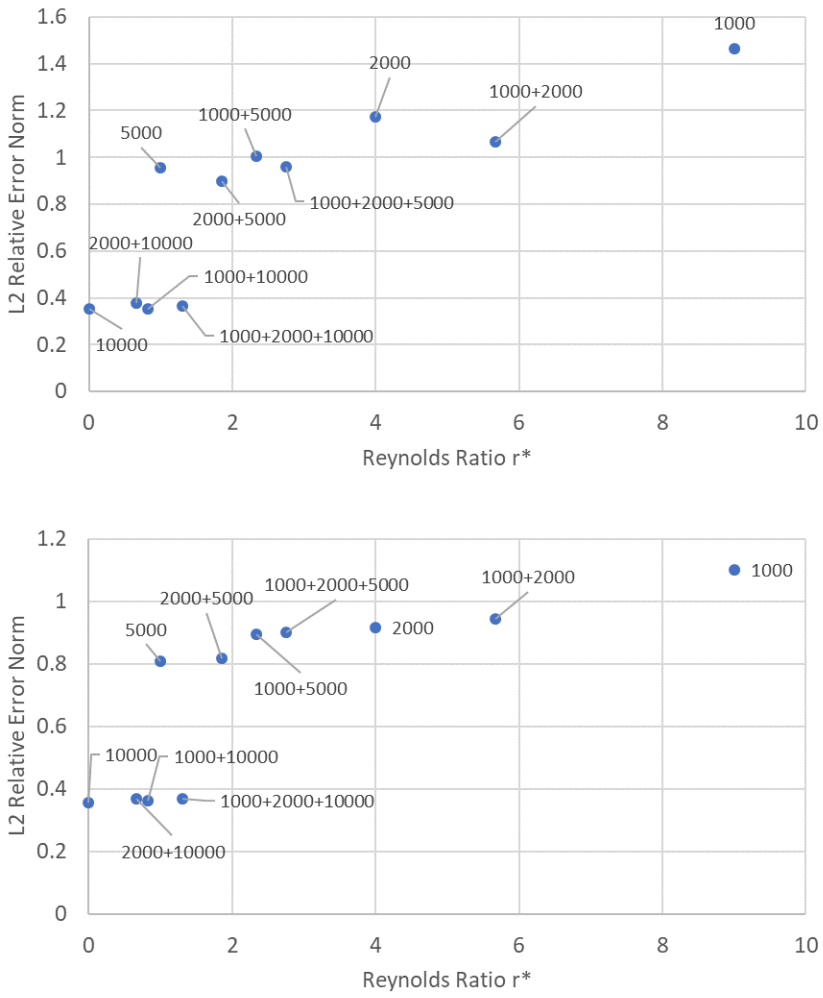
$$r^* = Re_{target}/\overline{Re}_{training} \quad \text{Eq. 91}$$

$$\overline{Re}_{training} = \sum_{i}^{n} Re_i \,/n \quad \text{Eq. 92}$$

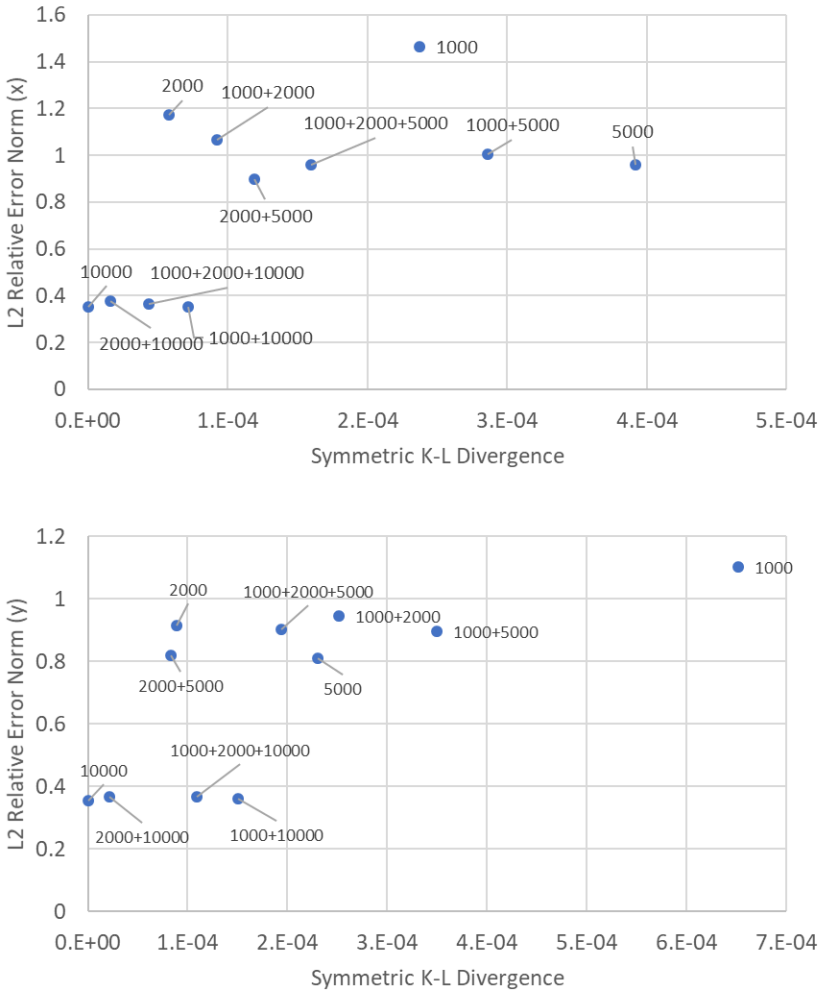
$$D_{KL}(P,Q) = \sum_{i} P(i)\log\left(\frac{P(i)}{Q(i)}\right) + Q(i)\log\left(\frac{Q(i)}{P(i)}\right) \quad \text{Eq. 93}$$

$$P(i) = \frac{1}{Nh_1h_2\dots h_d}\sum_{i=1}^{N}\prod_{j=1}^{d} k(\frac{x_j - y_{ij}}{h_j}) \quad \text{Eq. 94}$$

where Re_{target} is the Reynolds number of the target scenario; $\overline{Re}_{training}$ is the average Reynolds number of the training scenarios calculated by Eq. 92; $P(i)$ and $Q(i)$ are the probability density distribution of physical features y approximated by multivariate kernel distributions; $k(\cdot)$ is the one-dimensional kernel smoothing function; x is the d-dimensional random vector for the physical features. N is the number of samples drawn from each group of physical features. Figure 5-27 shows the plot of L2 relative error norm for the corrected velocity field against the Reynolds ratio r^* and symmetric K-L Divergence D_{KL} when the target scenario has $Re_{target} = 10000$. However, no strong mathematical correlation has been found between the suggested representation of similarity and the PEML performance.



(a) L2 relative error norm versus Reynold ratio r^* for x (upper) and y (lower) directions' velocity field



(b) L2 relative error norm versus symmetric K-L Divergence of all physical features for x (upper) and y (lower) directions' velocity field

Figure 5-27: Plot of NMSE for the corrected low-fidelity simulation against the Reynolds ratio r^* and K-L Divergence when the target scenario has $Re_{target} = 10000$.

5.4.3. Extended Framework Development

The framework development of this study mainly contributes to the Risk-Informed Evaluation Model Development and Assessment Process (REMDAP) framework for the Integrated Research Project (IRP) of "Development and Application of a Data-Driven Methodology for Validation of Risk-Informed Safety Margin Characterization Models" [182]. The objective of IRP is to "provide a mathematically defendable basis for calculating biases and their

uncertainties for a wide range of operating conditions that represent the intended range of model application" by combining the CSAU/EMDAP methodology with advanced DDM methods, including Reduce Order Modeling (ROM) [183], Simulation-Based Scaling (SBS) [139], Validation Data Planning, and Validation Data Management System (VDMS) [184]. Figure 5-28 shows the advanced methods and tools to facilitate the implementation of already demanding EMDAP in a risk-informed application. Orange boxes represent methodologies that will be developed in this project and their task number.

