Validation, Verification, and Uncertainty Quantification for Models with Intelligent Adversaries

Jing Zhang and Jun Zhuang

Contents

1 Introduction . . . 2
2 Model Verification vs. Validation . . . 3
2.1 Terminology . . . 3
2.2 Relationships Between Verification and Validation . . . 5
3 Validation, Verification, and UQ in the Literature . . . 5
3.1 Validation, Verification, and UQ in Model Development Process . . . 5
3.2 Quantitative Model Validation Techniques . . . 6
4 Validation for Intelligent Adversary Models . . . 7
4.1 Difficulties in Validating Intelligent Adversary Models . . . 7
4.2 Verification and Validation Methods for Intelligent Adversary Models . . . 8
4.3 Validate Intelligent Adversary Models Using Proxy Models . . . 13
5 Conclusion . . . 15
References . . . 16

Abstract

Model verification and validation (V&V) are essential before a model can be implemented in practice. Integrating model V&V into the process of model development can help reduce the risk of errors, enhance the accuracy of the model, and strengthen the confidence of the decision-maker in model results. Besides V&V, uncertainty quantification (UQ) techniques are used to verify

J. Zhang (✉)
Department of Industrial and System Engineering, New York State University at Buffalo, Buffalo, NY, USA
e-mail: jzhang42@buffalo.edu

J. Zhuang
Department of Industrial and Systems Engineering, New York State University at Buffalo, Buffalo, NY, USA
e-mail: jzhuang@buffalo.edu

R. Ghanem et al. (eds.), *Handbook of Uncertainty Quantification*,
DOI 10.1007/978-3-319-11259-6_44-1

and validate computational models. Modeling intelligent adversaries is different from and more difficult than modeling non-intelligent agents. However, modeling intelligent adversaries is critical to infrastructure protection and national security. Model V&V and UQ for intelligent adversaries present a big challenge. This chapter first reviews the concepts of model V&V and UQ in the literature and then discusses model V&V and UQ for intelligent adversaries. Some V&V techniques for modeling intelligent adversaries are provided which could be beneficial to model developers and decision-makers facing with intelligent adversaries.

Keywords
Decision making • Intelligent adversaries • Model validation and verification • Validation techniques

1 Introduction

Models have been extensively used in research when describing systems and predicting scenarios. Model verification and validation (V&V) can quantify confidence in the accuracy of model-based predictions under certain assumptions.

Verification refers to building the system right, while validation refers to building the right system [47]. Verification is conducted before validation. The verification process includes assessing code verification and calculation verification. Validation consists of conceptual model validity and operational validity. There are many approaches to validation, such as validation by assumption, validation by results, and validation by common sense. Validation techniques include animation, comparison to models, degenerate tests, event validity, extreme condition, face validity, fixed values, historical data validation, historical methods, internal validity, multistage validation, operational graphics, parameter variability, predictive validation, traces, and turning tests [58]. See the explanations of the techniques in Table 1.

Academia, industry, and government have been interested in model validation. Some terms related to the model, such as "reliability," "credibility," "confidence," and "applicability," have become common in academic and industrial studies, as well as government reports and implementations. A Standards Committee for the development of model V&V procedures for computational solid mechanics models has been formed by the American Society of Mechanical Engineers (ASME); a V&V model for all safety-related nuclear facility design, analyses, and operations has been supported by the Defense Nuclear Facilities Safety Board (DNFSB); and validation of complex models has been a key concern of the military simulation community for over three decades [39].

Considerable attention has been paid to model verification and validation. Numerous articles have appeared in the literature expressing different concerns of the validity of the models that have been proposed. The advances in the techniques of modeling and solution have impacted how people perceive model validation. For details about model V&V and about computational simulation models [57–59], see [44]. Validation methods, procedures for economic and financial models, urban and transportation models, government and criminology models, and medical and physiological models have also been studied [16].

Table 1 Common model validation techniques (Source: [58])

Techniques	Explanation
Animation	Use graphs to show the model's behavior through time
Comparison to models	Compare model results to the results of other valid models
Degenerate test	Test model behavior using appropriate values of input/internal parameters
Event validity	Compare model event to real system to see the similarity
Extreme condition	Check model plausibility in extreme and unlikely levels of the system
Face validity	Ask knowledgeable people about the reasonability of the model
Fixed values	Fix values for variables/parameters to check against easily calculated values
Historical data validation	Use part of the data to build model, and the rest of data to test model
Historical methods	Three historical methods: rationalism, empiricism, positive economics
Internal validity	Implement several runs to determine the amount of variability in the model
Multistage validation	Combine the three historical methods into a multistage process
Operational graphics	Display values of various performance measures
Parameter variability	Use sensitivity analysis to determine the parameters' effect
Predictive validation	Check the prediction of the model with the system behavior
Traces	Trace entities in the model to see whether the model logic is correct
Turning tests	Ask people to discriminate the outputs of the model and system

The V&V approach quantifies degree of the accuracy and confidence inferred from the comparison of the prediction from the model with the results from reality or experiments. There can be no validation if there is no experimental data with which to compare the result of the model [6]. However, for intelligent adversaries, deficiencies of data and the incompleteness of understanding adversaries' behavior hinder the modeler from building the model and obtaining credible predictions. Taking the characteristics of the intelligent adversaries into consideration, is it possible to obtain sufficient data to build and validate such models? If not, is model validation even possible for intelligent adversary analysis in the absence of outcome data? These questions will be discussed in this chapter.

2 Model Verification vs. Validation

2.1 Terminology

We first introduce the terms of "verification" and "validation" before discussing the relationships between them. Model V&V methods and procedures have been defined by multiple organizations. In the development of fundamental concepts and terminology for V&V, the Department of Defense (DoD) Modeling and Simulation Office (DMSO) has been the leader and played a major role in attempting to standardize the definitions of V&V [7, 8]. In addition, DMSO developed the

Table 2 Definitions of verification and validation by DMSO and IEEE (Source: [7, 24])

	Verification	Validation
DMSO	The process of determining that a model implementation accurately represents the developer's conceptual description and specifications	The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model
IEEE	The process of evaluating a system or component to determine whether the products of a given development phase satisfy the conditions imposed at the start of that phase	The process of evaluating a system or component during or at the end of the development process to determine whether it satisfies specified requirements

US fundamental concepts and terminology for model V&V applied to high-level systems such as ballistic missile defense and battle management simulations [71].

There is a variety of formal definitions. The Defense Modeling and Simulation Organization (DMSO) of the Department of Defense (DoD) and the Institute of Electrical and Electronics Engineers (IEEE) give the most widely used definitions of the terms of verification and validation [44]; see Table 2.

Software engineering and software quality assurance use the IEEE definitions [44]. By contrast, computational simulations in science and engineering and operation research use the DMSO definitions. The DMSO definitions are widely adopted by [1, 32, 33, 55, 63, 71] as well as in this chapter. [71] states "Software V&V is fundamentally different from model V&V. Software V&V is required when a computer program or code is the end product. Model V&V is required when a predictive model is the end product. A code is the computer implementation of algorithms developed to facilitate the formulation and approximate solution of a class of models."

Different papers have defined model V&V differently according to specific contexts. For example, [59] defines model validation as "substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model"; according to [15], model validation refers to activities "to establish how closely the model mirrors the perceived reality of the model use/developer team"; [36] defines model as "activities designed to determine the usefulness of a model; i.e., whether it is appropriate for its intended uses(s); whether the benefits of improving model usefulness exceed the costs; whether the model contributes to making "better" decisions; and possibly how well the particular model performs compared to alternative models"; [46] defines verification and validation as "the process of determining the accuracy with which a computational model can produce results deliverable by the mathematical model on which it is based: code verification and solution verification" and "the process of determining the accuracy with which a model can predict observed physical events (or the important features of a physical reality)" respectively.