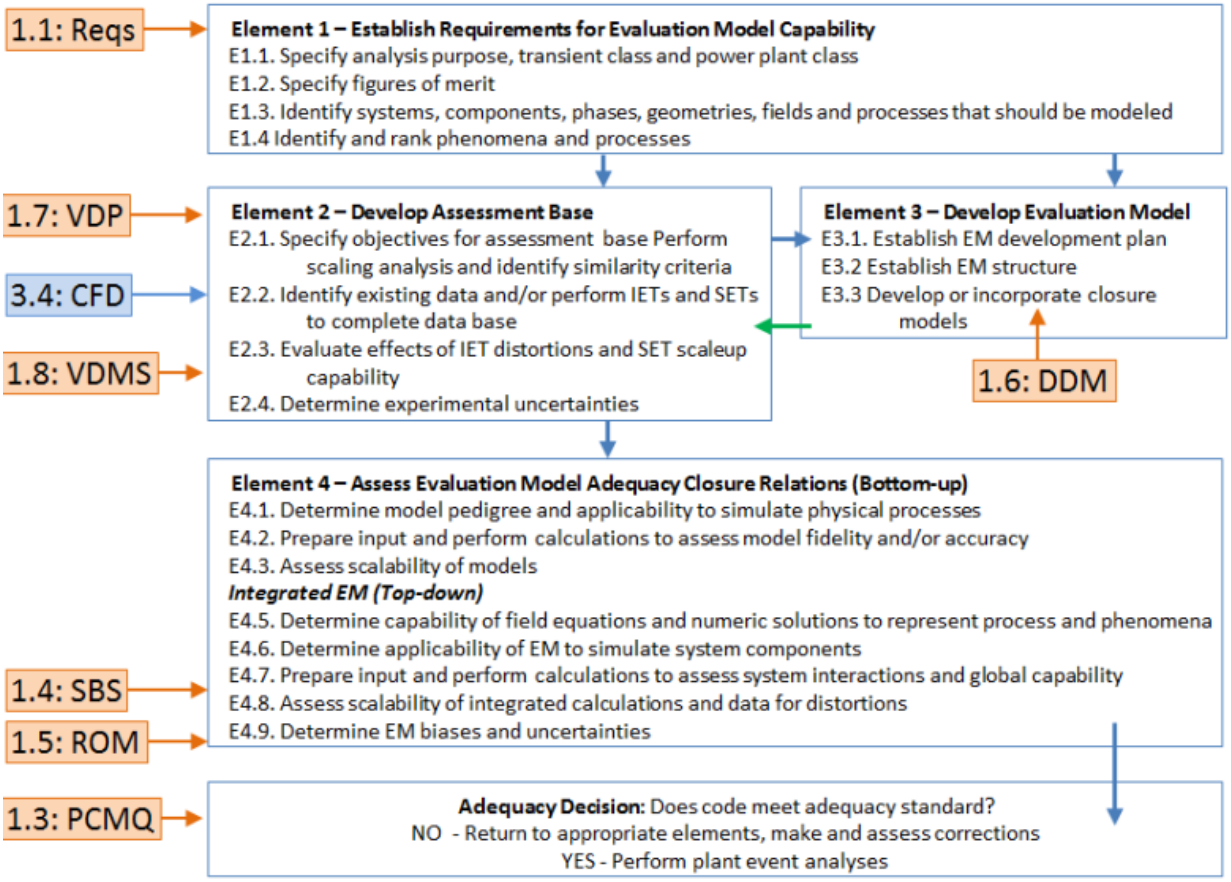


Figure 5-28: Risk-informed Evaluation Model Development and Assessment Process. Orange boxes denote developments in the project and indicate EMDAP elements where the advances are expected to impact [182]. "Reqs" is the abbreviation for "Requirements"; VDMS refers to Validation Data Management System; SBS refers to Simulation-Based Scaling; DDM refers to Data-Driven Modeling; ROM refers to Reduced Order Modeling.

More details of the IRP progress can be found in the first-year annual report [185] and milestone reports [186]. Moreover, to guide and test the methodology development in this study, experimental facilities for flooding and system thermal-hydraulics will be designed.

5.5. Issues and Challenges

For the concept of sufficient accuracy, the application is limited to the validation processes with large uncertainties (Level 2 and 3), but not deep uncertainties. Meanwhile, the concept of sufficient accuracy is more applicable when the validation is required to be both convincing and efficient. This study investigates the relationship of subjective beliefs and assumes that the maturity levels are objective and prescribed. However, in many scenarios, the designation of model maturity levels can be subjective, and its effects have not been looked into. Besides, this study uses Bayesian Inference to propagate the error range from prototypic systems to sub-scale models, which only applies to scenarios with Level 1 uncertainty. Also, the inverse UQ techniques depends heavily on data coverage and data quality, their effects on the propagation of sufficient accuracy have not been included.

For the concept of PCMQBN, the precision is the most important concern, and the Bayesian sensitivity analysis is proposed as a convenient resolution. It is found from the sensitivity analysis of Uncertainty [U] attribute that the PCMQBN estimations for the model credibility become closer to the real results as more uncertainty information is used. Further investigations will be included in the future work, where sensitivity study needs to be performed for all attributes, structures, and utility functions. In addition to the problem of precision, the attribute of coverage is not included. Besides, the effect of Process Quality Assurance (PQA) factors, classified as indirect evidence, is not considered. Due to the limited scope of this study, EMVs are determined by postulated monetary rewards and used to guide the validation decision-makings. Expected utility is still needed to properly connect to the concept of sufficient accuracy.

For the technique of PEML-based local data-driven UQ, this study has not found a quantitative relationship between the similarity index and the simulation uncertainties predicted by PEML. It is suggested that more metrics and index should be tested for determining the

similarity index. Besides, considering the subjective nature and practical issues of similarity judgements, a new concept is needed for identifying the physical characteristics between scenarios. In addition, it is found from Figure 4-22 that the interactions between SPH uncertainties and scenarios' physical characteristics are highly nonlinear. Also, the difficulties in determining numerical error makes SPH's UQ more complicated than implicit LES. It is suggested that the error analysis of SPH should start from the small and fundamental scales, like the Lagrangian operators and smoothing functions, etc.

5.6. Summary Remarks

This chapter first describes three new technologies for improving the regular validation framework EMDAP. The concept of sufficient accuracy is developed for guiding the designation of model credibility requirements according to the safety analysis for prototypic systems. The sufficient accuracy admits the qualitative nature of model validation, and it combines the probabilistic and deterministic analysis by representing the validation goal with subjective belief and maturity levels. Besides, a technique of PCMQBN is developed to formalize the validation decision-making process with mathematical languages. PCMQBN is able to support the adequacy decision according to expert knowledge for scenarios with large uncertainties (Level 2 and 3). Moreover, a local data-driven UQ method is developed based on PEML, and the objective is to quantify the uncertainty of SPH simulations. Random Forest is used as the machine-learning engine for constructing a surrogate between local physical features and simulation uncertainties. Case studies are performed for each new concept, where findings and limitations are discussed. Finally, new technologies are incorporated into the regular EMDAP, and a new REMDAP framework is suggested.

Figure 5-29 schematically illustrates the validation pyramid by the concept of sufficient accuracy. The propagation of validation goal is suggested according to the technique of bounded error, while the integration of model adequacy is suggested via the PCMQBN. At the same time, the local data-driven UQ technique aims to help the credibility assessment of sub-scale or integral model. However, more investigations are still needed, which include the verification of Machine Learning algorithms, the integration of sub-scale to full-scale errors, etc.

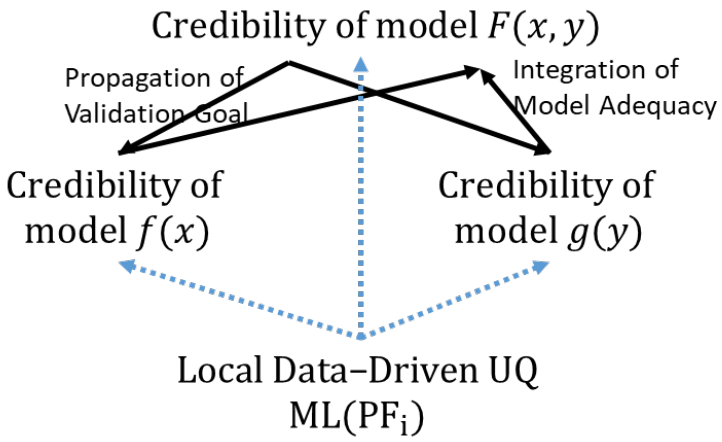


Figure 5-29: Schematic illustration of validation pyramid guided by the concept of sufficient accuracy. The propagation of validation goal is suggested to be achieved by the bounded error technique, while the integration of model adequacy is suggested to be achieved by PCMQBN. At the same time, the local data-driven UQ technique could help the credibility assessment of sub-scale or integral model.

6. CONCLUSION: SUMMARY OF FINDINGS AND FUTURE WORK

This chapter summarizes the findings and contributions. At the same time, recommendations are discussed for future works.

6.1. Summary

This study first assesses the credibility of Smoothed Particle Hydrodynamic (SPH) methods for the external-flooding and high-wind scenario (condition A2) according to the methodology of Code Scaling Applicability and Uncertainty (CSAU) and its regulatory guide Evaluation Model Development and Assessment Process (EMDAP). According to the author's knowledge, high-rank phenomena and numerical benchmarks are identified for both external-flooding and high-wind scenarios by the Phenomenon Identification and Ranking Table (PIRT). At the same time, performance measurement standards are established to decide if the code is adequate for the specific application. Next, by assuming that SPH simulation codes (assumption A1) are faithful representations of their underlying mathematical models, NEUTRINO and LAMMPS-SPH are used for predicting the quantity of interest of selected numerical benchmarks. However, due to the lack of data from integral effect tests, this study only estimates the performance of SPH methods for separate phenomena and processes (assumption A6), and the level of fidelity for supporting evidence is coarse.

It is found that for the designated external-flooding scenario, SPH methods can predict the hydrodynamic force acting on both stationary and moving structures with acceptable performances. For the phenomena with stationary structures, it is suggested that the particle size of NEUTRINO simulation should be 0.02m. At the same time, according to the scaling analysis by dimensionless number x^*, the database is sufficient for predicting a similar phenomenon in full-scale scenarios. Such an argument also applies to the phenomenon of wave propagation, and the suggested particle size is 0.02m. For the phenomenon with moving structures, the suggested particle size for a postulated scenario is 0.25m for the falling cube and 0.1m for the floating cube. Since there is only one dataset for this phenomenon, distortion analysis is not applicable (N/A), and the database is not sufficient. Moreover, performance measurements cannot be made regarding the SPH's integral

performance for the designated flooding scenario. Finally, the adequacy decision should be "inadequate" with respect to the available database and selected SPH implementations.

It is also found that for high-wind scenarios, the performance of SPH methods in simulating turbulence and vortex shedding phenomena is insufficient. Moreover, the issue of particle vacancy problems is severe when the Reynolds number grows. Though the model of particle shifting algorithm can fill the vacancy without increasing the computational costs, the error fluctuations are severe, and the simulations become less stable. Besides, the database is not sufficient for both phenomena, and the uncertainty in the application scenarios cannot be bounded. Moreover, performance measurements cannot be made regarding the SPH's integral performance for the designated high-wind scenario. Finally, the adequacy decision should be "inadequate" with respect to the available database and selected SPH implementations.

Table 6-1 summarizes the performance of SPH methods in simulating external-flooding and high-wind scenarios. The judgments on the simulation performance are made according to the performance standards in section 4.1.2. At the same time, the results of scaling analysis for the validation database and the accuracy assessments for SPH simulations are also included. The judgment on each scaling attribute is made based on the attribute definitions and database properties in section 4.4. Due to the limitation of database, the distortion analysis is not applicable (N/A). It is suggested that additional models, including the turbulence model, wall function, the pressure-velocity coupling model, and adaptive particle size, etc., should be developed. Also, more data from the large-scale flooding scenarios, wind tunnel, debris transport, and so on, should be gathered. Moreover, this study is built based on the author's best knowledge, diverse expert opinions are still needed. Besides, it is suggested that additional models, including the turbulence model, wall function, and adaptive particle size, etc., should be developed. At the same time, more data from the large-scale flooding scenarios, wind tunnel, debris transport etc. should be gathered.

Table 6-1: Performance summary of SPH simulations, where the performance standards and phenomena identifications are discussed in section 4.1. Results from accuracy assessment and database scaling analysis are also included. The accuracy is measured by L_1 relative error norm according to section 4.3, the qualitative judgments on relevancy, sufficiency, and distortion are made according to section 4.4 scaling analysis.

External-Flooding Scenario							
ID	Phenomenon Description	Performance	Accuracy (L_1 Error)	Validation Database			Suggested Particle Size
				Relevancy	Sufficiency	Distortion	
A	Response Time						
	Wave propagation	Reasonable	12.7%	Yes	Yes	Bounded	0.02m
B	Structural Loads						
	Hydrodynamic force on stationary structures	Excellent	3.6%	Yes	Yes	Bounded	0.02m
	Hydrodynamic force on moving structures	Falling – Insufficient	5.52%	Yes	No	N/A	Falling – 0.25m
		Floating – Insufficient	4.41%	Yes	No	N/A	Floating – 0.1m
High-Wind Scenario							
A	Velocity field						
	Turbulence	Insufficient	22.3%	Yes	No	Unbounded	0.01m
	Vortex Shedding	Insufficient	>100%	Yes	No	Unbounded	N/A

Besides, by assuming that the scaling methodology is consistent and improvable (Assumption A5), this study then identifies three issues of established validation framework, including the formulation and assessment of validation goal, obscure decision-making process for the code adequacy, and insufficient code performance. Three new technologies, including sufficient accuracy, Predictive Capability Maturity Quantification by Bayesian Network (PCMQBN), and local data-driven Uncertainty Quantification (UQ) are formalized and demonstrated with case studies. A flowchart of the improved validation framework is suggested by incorporating three new methodologies, and the scopes of two additionally required methodologies are defined. At the same time, the improved framework will contribute to the

development of Risk-Informed Evaluation Model Development and Assessment Process (REMDAP).

First, a concept of sufficient accuracy is developed for guiding the formalization of model credibility and the designation of validation goal, which are consistent with the risk-informed concept and safety-analysis purpose. The sufficient accuracy is designed for situations with large but not deep uncertainties (assumption A3 and A4), and it combines both probabilistic and deterministic analysis by representing the validation goal with subjective beliefs and objective maturity levels. Comparing to the regular approach, the sufficient accuracy explicitly connects accuracy requirements with safety analysis in designating validation goals. At the same time, the sufficient accuracy admits the qualitative and subjective nature of model validation under large uncertainties (assumption A3). It also enables a sophisticated documenting capability for all involving human artifacts for remaining convincing with subjective inputs. Moreover, the sufficient accuracy aims to avoid the excess conservatism by seeking the critical belief for accepting the models. In this study, the critical belief is defined as the situations where model predictions can always satisfy the safety requirements if the subjective beliefs from assessments are higher than the critical values. Two case studies are performed to demonstrate the concept and application of sufficient accuracy. Meanwhile, the critical beliefs on the model maturity are found according to the safety requirements and the sampling of uncertain scenarios. The acceptance domain and two extreme regions are identified. The case study also shows the process of propagating the validation goal from an integral model to sub-scale models. It is found that the critical beliefs assessed by sub-scale models are higher than those by the integral model. At the same time, the error ranges for predictions from sub-scale models are much narrower than those from prototypic models since errors propagated from sub-scale models are physics-constraint.

Second, a technique of PCMQBN is developed to formalize the validation's adequacy-decision-making process with mathematical languages. Considering the subjective nature of validation, the validation decision-making process is first defined as an argument process (assumption B1). Next, aiming at three components (evidence, structure, and acceptance criterion) of the validation arguments, PCMQBN further represents the argument process with Bayesian Decision theory (assumption B2-B5): the evidence is characterized according to the concept of

sufficient accuracy; the argument structure is constructed according to the validation cubic model; the acceptance criterion is made according to the sufficient-accuracy-guided validation goal. Next, the Bayesian network is used for graphically representing the adequacy decision-making process (assumption B6). When the scenarios have Level 1 uncertainty, the PCMQBN becomes a Bayesian inference process, where the maturity level of sub-attributes and the conditional probabilities can be directly inferred. For scenarios with Level 2 or 3 uncertainties, PCMQBN can support the integration of evidence and the argument towards the adequacy decision for application purposes. Besides, two case studies are prepared for demonstrating the application of PCMQBN in estimating the model credibility. It is found that for the same model, it is preferable to have the model parameters being calibrated by the QoI directly rather than being calibrated separately. Also, when the data of QoI are used directly for model calibrations (posterior information), the experiments that measure the QoI should be more important than others. Besides, it is found that when only prior information is used for model prediction, the model credibility depends heavily on the amount and coverage of prior information. If important information and highly correlated parameters cannot be accurately measured, a complicated model may result in bad predictions.

Third, it is found that the uncertainty of SPH simulations can hardly be represented precisely by parametric distributions (assumption C1). Therefore, a local data-driven UQ method is developed based on the technique of Physics-Evaluated Machine Learning (PEML). The objective is to quantify the total uncertainty of SPH simulations based on local simulation errors (assumption C2). Random forest is used as the machine-learning engine for constructing a surrogate between local physical features and simulation uncertainties. Five lid-driven cavity simulations with different Re_τ are set, the machine learning engine is trained and tested with different cases. The performance of the constructed surrogate has been qualitatively analyzed. It is found that for most cases, the constructed surrogate is able to reduce the local errors of low-fidelity SPH simulations. Besides, it is found that as more datasets are used for training, the performance of the constructed surrogate becomes better. Also, the performance of interpolated PEML is found to be better than that of extrapolated PEML. The interpolated PEML is defined as situations where the Reynolds number of the testing cavity flow is covered by the Reynolds number of training cases, and the extrapolated PEML is the situation where the Reynolds number of the testing cavity is outside the range of training cases. Moreover, this study also tests the

capability of feedforward neural network as the machine-learning engine, and it turns out that random forest is faster and more robust than the neural network.

In conclusion, this study first examines the CSAU/EMDAP approach for systematically assessing the credibility of SPH methods in RISMC applications. It is found that with the collected database and selected SPH implementations, SPH methods are expected to be "inadequate" in simulating the designated external-flooding and high-wind scenarios. Next, this study identifies three framework issues during the assessment, and new methodologies are suggested. Finally, an improved framework is suggested by incorporating the formalized methodologies, and potential challenges are discussed.

6.2. Contributions

Major contributions of this study include:

(1) The systematic validation of SPH simulations for the transient and accident analysis.

Previously, the validation of SPH simulations is performed independently. Without a specific intended use, validation goals and investigated phenomena of previous validations are rarely connected. In this study, to inform the nuclear safety analysis, SPH method is selected as the simulation tool for predicting the generation, propagation, and interaction of external hazards, including external flooding and high wind, with the nuclear SSCs. Next, by collecting the available database and applying selected SPH implementations, this study systematically assesses the credibility of SPH methods for the designated scenarios of external flooding and high wind. Besides, to ensure the reliability and consistency, this study employs the CSAU/EMDAP validation framework.

(2) The development of the concept of sufficient accuracy for guiding the validation of RISMC tools.

Previously in the EMDAP/CSAU framework, the validation goal is usually defined by simulation accuracy/error information calculated using validation metrics, and the subjective

nature of validation under large uncertainties is not discussed. Moreover, the designation and propagation of validation goals are not formalized, especially when multiple phenomena, processes, and components are involved. In this study, the subjective nature of validation under large uncertainties are stressed, and a concept of sufficient accuracy is suggested for formalizing the representation of model credibility. To be specific, the "accuracy" is represented by the maturity level that objectively tracks the model's quality, and the "sufficient" is characterized by the decision-makers' subjective beliefs on the model's quality. Guided by the concept of sufficient accuracy, the validation goal can be directly related to the safety goal of risk-informed safety analysis methodologies. At the same time, the formalization of sufficient accuracy concepts helps the designation and propagation of validation goals for multi-scale models. The sufficient-accuracy-guided validation goal also enables a formalized decision-making process for model adequacy. Case studies are performed to demonstrate and evaluate its applicability, while potential issues are also discussed.