2.2 Relationships Between Verification and Validation

Model verification and validation (V&V) are the primary processes for quantifying and building credibility in numerical models and essential parts of the model development process if models are accepted and used to support decision-making. Verification and validation are often mentioned together in requirements to test models, yet they are fundamentally different.

Verifying a model means checking if the model produces the intended output as a function of the inputs, and whether it is mathematically correct. The focus here is on the model implementation and coding [18]. Verification is a critical activity, but is not the same as validation. Fundamentally, model validation is subjective and different perspectives on model validation could have very different meanings. The assertion "the model was judged valid" can mean almost anything, since the modelers choose "the validity tests, the criteria for passing those tests, what models outputs to validate, what setting to test in, what data to use, etc."[37].

Verification is a matter of asking "Did I build the thing right?" and "Have the model and the simulation been built so that they fully satisfy the developer's intent?". By contrast, validation asks "Did I build the right thing?" and "Will the model be able to adequately support its intended use?" "Is its fidelity appropriate for that?" [49].

The purpose of model verification and validation is to assess and improve credibility, accuracy, and trustworthiness. Model verification and validation cannot certify a model to be accurate for all scenarios; but can provide evidence that a model is sufficiently accurate for its intended use [71].

3 Validation, Verification, and UQ in the Literature

3.1 Validation, Verification, and UQ in Model Development Process

[76] introduces basic steps in the modeling process, which includes (1) describing problem, (2) isolating system, (3) adopting supporting theory, (4) formulating model, (5) analyzing data requirements, collecting data, (6) developing computer program, (7) debugging computer program, (8) developing alternative solutions, (9) evaluating model output/results, (10) presenting results/plans, (11) developing model maintenance procedures, and (12) transferring system to users. According to [76], steps (1)–(7) are covered by model verification, and steps (1)–(11) are covered by model validation. Yet, the absence of the necessary information often makes it hard to follow the steps to validate models. [15] adopts this modeling process which aims at indicating how "the research community concerned policy models is attempting to develop and test procedures for improving the role of models as decision aids."

Over the past three decades, new dimensions have been brought to the notion of model V&V by large-scale computer-based mathematical and simulation models [30]. The "Sargent Circle" in simulation validation is one of the earliest and most influential among all the paradigms of the relationships among V&V activities [56].

In the Sargent Circle [56], conceptual model validation is defined as "determining that the theories and assumptions underlying the conceptual model are correct and that the model representation of the problem entity is reasonable for the intended purpose of the model"; computerized model verification is defined as "assuring that the computer programming and implementation of the conceptual model is correct"; and operational validation is defined as "determining that the model's output behavior has sufficient accuracy for the model's intended purpose over the domain of the model's intended applicability." Most validation takes place in operational validation. In order to obtain a high degree of confidence in the model and its results, the model developers need to compare the input-output behaviors of the model and the system. There are three basic comparison approaches: (1) graphs of the model and system behavior data, (2) confidence intervals, and (3) hypothesis tests. For details of the methodologies, see [2, 4, 31, 33].

A detailed schematic of the model V&V in model development process is given by [71]. There are two branches in the procedure of model V&V: the first is to obtain relevant and high-quality experimental data via physical testing; and the second is to develop and exercise the model.

Because of inherent randomness, uncertainties cannot be ignored. Uncertainty quantification (UQ) exists in the processes of both performing experiments and developing models, which emphasizes the importance of UQ in improving confidence in both the experiment and model outcomes.

By comparison with experimental data, validation is able to quantify the confidence in the predictive capability of the model. For more definitions, uncertainties, and model explanations, see [1, 45, 52, 53, 70, 71].

3.2 Quantitative Model Validation Techniques

With the development of computing capacity, computational models become extensively used to solve practical problems in various disciplines and play an important role as predictive models for complex systems. Imprecise data and model assumptions could impact the quality of the model prediction. It is important to quantify the uncertainty in the model prediction [55]. Although qualitative validation methods such as graphical comparison between model prediction and experimental data are widely used, statistics-based quantitative methods are essential to systematically account for the uncertainty in both model prediction and experimental observation [43]. [21] claims that the ability to quantify uncertainty is essential for the success of any model validation.

Many previous papers have studied the application of statistical hypothesis testing methods in the context of model validation [22, 51], as well as the validation metrics, which provide quantitative measures of agreement between a predictive model and physical/experimental observations [13, 33, 42]. Yet there remain some unclear issues in the practice of model validation. [32] studies the quantitative model validation techniques from the perspectives of both hypothesis testing-based and non-hypothesis testing-based methods and gives a systematic procedure for quantitative model validation. The quantitative model validation techniques include

classical hypothesis testing, Bayesian hypothesis testing, confidence intervals, reliability-based metric, and area metric-based method. For details and examples of these quantitative model validation techniques, see [13, 26, 32].

4 Validation for Intelligent Adversary Models

4.1 Difficulties in Validating Intelligent Adversary Models

Model validation, together with verification, is critical for models intended to be used in practice. They are required by many organizations such as the US Department of Defense that uses adversary models [18]. After the attacks on September 11, 2001, billions of dollars have been spent on homeland security. To better understand intelligent adversary behaviors and to better study the strategic interactions between defenders and adversaries (e.g., attackers and terrorists), numerous models have been developed. Unfortunately, because of the deficiency of empirical data, few (if any) such models have yet been validated, which limits the application of those models in practice.

In defense and homeland security, decisions are made about allocating resources to prevent attacks by adversaries and protect the public. Risks from such intelligent adversaries, like terrorists, must be assessed prior to guiding defensive resource allocation; otherwise the effectiveness of resource allocation would be unreliable. Intelligent adversary risk assessment aims to prevent adversary attacks or mitigate the effects of adversary attacks by allocating resource efficiently. The risk assessment has significant importance in protecting the public safety. When a model is adopted to predict the adversaries' behavior and to instruct resource allocations, decision-makers should be confident that the model adequately represents the real situation. Poor risk assessment could lead to ineffective resource allocations and vulnerable targets.

Intelligent adversary risk assessment models are increasingly being developed and studied. Those models need to be validated. Probabilistic risk assessment (PRA)/ event-tree-based methods [12,74], decision-analytic methods [14,50], game-theoretic methods [20, 28, 79], and statistical machine-learning methods [11, 35] have been proposed for modeling intelligent adversaries. However, these models are complex and may not be directly tested by comparing model predictions with the outcome of events in the real world, since there are too few comparable adversary attack data to support statistical inferences about the model validity. The Committee on Methodological Improvements to the Department of Homeland Security's Biological Agent Risk Analysis cautions that "there may be insufficient scientific knowledge to verify or validate these models" [41]. When modeling the intelligent adversary, the modeler needs to consider the strategy and the rationality of the adversaries; in addition, it is difficult to model the consequences resulting from terrorist incidents, since it is hard to assess when, where, and how the terrorists would strike. Meanwhile, as technologies evolve, adaptive terrorists could mount new types of attacks. In the reality of counterterrorism, there are

lots of uncertainties; and throughout the built models, there are assumptions and dubious parameters. Unless the models are validated, the model results may not be trustworthy.

It is possible to make empirical observations to compare with the prediction results from the corresponding models. For example, [75] proposes a prospect theory model of coaches's utility and estimates the models' parameters using the data from the 2009 NFL season; [19] studies parking choice models, which are first calibrated based on the collected data from video-recorded observations from a set of parking lots on the University at Buffalo north campus and then used to predict the drivers' behavior. Intelligent adversaries are more difficult to model, since they are adaptive, and may have unknown preferences, beliefs, and capabilities. [17] indicates that "the key difference between risk assessment for situations with intelligent adversaries and traditional risk assessment problems is that intelligent adversaries adapt. They adapt to observed, perceived, and imputed likely future actions by those defending the system they are attempting to damage. This adaptive behavior must be considered if risk assessment models are to provide accurate estimates of future risk from intelligent adversaries and appropriately support risk management decision making."