(3) The application of Physics-Evaluated Machine Learning in quantifying the uncertainty of SPH methods based on local information.

Previously, the characterization of SPH uncertainties is a great challenge since the particle will distribute in a disordered pattern after the simulation starts. The well-established UQ methods for scientific computing require a separation of simulation uncertainties with respect to their sources and aim to characterize each uncertainty term with parametric distribution. However, all uncertainty terms of SPH methods are tightly coupled, and they can hardly be represented precisely by parametric distributions. In this study, the concept of Physics-Evaluated Machine Learning (PEML) is applied to construct a surrogate model between the local error of coarse-grid SPH simulations and the coarse-grid local physical features using Machine Learning algorithm. The performance of PEML has been qualitatively analyzed with respect to the similarity between training and testing cases, data amount, and machine learning engine. Case studies are performed to demonstrate and evaluate its applicability, while potential issues are also discussed.

(4) The application and improvement of PCMQBN for the model credibility and adequacy assessment.

Previously, the decision-making process towards model adequacy is heuristic and obscure. When evidence from full-scale facilities and prototypic conditions are not available, it becomes hard to remain convincing when the evidence from reduce-scale facilities and separate phenomena are integrated towards the adequacy decision. In this study, the validation process is formalized as an argument process. Next, Bayesian decision theory is employed for representing the validation argumentation with mathematical languages. At the same time, based on the concept of sufficient accuracy, the model credibility is represented by a combination of objective maturity level and subjective belief. Next, the integration of evidence is represented graphically by Bayesian Network. Finally, monetary values are assigned to each maturity level, and the adequacy decision can be made transparently and consistently by comparing the expected monetary rewards against the criteria. Case studies are prepared for demonstrating and evaluating the applicability of this method, while limitations and potential issues are also discussed.

(5) The improvement of CSAU/EMDAP validation framework by incorporating new methods and techniques.

In this study, the concept of sufficient accuracy is developed for adapting the well-established CSAU/EMDAP framework to the risk-informed concept. The sufficient accuracy stresses the subjective nature of validation under large uncertainties. And it helps the designation and propagation of validation goals through multi-scale models. At the same time, the Predictive Capability Maturity Quantification by Bayesian Network (PCMQBN) is incorporated for transparently and consistently representing the decision-making process of model adequacy. Moreover, the Physics-Evaluated Machine Learning (PEML) technique is introduced for improving the uncertainty quantification process for SPH methods with novel machine learning algorithms. Meanwhile, an improved CSAU/EMDAP framework is suggested by incorporating the formalized methodologies. Besides, scopes are defined for two additionally required methodologies, including the Validation Data Plan (VDP) and Uncertainty Scaling. At last, this study aims to contribute to the development and assessment of Risk-Informed Evaluation Model Development and Assessment Process (REMDAP) framework.

6.3. Recommendations for Future Work

Based on the progress of this study, further improvements can still be made, including:

(1) To improve the capability and robustness of PCMQBN

In this study, the evidence characterizations, validation structures, and acceptance criteria are prepared based on the author's knowledge. Mathematical Models for validation cubic and weight factor are tentative. Proper formalization and sophisticated analysis for complex case studies are still needed. Besides, for Bayesian applications, strong assumptions must be made for the precision of probability and utility such that they can be represented precisely with a single additive value [53]. Under the large uncertainty of the validation process, additional analysis is needed to make the results convincing. One convenient approach is called Robust Bayes [75], where multiple validation assessments are performed, and the classes of results are ensembled to inform the final decision of model credibility. Besides, the sensitivity study is a unique feature enabled by the quantified framework. Once an agreement is reached on the validation structures, acceptance criteria, and evidence characterizations, PCMQBN validation can be performed transparently and independently by individuals from the expert community. And the "defense-in-depth" can be assured for the validation quality.

(2) To improve the capability and validation of SPH simulations

For the SPH methods, it is recommended that additional models, including turbulence model, wall function, the pressure-velocity coupling model, and adaptive particle size, etc., should be developed. Also, more data from the large-scale flooding facilities should be gathered. In this study, as a part of the integrated research project, a flooding test facility has been constructed in George Washington University and is ready to generate data for various phenomena. At the same time, teams from Centroid Lab Inc. has been actively developing and incorporating new models for NEUTRINO software. Collaborations with both teams will continue improving the capability of SPH and helping the validation progress.

(3) To improve the local data-driven UQ technique and uncertainty extrapolations

In section 5.3, only qualitative relationships are constructed between the PEML-predicted total uncertainty and the training case. Characterization of training cases by similarity index, like Reynolds ratio and K-L divergence, does not show a good result. More metrics and methods should be tested for determining the similarity index. It is also suggested that a new concept named "physical coverage" should be used for extrapolating the simulation uncertainty rather than the similarity. Moreover, considering the complex nature of SPH methods, advanced techniques, like physics-integrated ML, should be tested for quantifying the SPH simulation uncertainty.

(4) To formalize and develop the concept of VDP and uncertainty scaling

In section 5.4, a schematic flowchart has been prepared as the improved EMDAP framework. Two important components, VDP and uncertainty scaling, has been investigated concerning their requirements and preliminary developments. Further developments are still needed for formalizing their assumptions, theories, and procedures. Case studies are also needed for demonstrating and validating the proposed concepts.

7. REFERENCES

[1] BNN, "Fukushima accident rates as major disaster on international scale," Baltic News Network, 16 March 2011. [Online]. Available: http://bnn-news.com/fukushima-accident-rated-international-scale-major-disaster-20982. [Accessed 29 May 2017].

[2] U.S. NRC, "ESBWR Design Control Document," U.S. NRC, Washington DC, 2010.

[3] U.S. NRC, "Westinghouse AP1000 Design Control Document Rev. 19," U.S. NRC, Washington DC, 2011.

[4] D. Wagner, M. Casada and J. Fussell, "Methodology for flood risk analysis for nuclear power plants," in *Low-Probability High-Consequence Risk Analysis*, New York, Springer US, 1984, pp. 65-80.

[5] R. H. Perkins, M. T. Bensi, J. Philip and S. Sancaktar, "Screening analysis report for the proposed generic issue on flooding of nuclear power plant sites following upstream dam failures," U.S. Nuclear Regulatory Commission, Washington D.C., 2011.

[6] "Flood Risk at Nuclear Power Plants," Union of Concerned Scientists, [Online]. Available: http://www.ucsusa.org/nuclear-power/nuclear-power-accidents/flood-risk-at-nuclear-power-plants#.WLRYSW_yupo. [Accessed 27 Feburary 2017].

[7] U.S. NRC, "Recommendations for Enhancing Reactor Safety in the 21st Century, The Near-Term Task Force Review of Insights from the Fukushima Dai-ichi Accident," U.S. NRC, Washington DC, 2011.

[8] U.S. NRC, "Recommended Action to be taken without Delay from the Near-Term Task," U.S. NRC, Washington DC, 2011.

[9] U.S. NRC, "Guidence for Performing the Integrated Assessment for External Flooding," U.S. NRC, Washington DC, 2012.

[10] U.S. NRC, "Guidance For Assessment of Flooding Hazards Due to Dam Failure," U.S. NRC, Washington DC, 2013.

[11] U.S. NRC, "Guidance for Performing a Tsunami, Surge, or Seiche Hazard Assessment," U.S. NRC, Washington DC, 2012.

[12] U.S. NRC, "What are the Lessons Learned from Fukushima?," U.S. NRC, 08 May 2017. [Online]. Available: https://www.nrc.gov/reactors/operating/ops-experience/japan-dashboard/priorities.html. [Accessed 28 May 2017].

[13] G. Hanson, K. Cook and S. Hunt, "Physical modeling of overtopping erosion and breach formation of cohesive embankments," *Transactions of the ASAE,* vol. 48, no. 5, pp. 1783-1794, 2005.

[14] W. Wu, "Earthen embankment breaching," *Journal of Hydraulic Engineering,* vol. 137, no. 12, pp. 1549-1564, 2011.

[15] Flash Back, "2004 Indian Ocean Earthquake and Tsunami Documentary," Flash Back, 7 October 2015. [Online]. Available: https://www.youtube.com/watch?v=UcKlaClpR34. [Accessed 29 May 2017].

[16] ADCIRC, "ADCIRC a (parallel) advanced circulation model for oceanic, coastal and estuarine waters," The University of North Carolina at Chapel Hill, Chapel Hill, 2006.

[17] K. Emanuel, S. Ravela, E. Vivant and C. Risi, "A Statistical-Deterministic Approach to Hurricane Risk Assessment," *Bulletin of American Meteorological Society,* vol. 87, pp. 299-314, 2006.

[18] U.S. NRC, "Effect of Hurricane Andrew on Turkey Point Nuclear Generating Station and Lessons Learne," United States Nuclear Regulatory Commission, Washington, 1994.

[19] D. Lochbaum, "Fission Stories #48: Hurricane Andrew vs. Turkey Point," Union of Concerned Scientists, 12 July 2011. [Online]. Available: http://allthingsnuclear.org/dlochbaum/fission-stories-48-hurricane-andrew-vs-turkey-point. [Accessed 27 Feburary 2017].

[20] F. Hebdon, "Effect of Hurricane Andrew on the Turkey Point Nuclear Generating Station from August 20-30, 1992," United States Nuclear Regulatory Commission, Washington, 1993.

[21] U.S. NRC, "Procedural and submittal guidance for the individual plant examination of external events (IPEEE) for severe accident vulnerabilities," U.S. NRC, Washington DC, 1991.

[22] U.S. NRC, "Tornado Missile Protection," U.S. NRC, Washington, 2015.

[23] U.S. NRC, "Design Basis Tornado and Tornado Missiles for Nuclear Power Plants," United States Nuclear Regulatory Commission, Washington, 2001.

[24] Union of Concerned Scientists, "Hurricane Andrew vs. Turkey Point," 12 July 2011. [Online]. Available: http://allthingsnuclear.org/dlochbaum/fission-stories-48-hurricane-andrew-vs-turkey-point. [Accessed 25 Oct 2016].

[25] U.S. NRC, "NEDO-33201, Revision 6, ESBWR Design Certification Probabilistic Risk Assessment," U.S. NRC, Washington DC, 2010.

[26] C. Smith, C. Rabiti, R. Martineau and R. Szilard, "Risk-Informed Safety Margins Characterization (RISMC) Path Technical Program Plan," Idation National Laboratory, Idaho Falls, 2015.

[27] U.S. NRC, "PRA Procedures Guide – A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants," United States Nuclear Regulatory Commission, Washington, 1983.

[28] U.S. NRC, "Tornado Climatology of the Contiguous United States," U.S. NRC, Washington, 2007.

[29] Storm Prediction Center, "The Enhanced Fujita Scale (EF Scale)," 26 October 2014. [Online]. Available: http://www.spc.noaa.gov/efscale/. [Accessed 27 October 2016].

[30] U.S. NRC, "Standard Review Plan," U.S. NRC, Washington DC, 1981.

[31] EPRI, "Tornado Missile Risk Analysis," Electric Power Research Institute, Palo Alto, 1978.

[32] L. A. Twisdale and W. L. Dunn, "Probabilistic Analysis of Tornado Wind Risks," American Society of Civil Engineers, Raleigh, 1983.

[33] U.S. NRC, "Use of tormis computer code for assessment of tornado missile protection," U.S. NRC, Washington DC, 2008.

[34] K. Hope, N. Povroznyk and R. Schneider, "Tornado Missile Strike Calculator: : An excel-based stochastic model of tornado-driven missile behavior for use in high winds PRA," in *International Topical Meeting on Probabilistic Safety Assessment and Analysis*, Sun Valley, 2015.

[35] C. Rabiti, A. Alfonsi, J. Cogliati, D. Mandelli, R. Kinoshita and C. Smith, "A novel method of controlling thermal hydraulics codes using RAVEN," LWRS Newsletter, 2014.

[36] R. Youngblood and C. Smith, "Technical approach and results from the fuels pathway on an alternative selection case study," Idaho National Laboratory, Idaho Falls, 2013.

[37] J. Lee and N. McCormick, Risk and safety analysis of nuclear systems, Wiley, 2011.

[38] S. Prescott, D. Mandelli, R. Sampath, C. Smith and L. Lin, "3D Simulation of External Flooding Events for the RISMC Pathway," Idaho National Laboratory , Idaho Falls, 2015.

[39] S. Prescott, R. Sampath, C. Smith and T. Yang, "3D Modeling Engine Representaiton Summary Report," Idaho National Laboratory, Idaho Falls, 2014.

[40] C. Smith, D. Mandelli, S. Prescott, A. Alfonsi, C. Rabiti, J. Cogliati and R. Kinoshita, "Analysis of pressurized water reactor station blackout caused by external flooding using the RISMC toolkit," Idaho National Laboratory, Idaho Falls, 2014.

[41] AIAA, "Guide for the Verification and Validation of Computational Fluid Dynamics Simulations," American Institute of Aeronautics and Astronautics, Reston, 1998.

[42] U.S. NRC, "Transient and accident analysis methods," U.S. Nuclear Regulatory Commission, Washington D.C., 2005.

[43] W. Oberkampf and C. Roy, Verification and Validation in Scientific Computing, Cambridge University Press, 2010.

[44] N. Zuber, G. E. Wilson, M. Ishii, M. Wulff, B. Boyack and A. Dukler, "An integrated structure and scaling methodology for severe accident technical issue resolution: Development of methodology," *Nuclear Engineering and Design,* vol. 186, pp. 1-21, 1998.

[45] B. Boyack, I. Catton, R. Duffey, P. Griffth, K. Katsma, G. Lellouche, S. Levy, U. Rohatgi, G. Wilson, W. Wulff and N. Zuber, "Quantifying reactor safety margins part 1: An overview of the code scaling, applicability and uncertainty evaluation methodology," *Nuclear Engineering and Design,* vol. 119, no. 1, pp. 1-15, 1990.

[46] T. Olivier and S. Nowlen, "A phenomena identification and ranking table (PIRT) exercise for nuclear power plant fire modeling applications (NUREG/CR-6978)," U.S. NRC, Washington DC, 2008.

[47] C. Fletcher, P. Bayless, C. Davis, M. Ortiz, S. Sloan, R. Shaw, R. Shultz, C. Slater, G. Johnsen, J. Adams, L. Ghan and D. Bessette, "Adequacy Evaluation of RELAP5/MOD3, Version 3.2.1.2. for simulating AP600 Small Break Loss-of-Coolant Accidents," Idaho National Engeneering and Environment Laboratory, Idaho Falls, 1997.

[48] W. Oberkampf, M. Pilch and T. Trucano, "Predictive Capability Maturity Model for computational modeling and simulation (SAND2007-5948)," Sandia National Laboratories, Albuquerque, 2007.

[49] J. Kaizer, "Fundamental Theory of Scientific Computer Simulation Review," U.S. NRC, Washington DC, 2013.

[50] V. A. Mousseau and N. Dinh, "CASL Verification and Validation Plan," Consortium for Advanced Simulation of Light Water Reactors, CASL-U-2016-1116-000, 2016.

[51] M. Morgan and M. Henrion, Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis, 1st. edition, Cambridge: Cambridge University Press, 1990.

[52] D. Vose, Risk analysis: a quantitative guide, 3rd edition, New York: Wiley, 2008.

[53] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall/CRC, 1990.

[54] C. Roy and W. Oberkampf, "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing," *Computer Methods in Applied Mechanics and Engineering,* vol. 200, pp. 2131-2144, 2011.

[55] P. Roache, Verification and Validation in Computational Science and Engineering, Albuquerque: Hermosa Publishers, 1998.

[56] R. Bank, "Hierarchical bases and the finite element method," *Acta Numerica,* vol. 5, pp. 1-43, 1996.

[57] P. Cavallo and N. Sinha, "Error quantification for computational aerodynamics using an error transport equation," *Journal of Aircraft,* vol. 44, pp. 1954-1963, 2007.

[58] I. Celik, U. Ghia, P. Roache and C. J. Freitas, "Procedure for estimation and reporting of uncertainty due to discretization in CFD applications," *Journal of Fluid Engineering.*

[59] P. Roache, K. Ghia and F. White, "Editorial policy statement on control of numerical accuracy," *Journal of Fluid Engineering,* vol. 108, p. 1, 1986.

[60] R. Howard, "Decision analysis: applied decision theory," in *Proceedings of the Fourth International Conference on Operations Research*, 1966.

[61] G. Parnell, T. Bresnick, S. Tani and E. Johnson, Handbook of Decision Analysis, Wiley, 2013.

[62] R. Howard, "Decision analysis: practice and promise," *Management Science,* vol. 34, no. 6, pp. 679-695, 1988.

[63] X. Jiang and S. Mahadevan, "Bayesian risk-based decision method for model validation under uncertainty," *Reliability Engineering and System Safety,* vol. 92, pp. 707-718, 2007.

[64] R. Zhang and S. Mahadevan, "Bayesian methodology for reliability model acceptance," *Reliability Engineering and System Safety,* vol. 80, pp. 95-103, 2003.

[65] R. Rebba and S. Mahadevan, "Computational methods for model reliability assessment," *Reliability Engineering and System Safety,* no. 93, pp. 1197-1207, 2008.

[66] H. Courtney, 20/20 Foresight: crafting strategy in an uncertain world, Boston: Harvard School Press, 2001.

[67] S. Makridakis, R. Hogarth and A. Gaba, "Forecasting and uncertainty in the economic and business world," *International Journal of Forecasting ,* vol. 25, pp. 794-812, 2009.

[68] W. Walker, R. Lempert and J. Kwakkel, "Deep Uncertainty," 2012.

[69] F. Hillier and G. Lieberman, Introduction to operations research, Boston: McGraw-Hill, 2001.

[70] A. Patt and D. Schrag, "Using specific language to describe risk and probability," *Climatic Change,* vol. 61, pp. 17-30, 2003.

[71] R. Lempert, S. Popper and S. Bankes, "Shaping the next one hundred years: new methods for quantitative long-term strategy analysis, MR-1626-RPC," The RAND Pardee Center, Santa Monica, 2003.

[72] N. Taleb, The black swan: the impact of the highly improbable, Random House, 2007.