Also, [3] claims that adversary risk analysis has three special uncertainties: (1) aleatory uncertainty (randomness of outcomes), (2) epistemic uncertainty (strategic choices of an intelligent adversary), and (3) concept uncertainty (beliefs about how the problems are framed).

Validating models for the probable behaviors of intelligent adversaries may not mean comparing the model results with existing data or experimental data as is in validating traditional models. This is because the data, if any, is often incomplete and sometimes classified. In terms of validating counter-terrorism models, [65] states that "for terrorist acts, validation is only possible in a limited sense and may be more correctly characterized as ensuring the models are reasonable or credible, performing sanity checks, ensuring consistency with what is known about terrorist groups, and not being able to invalidate the model. Validity, in this case is viewed as a range, not a binary valid/invalid assessment."

4.2 Verification and Validation Methods for Intelligent Adversary Models

4.2.1 Basic Necessary Conditions for Intelligent Adversary Risk Analysis

[17] proposes four basic necessary conditions for intelligent adversary risk analysis:

1. Adversary models must be descriptively accurate representations of future adversary actions to the best of the then-current knowledge of the defender;
2. Adversary models must be computationally tractable to support risk management decisions in the particular situations being addressed;

3. Adversary models must explicitly address uncertainty and represent any uncertainty in the predicted adversary actions;
4. There must exist one or more defensible methods for gaining confidence in the models for practical use.

The methods for gaining confidence in [17] include "validation by common sense" and "limited case-based validation for historic situation when data is available."

4.2.2 Conceptual Processes to Provide Increased Confidence in Intelligent Adversary Models

[65] presents three conceptual processes that may be used to provide increased confidence in intelligent adversary models, but may fall short of traditional validation:

1. Minimally required components, which include terrorist objectives, attack logistics, decision criteria, and adaptation. Models that fail to address these key factors are considered less credible or valid;
2. Use of analogy. To obtain the real data could be hard or even impossible. However, we may infer something about the likely future behaviors of intelligent adversaries through appropriate use of existing databases, and historical attacks may be helpful in validating current models. Validating through analogy rather than using direct data is a reasonable approach. "Adversaries that are influenced by bias or have philosophical or religious perspectives will apply those perspectives to all their planning."
3. Use of uncertainty. In modeling intelligent adversaries, we should consider the uncertainties in the structure and/or parameters. Parameter values should be evaluated and measured properly.

4.2.3 Transparent Risk Assessment to Improve Confidence

"Risk assessment transparency improves confidence" is suggested by [41], where the bioterrorism risk assessment (BTRA) model [10] is reviewed. The current use of the word "transparency" is summarized by [48] as "letting the truth be available for others to see if they so choose, or perhaps think to look, or have the time, means, and skills to look" and involving "active disclosure."

In establishing confidence and trust in the methods of outputs from risk assessment models, transparency is a major factor. Achieving transparency requires the assumptions, model's mathematical and structural foundations, and the sources of data used in the analysis to be made explicit. In [41], it is emphasized that "the accuracy of quantitative bioterrorism risk assessment models and the confidence placed in them depend on the validity of the assumptions and the availability of sound data for each of the biological agentsbeing analyzed."

4.2.4 Importance of Sensitivity Analysis for Validation

[41] suggests that "sensitivity analysis is important for validation." [54] defines sensitivity analysis as the determination of how "uncertainty in the output of a model

(numerical or otherwise) can be apportioned to different sources of uncertainty in the model input."