[73] J. Rosenhead and J. Mingers, Rational analysis for a problematic world revisited, West Sussex: Wiley, 2001.

[74] K. Heijden, Scenarios: the art of strategic conversation, Wiley, 1996.

[75] J. Berger, "The robust Bayesian viewpoint (with discussion)," *Robustness of Bayesian Analyses,* pp. 63-144, 1984.

[76] G. Ganzenmuller and M. O. Steinhauser, "The implementation of smooth particle hydrodynamics in LAMMPS," 17 July 2011. [Online]. Available: https://lammps.sandia.gov/doc/PDF/SPH_LAMMPS_userguide.pdf. [Accessed 29 October 2018].

[77] R. Sampath, "Neutrino document (release 1.0)," March 2018. [Online]. Available: https://media.readthedocs.org/pdf/neutrinodocs/master/neutrinodocs.pdf. [Accessed 10 October 2018].

[78] CentroidPic Inc., "NEUTRINO," Centroid LAB Inc., 2015. [Online]. Available: http://www.neutrinodynamics.com. [Accessed 20 October 2018].

[79] R. A. Gingold and J. J. Monaghan, "Smoothed Particle Hydrodynamics: theory and application to non-spherical star," *Monthly Notices of the Royal Astronomical Society,* pp. 375-389, 1977.

[80] L. B. Lucy, "A numerical approach to the testing of the fission hypothesis," *Astronomical Journal,* vol. 82, no. 12, pp. 375-389, 1977.

[81] D. Violeau and R. Issa, "Numerical modelling of complex turbulent free-surface flows with the SPH method: an overview," *International Journal for Numerical Methods in Fluids,* vol. 53, no. 2, pp. 277-304, 2007.

[82] B. Ataie-Ashtiani and G. Shobeyri, "Numerical simulation of landslide impulsive waves by incompressible smoothed particle hydrodynamics," *International Journal For Numerical Methods in Fluids,* vol. 56, pp. 209-232, 2008.

[83] T. Capone, A. Panizzo and J. Monaghan, "SPH modelling of water waves generated by submarine landslides," *Journal of Hydraulic Research,* vol. 48, pp. 80-84, 2010.

[84] R. Vacondio, B. Rogers, P. Stansby and P. Mignosa, "Shallow water SPH for flooding with dynamic particle coalescing and splitting," *Advances in Water Resources,* vol. 58, pp. 10-23, 2013.

[85] J. Wu, H. Zhang, R. Dalrymple and A. Herault, "Numerical modeling of dam-break flood in city layouts including underground spaces using GPU-based SPH method," *Journal of Hydrodynamics,* vol. 25, no. 6, pp. 818-828, 2013.

[86] A. Souto-Iglesias, L. Delorme, L. Perez-Rojas and S. Abril-Perez, "Liquid moment amplitude assessment in sloshing type problems with smooth particle hydrodynamics," *Ocean Engineering,* vol. 33, pp. 1462-1484, 2006.

[87] B. Cartwright, A. Chhor and P. Groenenboom, "Numerical simulation of a helicopter ditching with emergency flotation devices," in *Proceeding of 10th SPHERIC International Workshop*, Parma PR, 2010.

[88] S. Marrone, B. Bouscasse, A. Colagrossi and M. Antuono, "Study of ship wave breaking patterns using 3D parallel SPH simulations," *Computers and Fluids,* vol. 69, pp. 54-66, 2012.

[89] A. Zhang, X. Cao, F. Ming and Z. Zhang, "Investigation on a damaged ship model sinking into water based on three dimensional SPH method," *Applied Ocean Reseawrch,* vol. 42, pp. 24-31, 2013.

[90] L. Qiu, Y. Liu and Y. Han, "A 3D Simulation of a Moving Solid in Viscous Free-Surface Flows by Coupling SPH and DEM," *Mathematical Problems in Engineering,* 2017.

[91] C. Altomare, A. Crespo, J. Dominguez, M. Gomez-Gesteira, T. Suzuki and T. Verwaest, "Applicability of smoothed particle hydrodynamics for estimation of sea wave impact on coastal structures," *Coastal Engineering ,* vol. 96, pp. 1-12, 2015.

[92] S. Lind, P. Stansby, B. Roger and P. Lloyd, "Numerical predictions of water-air wave slam using incompressible-compressible smoothed particle hydrodynamics," *Applied Ocean Research,* vol. 49, pp. 57-71, 2015.

[93] X. Ni, W. Feng and D. Wu, "Numerical simulations of wave interactions with vertical wave barriers using the SPH method," *International Journal for Numerical Methods in Fluids,* vol. 76, no. 4, pp. 223-245, 2014.

[94] D. Violeau and B. D. Rogers, "Smoothed particle hydrodynamics (SPH) for freesurface flows: past, present and future," *Journal of Hydraulic Research,* vol. 54, pp. 1-26, 2016.

[95] R. A. Dalrymple and B. D. Rogers, "Numerical modeling of water waves with the SPH method," *Coastal Engineering,* pp. 141-147, 2006.

[96] A. Leroy, "A new incompressible SPH model : towards industrial applications (PhD Thesis)," University Paris-Est, Paris, 2015.

[97] L. Lin, "Assessment of the smoothed particle hydrodynamics method for nuclear thermal-hydraulic applications," North Carolina State Uninversity, Raleigh, 2016.

[98] F. Philip, "A new class of accurate, mesh-free hydrodynamics simulation methods," *Monthly Notices of the Royal Astronomical Society,* vol. 450, pp. 53-110, 2015.

[99] J. Morris, "A study of the stability properties of smooth particle hydrodynamics," in *Proceedings of the Astronomical Soceity of Australia*, 1995.

[100] J. Read, T. Hayfield and O. Agertz, "Resolving mixing in smoothed particle hydrodynamics," *Monthly Notices of the Royal Astronomical Society,* vol. 405, pp. 1513-1530, 2010.

[101] W. Dehnen and H. Aly, "Improving convergence in smoothed particle hydrodynamics simulations without pairing instability," *Monthly Notices of the Royal Astronomical Society,* vol. 425, no. 2, pp. 1068-1082, 2012.

[102] G. Dilts, "Moving-least-squares-particle hydrodynamics - I. consistency and stability," *International Journal for Numerical Methods in Engineering,* vol. 44, pp. 1115-1155, 1999.

[103] G. Dilts, "Moving least-squares particle hydrodynamics II: conservation and boundaries," *International Journal for Numerical Methods in Engineering,* vol. 48, pp. 1503-1524, 2000.

[104] J. Morris and J. Monaghan, "A switch to reduce SPH viscosity," *Journal of Computational Physics,* vol. 136, pp. 41-50, 1997.

[105] J. Monaghan, "On the problem of penetration in particle methods," *Journal of Computational Physics,* vol. 82, pp. 1-15, 1989.

[106] J. J. Monaghan, "Smoothed Particle Hydrodynamics," *Annual Revision Astronomy Application,* vol. 30, pp. 543-574, 1992.

[107] R. Xu, P. Stansby and D. Laurence, "Accuracy and stability in incompressible SPH (ISPH) based on the projection method and a new approach," *Journal of Computational Physics, 228(18),* p. 6703–6725, 2009.

[108] S. Lind, R. Xu and S. P., "Incompressible smoothed particle hydrodynamics for free-surface flows: A generalised diffusion-based algorithm for stability and validations for impulsive flows and propagating waves[J]," *Journal of Computational Physics,* pp. 1499-1523, 2012.

[109] X. Hu and N. Adams, "An incompressible multi-phase SPH method," *Journal of Computational Physics,* vol. 227, pp. 264-278, 2007.

[110] M. Ellero, M. Serrano and P. Espanol, "Incompressible smoothed particle hydrodynamics," *2007,* vol. 226, pp. 1731-1752, 2007.

[111] S. Shao and E. Lo, "Incompressible SPH method for simulating Newtonian and non-Newtonian flows with a free surface," *Advances in Water Resources,* vol. 26, pp. 787-800, 2003.

[112] S. Cummins and M. Rudman, "An SPH Projection Method," *Journal of Computational Physics,* vol. 152, pp. 584-607, 1999.

[113] A. Mahmoudi, H. H. Zadeh and M. J. Ketabdari, "A comparison between performaces of turbulence models on simulation of solitary wave breaking by WCSPH method," *Journal of Civil Engineering and Urbanism,* 2014.

[114] G. Batchelor, Introduction to Fluid Mechanics, Cambridge: Cambridge University Press, 2000.

[115] J. P. Morris, P. J. Fow and Y. Zhu, "Modeling low Reynolds number incompressible flows using SPH," *Journal of Computational Physics,* vol. 136, pp. 214-226, 1997.

[116] M. Robinson, "Turbulence and viscous mixing using smoothed particle hydrodynamics (PhD Thesis)," Unversity of Monash, Melbourne, 2009.

[117] A. Valizadeh and J. Monaghan, "Smoothed particle hydrodynamics simulations of turbulence in fixed and rotating boxes in two dimensions with no-slip boundaries," *Physics of Fluids,* vol. 24, no. 3, 2012.

[118] A. Mayrhofer, D. Laurence, B. Rogers and D. Violeau, "DNS and LES of 3-D wall-bounded turbulence using Smoothed Particle Hydrodynamics," *Computers & Fluids,* vol. 115, pp. 86-97, 2015.