There exist a lot of uncertainties or even errors in the model variables and parameters. The decision- maker needs to know how the uncertainties would impact the outputs of the model and therefore the confidence in the model. Many researchers have used sensitivity analysis to test uncertainties and evaluate the validity of the proposed models [60, 61, 80]. Sensitivity analysis has become an important approach to the testing and validation of risk assessment models of complex systems [5].

Recent studies [60, 61, 81] use sensitivity analysis for risk assessment to show how the results of a certain strategy would change when the parameter values change. In the future, it would be important to use more sensitivity analysis in risk management. This would be helpful to see, for example, what countermeasure strategy would be adopted if the modeler knew more about the intelligent adversaries' behavior.

4.2.5 Comparing Models to Obtain Validity

Many models have been developed to study intelligent adversaries, and there are many variants and examples of these models in the literature [27,60,61,66,79,81]. If some model has been proved to be valid, comparing other models with the validated model may be an effective way to do model verification and validation.

In [38], a comparative analysis of probabilistic risk analysis (PRA) and intelligent adversary methods for counterterrorism risk management is conducted. Defender event tree and Bayesian network, attacker event tree and Bayesian network, defender decision tree, attacker decision tree, sequential games, intelligent adversary risk analysis, adversary risk analysis, and simulation games are reviewed. [38] considers each application on the same two illustrative example decisions. With respect to risk assessment, [38] states "Defender event trees and decision trees that represent attacker decisions as probabilities estimate lower expected consequences than attacker event trees and decision trees for the highest expected consequence attack, that is, the one that the attacker would choose." As for risk communication, it is concluded "Event trees, influence diagrams with just probability nodes, and Bayesian networks with only probability nodes are all equivalent as they are following the laws of probability even though they use different solution algorithms." At last, in terms of risk management, [38] gets the same conclusion with [41] that "event trees are less useful for assessing the risk posted by intelligent, adaptive adversaries."

4.2.6 Simulation Validation with Intelligent Adversary Models

Simulation is a powerful tool for the analysis of complex process and systems. A growing number of simulation systems have been created to analyze the threats that terrorist attacks pose for public safety [34].

A National Research Council report [40] urges the Department of Homeland Security (DHS) to better validate its terrorism risk models. [39] reports the RAND's approach to validating the Risk Management Analysis Tool, or RMAT, which is

one of the Transportation Security Administration's (TSA) principal terrorism risk modeling tools developed by the TSA and Boeing Company. RMAT is one of the growing class of quantitative models which are complex and could not be validated by comparing model predictions to the outcome of events in the real world, since the reference statistical data is limited. According to [39], "RMAT simulates terrorist behavior and success in attacking vulnerabilities in the domestic commercial air transportation system, drawing on estimates of terrorist resources, capabilities, preferences, decision processes, intelligence collection, and operational planning" and "to estimate the terrorism risk-reduction benefits attributable to new and existing security programs, technologies, and procedures."

Complex simulation is used to test the validity of the RMAT model [39]. In validating the defender model in RMAT, four areas are addressed [39]:

1. Identifying and evaluating the validity of key assumptions implicit in the overall system design;
2. Comparing the world representation in RMAT to external sources;
3. Assessing the completeness of the attack scenarios considered in RMAT, including both weapon-target pairings and pathways by which attacks are carried out;
4. Comparing the attack consequences modeled in RMAT to external sources." Regarding the data, diverse forms of evidence are used to validate the data, such as "logic, subject matter expert judgments, and literature searches.

4.2.7 Using Experimental Data to Validate Intelligent Adversary Models

Validation includes comparing the model output with the real data or experimental results. In traditional models, experiments usually mimic the real situation and obtain reliable data. However, for intelligent adversary models, it is typically impossible to find data from experiments in which the conditions correspond exactly to the scenario because of the uncertain and adaptive nature of intelligent adversaries. It may also risk people's lives and public property to do some of such experiments. However, data gained based on laboratory experiments could provide insights into the behaviors of both the defender and attacker during certain hazard and emergency situations, which could be used to validate models.