[119] W. Welton and S. Pope, "PDF Model Calculations of Compressible Turbulent Flows Using Smoothed Particle Hydrodynamics," *Journal of Computational Physics,* vol. 134, pp. 150-168, 1997.

[120] S. Adami, X. Hu and N. Adams, "Simulating three-dimensional turbulence with SPH," in *Proceedings of the 8th SPHERIC International Workshop*, Trondheim, 2012.

[121] D. Violeau, "One and two-equations turbulent closures for Smoothed Particle Hydrodynamics," in *Proceeding of the 6th International Conference in Hydroinformatics*, 2004.

[122] S. Shao, "Incompressible SPH simulation of wave breaking and overtopping with turbulence modeling," *International Journal for Numerical Methods in Fluids,* pp. 597-621, 2006.

[123] J. Arai, S. Koshizuka and K. Murozono, "Large eddy simulation and a simple wall model for turbulent flow calculation by a particle method," *Internal Journal For Numerical Methods in Fluids,* pp. 772-787, 2013.

[124] D. Violeau, S. Piccon and J. Chabard, "Two attempts of turbulence modelling in smoothed particle hydrodynamics," *Advances in Fluid Modelling and Turbulence Measurements,* pp. 339-346, 2002.

[125] J. Gray and J. J. Monaghan, "Numerical modelling of stress fields and fractures around magmachambers," *Journal of Volcanology and Geothermal Research,* vol. 135, no. 3, pp. 259-283, 2004.

[126] S. Simon, "Smooth Particle Hydrodynamics applied to fracture mechanics (master thesis)," Uppsala University, Uppsala, 2013.

[127] B. Maurel and A. Combescure, "An sph shell formulation for plasticity and fracture analysis in explicit dynamic," *International Journal for Numerical Methods in Engineering,* vol. 76, no. 7, pp. 949-971, 2008.

[128] J. Lin, H. Naceur, D. Coutellier and A. Laksimi, "Efficient Meshless SPH Method for the Numerical Modeling of Thick Shell Structures undergoing large deformations," *International Journal of Non-Linear Mechanics,* vol. 65, no. 1, pp. 1-13, 2014.

[129] Bullet Physics Library, "Real-Time Physics Simulation," WordPress, 2017. [Online]. Available: http://bulletphysics.org/wordpress/. [Accessed 10 April 2017].

[130] C. Antoci, M. Gallati and S. Sibilla, "Numerical simulation of fluid–structure interaction by SPH," *Computers & Structures,* vol. 85, no. 11-14, pp. 879-890, 2007.

[131] N. Akinci, "Interface handling in smoothed particle hydrodynamics (PhD thesis)," Albert-Ludwigs-Universität, Freiburg, 2014.

[132] J. J. Monaghan, "Particle methods for hydrodynamics," *Computer Physics Reports,* vol. 3, no. 2, pp. 71-124, 1985.

[133] D. Violeau, Fluid Mechanics and the SPH Method: Theory and Applications, Oxford: OXFORD UNIVERSITY PRESS, 2012.

[134] R. Fatehi and M. T. Manzari, "Error estimation in smoothed particle hydrodynamics and a new scheme for second derivatives," *Computers & Mathematics with Applications,* vol. 61, pp. 482-498, 2011.

[135] V&V/UQ Team, "How to PIRT," Sandia National Laboratories, Albuquerque.

[136] D. Goring, "Tsunamis-the propagation of long waves onto a shelf," California Institute of Technology, Pasadena, 1978.

[137] B. Ren, M. He, P. Dong and H. Wen, "Nonlinear simulations of wave-induced motions of a freely floating body using WCSPH method," *Applied,* vol. 50, pp. 1-12, 2015.

[138] C. Smith, S. Prescott, J. Coleman, E. Ryan, B. Bhandari, D. Sludern, C. Pope and R. Sampath, "Progress on the industry application external hazard analyese early demonstration (INL/EXT-15-36749)," Idaho National Laboratory, Idaho Falls, 2015.

[139] P. Athe, N. Dinh and H. Abdel-Khalik, "Investigation of similarity metrics for simulation based scaling analysis," in *Advances in Thermal Hydraulics*, New Orleans, 2016.

[140] G. Katell and B. Eric, "Accuracy of solitary wave generation by a piston wave maker," *Journal of Hydraulic Research,* vol. 40, no. 3, pp. 321-331, 2002.

[141] S. J. Cummins, T. B. Silverster and P. W. Cleary, "Three-dimensional wave impact on a rigid structure using smoothed particle hydrodynamics," *International Journal for Numerical Methods in Fluids,* vol. 68, pp. 1471-1496, 2012.

[142] F. Aureli, S. Dazzi, A. Maranzoni, P. Mignosa and R. Vacondio, "Experimental and numerical evaluation of the force due to the impact of a dam-break wave on a structure," *Advances in Water Resources,* vol. 76, pp. 29-42, 2015.

[143] T. Wu, C. Chu, C. Huang, S. Chien and M. Chen, "A two-way coupled simulation of moving solids in free-surface flows," *Computers & Fluids,* vol. 100, pp. 347-355, 2014.

[144] I. Hadzic, J. Hennig, M. Peric and Y. Xing-Kaeding, "Computation of flow-induced motion of floating bodies," *Applied Mathematical Modelling,* vol. 29, pp. 1196-1210, 2005.

[145] Sandia National Laboratory, "LAMMPS-SPH," Sandia National Laboratory, [Online]. Available: http://lammps.sandia.gov.

[146] CD-adapco, "STAR-CCM+ v.11.02 User Guide," CD-adapco, 2016.

[147] A. Khayyer, H. Gotoh and Y. Shimizu, "Comparative study on accuracy and conservation properties of two particle regularization schemes and proposal of an optimized particle shifting scheme in ISPH context," *Journal of Computational Physics,* vol. 332, pp. 236-256, 2017.

[148] The OpenFOAM Foundation, "OpenFOAM v6 User Guide," 2 March 2017. [Online]. Available: https://cfd.direct/openfoam/user-guid. [Accessed 19 November 2018].

[149] R. Bouffanais and M. Deville, "Large-eddy simulation of the flow in a lid-driven cubical cavity," *Physics of Fluids,* vol. 19, no. 5, 2007.

[150] F. Grinstein, L. Margolin and W. Rider, Implicit Large Eddy Simulation: computing turbulent fluid dynamics, Cambridge University Press, 2007.

[151] N. Dinh, "Validation data to support advanced code development," in *The 15th International Topical Meeting on Nuclear Reactor Thermal Hydraulics (NURETH-15)*, Pisa, 2013.

[152] N. Dinh, "CIPS Validation Data Plan (INL/EXT-12-25347)," Idaho National Laboratory,, Idaho Falls, 2012.

[153] P. Athe, "A framework for predictive capability maturity assessment of simulation codes," North Carolina State University , Raleigh, 2018.

[154] T. Theofanous, "On the proper formulation of safety goals and assessment of safety margins for rare and high-consequence hazards," *Reliability Engineering and System Safety,* vol. 54, pp. 243-257, 1996.

[155] H. Haario, M. Laine, A. Mira and E. Saksman, "DRAM: Efficiency adaptive MCMC," *Statistics and Computing,* vol. 16, pp. 339-354, 2006.

[156] U.K. MoD, "Interim Standard 00-55: The procurement of safety critical software in defence equipment," UK Ministry of Defence, 1991.

[157] P. Athe and N. Dinh, "A Framework to Support Assessment of Predictive Capability Maturity of Multiphysics Simulation Codes," U.S. Department of Energy, Washington DC, 2017.

[158] L. Sun, "Establishing confidence in safety assessment evidence," Department of Computer Science, University of York, 2012.

[159] J. Pearl, Bayesian networks: A model of self-activated memory for evidential reasoning, Los Angeles: University of California, 1985.

[160] J. Koski and J. Noble, "A Review of Bayesian Networks and Structure Learning," *Mathematica Applicanda,* vol. 40, no. 1, pp. 53-103, 2012.

[161] P. Larrañaga, H. Karshenas, C. Bielza and R. Santana, "A review on evolutionary algorithms in Bayesian network learning and inference tasks," *Information Sciences,* vol. 233, pp. 109-125, 2013.

[162] S. Kwag, A. Gupta and N. Dinh, "Probabilistic risk assessment based model validation method using Bayesian network," *Reliability Engineering and System Safety,* vol. 169, pp. 380-393, 2018.

[163] N. Dinh, "CASL Validation Data: An Initial Review (CASL-VUQ-L3)," Idaho National Laboratory (INL/EXT-11-21017), Idaho Falls, 2011.

[164] Agena Ltd., "AgenaRisk 7.0 Revision 3023".

[165] J. Salvatier, C. Fonnesbeck and T. Wiecki, "PyMC3," Sphinx, 2016. [Online]. Available: http://docs.pymc.io/index.html. [Accessed 7 November 2017].

[166] U.S. Department of the Interior, "Water measurement manual - a water resources technical publication," U.S. Department of the Interior, Washington DC, 2001.

[167] M. Bos, "Discharge measurement structures," International Institute for Land Reclamation and Improvement, Netherlands, 1989.

[168] S. Lind, R. Xu, P. Stansby and B. Rogers, "Incompressible smoothed particle hydrodynamics for free-surface flows: A generalised diffusion-based algorithm for stability and validations for impulsive flows and propagating waves," *Journal of Computational Physics,* vol. 231, no. 4., pp. 1499-1523, 2012.