An experiment was conducted in [23] to assess the extent to which individual decisions are consistent with theoretical predictions of misaligned profiling. The experiments are motivated, in part, by the counterintuitive nature of equilibrium patterns of the randomized strategies. In particular, the theory produces a paradox of misaligned profiling: in equilibrium the high reliability categories are searched more intensively, even though they are used less intensively by the terrorist organization. Field experiments with professional security officials to test these model predictions would be expensive and controversial, if possible at all. The results would be classified. Instead, [23] relies on laboratory experiments, which provide the ability to replicate and control the environment. The results of the experiment reveal behavioral patterns that are consistent with the predicted patterns. [23] provides theoretical analysis and experimental validation to guide policy makers to improve

the effectiveness of targeted profiled screening and investigates the efficient profiling andcounterterrorism policy.

4.2.8 Using Historical Data to Conduct Validation

In addition to experimental data, historical data, which may have been recorded in databases and government reports, can provide other opportunities to validate intelligent adversary models.

In terms of terrorism, there are some databases available that recorded the terrorists' attacks, such as the Global Terrorism Database (GTD) [64], the International Terrorism: Attributes of Terrorist Events (ITERATE) [25], and Terrorism in western Europe: Events Data (TWEED) [69]. Among them, the GTD records incidents from 1970 and includes both domestic and international terror incidents. LaFree and Dugan [29] states that the GTD "have by far the largest number of events than any of the other data sets." Reports from the governments, such as the Federal Emergency Management Agency (FEMA) [72], the National Research Council (NRC) [73], and US Government Accountability Office (U.S. GAO) [77], could also provide useful data to do research on intelligent adversaries.

[78] presents a class of multi-period and multi-target attacker-defender games where the attackers have multiple attacking options. The attack types considered in [78] include assassination, armed assault, bombing/explosion, facility/infrastructure attack, hijacking, hostage taking, and unarmed assault, which is summarized based on the GTD categories. The percentage of the attack types used by the attackers is shown in Fig. 1. Different attack types would impact the attack success probabilities, consequences, as well as the effectiveness of defensive resource allocation. Sequential games are studied when the defender is faced with multiple attack types and adaptive attackers and how the defender would distribute a limited amount of resources to protect multiple urban areas. The objective of the defender is to minimize the total expected loss.

Based on the historical data from the GTD and UASI (the Urban Area Security Initiative, a Department of Homeland Security grant program), the parameter values are estimated in [78], such as the economic loss, fatality loss, success probability,

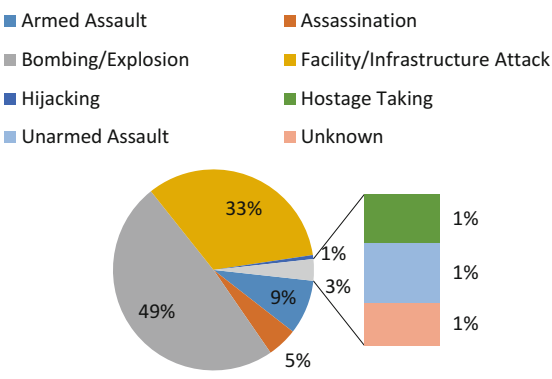

Fig. 1 Percentage of attack types in the USA (Source: [64])

and the defense cost-effectiveness for different attack types and targets. The authors estimate that basing defensive planning on the proposed model results in the lowest expected loss, having an expected loss which is 8–57 % lower than the single attack-type model and 82–96 % lower than the results of the real allocation [78].