[169] H. Xiao, "Quantification of Model Uncertainty in RANS Simulations: A Review," 1 October 2018. [Online]. Available: https://arxiv.org/pdf/1806.10434.pdf. [Accessed 11 November 2018].

[170] Y. Liu and N. Dinh, "Development of a Data-driven Analysis Framework for Pool Boiling and Subcooled Flow Boiling Problems," U.S. Department of Energy, Washington DC, 2017.

[171] L. Lin and N. Dinh, "Development of validation methoddology for RISMC models," North Carolina State University Raleigh, 2018.

[172] E. Parish and K. Duraisamy, "A paradign for data-driven predictive modeling using field inversion and machine learning," *Journal of Computatinal Physics,* vol. 305, pp. 758-774, 2016.

[173] J. Wang, J. Wu and H. Xiao, "Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data," *Physical Review Fluids,* vol. 2, 2017.

[174] C. Chang and N. Dinh, "Classification of machine learning frameworks for data-driven thermal fluid models," *International Journal of thermal sciences (accepted for publication).*

[175] T. Chiang, W.H., Sheu and R. Hwang, "Effect of Reynolds number on the eddy structure in a lid-driven cavity," *International Journal for Numerical Methods in Fluids,* vol. 26, pp. 557-579, 1998.

[176] A. Prasad and J. Koseff, "Reynolds number and end-wall effects on a lid-driven cavity flow," *Physics of Fluids,* vol. 208, 1988.

[177] L. Breiman, J. Friedman, C. Stone and R. Olshen, Classification and regression trees, Boca Raton: CRC Press, 1984.

[178] M. Hagan and M. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transaction on Neural Networks,* vol. 5, no. 6, pp. 989-993, 1994.

[179] M. Beale, M. Hagan and H. Demuth, "Deep learning toolbox: user's guide," 2018. [Online]. Available: https://www.mathworks.com/help/pdf_doc/deeplearning/nnet_ug.pdf. [Accessed 17 November 2018].

[180] S. Bottaro and K. Lindorff-Larsen, "Biophysical experiments and biomolecular simulations: A perfect match," *SCIENCE,* vol. 361, no. 6400, pp. 355-360, 2018.

[181] H. Raiffa and H. Schlaifer, Applied Statistical Decision Theory, Boston : Harvard University Press, 1961.

[182] N. Dinh, A. Gupta, J. B. I. Bolotnov, M. Avramova, H. Abdel-Khalik, X. Sun, P. Bardet, R. Sewell, R. Sampath, J. Lane, A. Bui, S. Prescott, C. Rabiti, R. Youngblood and W. Ren, "Development and Application of a Data-Driven Methodology for Validation of Risk-Informed Safety Margin Characterization Models," North Carolina State University, Raleigh, 2015.

[183] A. Alfonsi, C. Rabiti, D. Mandelli, J. Cogliati, C. Wang, P. Talbot, D. Maljovec and C. Smith, "RAVEN theory manual and user guide," Idaho National Laboratory, Idaho Falls, 2017.

[184] W. Ren, "Nuclear enerugy knowledge based for advanced modeling and simulation - functionalities and operation (Beta) - NE-KAMS Version Beta," Oak Ridge National Laboratory, Oak Ridge, 2012.

[185] N. Dinh, A. Gupta, I. Bolotnov, J. Baugh, M. Avramova, A. Bui, H. Abdel-Khalik, X. Sun, P. Bardet, J. Lane, R. Sewell, R. Sampath, S. Prescott, C. Rabiti, R. Youngblood and W. Ren, "Development and application of a data-driven methodology for validation of risk-informed safety margin characterization (M2NU-16-NC-NCSU-0.0301-152 Year 1 Status Report)," North Carolina State University, Raleigh, 2018.

[186] L. Lin, P. Athe and N. Dinh, "Predictive capability maturity assessment with Bayesian network," North Carolina State University , Raleigh, 2017.

[187] B. Finetti, Theory of probability: a critical introductory treatment, New York: Wiley, 2017.

[188] Wind Science and Engineering Center, "Enhanced F scale for tornado damage," Texas Tech University, Lubbock, 2004.

[189] R. Wang and D. Strong, "Beyond Accuracy: what data quality means to data consumers," *Journal of Management Information Systems,* vol. 12, no. 4, pp. 5-34, 1996.

[190] R. Howard, "Information Value Theory," *IEEE Transactions on Systems,* vol. 2, no. 1, pp. 22-26, 1966.

GLOSSARY

Terminology	Definition
Acceptance Domain	An accuracy requirement for simulating the generation, propagation, and interaction of accident scenarios. The acceptance domain is inherited from the concept of sufficient accuracy and arisen from the application requirements. Other than the practical consideration, no explicit restriction is placed on the acceptance domain.
Bounded Error	An error analysis for the M&S based on the discrepancy between the model prediction and the true value. The bounded error is inherited from the concept of sufficient accuracy and estimated by comparing the model prediction against the true value or high-fidelity data. The bounded error has to be consistent, therefore, the analyzing process should be transparent, complete and robust.
Coverage	The coverage is direct evidence that is assessed by comparing the range of parameters for each phenomenon in the phenomenology pyramid (PP) with the range of parameter for the corresponding model in the model pyramid (MP) and data in data pyramid (DP), the coverage information for each component can be obtained [153].
Consistency	In classical deductive logic, a consistent theory is one that does not contain a contradiction.
Confidence	A subjective belief on a hypothesis based on the current states of knowledge. It is quantified by probability in this study [187].
Credibility	When a simulation model and its results can be accepted by decision-makers for certain application, it can be called credible. In this study, credibility composes the accuracy of model prediction against measured data, the quality of database, the scaling analysis of database, and the acceptance criterion of decision-makers
Data Applicability	Data applicability is direct evidence that concerns the data quality
Deep Uncertainty	The condition in which analysts do not know or the parties to a decision cannot agree upon (1) the appropriate models to describe interactions among a system's variables, (2) the probability distributions to represent uncertainty about key parameters in the models, and/or (3) how to value the desirability of alternative outcomes [68]
Enhanced Fujita (EF) Scale	Published by the National Weather Service [188], the Enhanced Fujita scale is implemented in place of the Fujita scale, and it rates the intensity of tornadoes in the U.S. and Canada based on the damage they cause.
EF 1 scale tornado	Wind speed of EF1 tornado ranges from 138 – 177 km/h. It causes moderate damage, including stripping roofs, overturning and damaging mobile homes, etc. [188]
Extrapolated PEML	The extrapolation PEML is defined as the case where the target fluid characteristics (like Re_τ) falls outside the range of training groups' fluid characteristics

Expected Value of Sample Information (EVSI)	The expected increase in utility that a decision-maker could obtain from gaining access to a sample of additional observations before making a decision [181]
Interpolated PEML	The interpolation PEML is defined as the case where the target fluid characteristics (like Re_τ) falls within the range of training groups' fluid characteristics
Maturity Level	Maturity level is an attempt to objectively track intellectual artifacts, or evidence, obtained in an assessment of an M&S effort [48]. The levels of maturity are determined based on intrinsic information quality and contextual information quality [189]
Nonparametric	Nonparametric technique is based on either distribution-free or distributions with unspecified parameters, like mean and variance. It does not assume any fixed structure of a model.
Predictive Capability Maturity Model (PCMM)	A decision model for assessing the level of maturity of computational modeling and simulation efforts according to six contributing elements [48]
Predictive Capability Maturity Quantification (PCMQ)	The quantified version of PCMM developed by Athe, et al. [157]
Physics-Evaluated Machine Learning	Physics-Evaluated Machine Learning (PEML) is also known as Type II ML framework [174]. Based on prior knowledge, PEML analyzes the discrepancies between observations and simulations. The objective is to inform simulations by observed data such that a consistency can be achieved between data and model.
Relevance [R]	Relevance is direct evidence that is determined based on the geometric similarity and material scaling. It determines the degree of applicability of data based on the preconceived view of phenomenology/process [152]
Subjective	In philosophy, the subjectivity is contrasted to the objectivity, while in this study, the subjective is an interpretation for probability, which should be rationally changed to account for availability of related evidence.
Sufficient Accuracy	A risk-informed concept for M&S validations by adapting the accuracy requirements according to scenarios and tools' limitation. The sufficient accuracy is the theoretical basis for the risk-informed validation framework. It is designated for situations where the effect of uncertainties is dominant. The "sufficient" is measured by the utility on the model accuracy. It requires the validation to be transparent, complete, consistent and improvable.
(Physics) Scaling [S]	Physics scaling is direct evidence that reflects the degree of similarity between phenomena in phenomenology pyramid and experiments in data pyramid on the basis of physics scaling. It determines the gap between the test facility and reactor behavior (phenomena at reactor conditions) [153]

Uncertainty [U]	Data uncertainty is direct evidence that consists of uncertainty in the measured data due to instrumentation errors and limited resolution of measurement instruments [153].
Utility	Utility is the sum of all pleasure minus the suffering that results from an action. And the utility function represents a consumer's preference ordering over a choice set. Considering the difficulty of eliciting utility and the controversy of theory itself, this study uses simple monetary rewards for ranking alternatives and making best decisions.
Validation Result	Validation result is direct evidence that concerns the comparison of model prediction and measurement. Validation result is quantified based on the assessment of two sub-attributes: coverage and validation test result
Value of Information	Value of information refers to the numerical values that are assigned to the elimination or reduction of any uncertainty [190]