4.3 Validate Intelligent Adversary Models Using Proxy Models

When modeling intelligent adversaries, it is important to construct a representation of preferences. However, sometimes it is difficult or even impossible to get direct elicitation from adversaries. Therefore, we could use indirect elicitation to construct and infer adversary motivations, objectives, preferences, capabilities, and beliefs. Proxy models are used by [27] to infer and validate models of adaptive adversaries. In [27], an adversary objective hierarchy and multi-attribute utility (MAU) models are constructed by proxy, using judgments from an adversary value expert (AVE). Past adversary behavior, public statements by the adversary, adversary web sites, and intelligence could be useful sources for the proxy to validate the behavior of the adversaries.

The proxy MAU models provide a relatively complete and accurate representation of the adversaries' values, including objectives, trade-offs, risk attitudes, and beliefs about consequence impacts. [27] conducts two validation studies; good convergence between the proxy model and the model assessed by direct contact is demonstrated in both cases, which indicate that the proxy model may provide insights on intelligent adversaries if constructed and implemented properly.

4.3.1 Evaluating Effectiveness of Real-World Deployments to Validate Intelligent Adversary Models

Game theory has been playing an important role in modeling adversary behaviors, and Stackelberg games have been widely used to study terrorism and are in active use for resource deployment scheduling systems by law enforcements around the USA. In a Stackelberg game, there are two players, a leader (defender) and a follower (attacker); the leader chooses a strategy first and the follower subsequently decides his own strategy after observing the leader's strategy. According to [62, 66, 67], "the Stackelberg games models have been used to assist the LAX Airport police in scheduling airport entrance checkpoints and canine patrols of the terminals, the Federal Air Marshals Service (FAMS) to schedule marshals on international flights, the United States Coast Guard (USCG) in scheduling patrols around Boston Harbor, Ports of NY/NJ, and Ports of LA/LB, the Los Angeles Sheriff's Department (LASD) for patrolling of the Metro trains, and (in-discussion) the patrolling of the Gulf of Mexico for illegal fishing for the USCG." This system has been expanded to all ports in the USA due to the success of the patrolling schedules [67].

Despite the fact that Stackelberg game-based applications have been deployed in practice, measuring the effectiveness of the applications remains a difficult problem. And the data available about the deterrence of real-world terrorist attacks is very

limited. [67] suggests several methods to evaluate the effectiveness of Stackelberg games in real-world deployments, including:

1. Computer simulations of checkpoints and canine patrols;
2. Tests against human subjects, including USC students, an Israeli intelligence unit, and on the Internet Amazon Turk site (which provides some insights into adversary bounded rationality);
3. comparative analysis of predictability of schedules and methodologies before and after implementation of a Stackelberg strategy;
4. Red team/adversary team;
5. Capture rates of guns, drugs, outstanding arrest warrants, and fare evaders;
6. User testimonials.

4.3.2 Validation With Subject Matter Experts

According to [9], the Department of Defense's Modeling and Simulation Coordination Office defines a subject matter expert (SME) as "an individual who, by virtue of position, education, training, or experience, is expected to have a greater-than-normal expertise or insight relative to a particular technical or operational discipline, system or process." However, [68] shows that many SMEs make very poor predictions "at predicting elections, wars, economic collapses, and other events" and are not accountable enough for the accuracy of the forecasts, but the forecasting skills could be improved through learning and practicing.

Experts may offer help to do subjective (and possibly mistaken) model V&V. Experts may have a relatively high level of knowledge about what, when, where, and how the intelligent adversaries may behave. In general, experts may have better knowledge or more plausible-sounding guesses and narratives about the variables and parameters than others and know where the potential uncertainties may exist. Using experts to do model verification and validation is a qualitative technique, and the judgments made by an expert may be subjective and error prone. Also, each expert may have a different understanding of the model and could use different approaches to validate the same model. The benefit of using SME is that the decision-maker may gain different perspectives on the model and would have a comprehensive idea of the situation being modeled.

[34] discusses the validation of a counterterrorism simulation of improvised explosive device (IED) incidents using the SME and concludes "it important to use the expertise of domain experts not only to compare the simulations to previous attacks of which they have knowledge, but also to use their knowledge to create new scenarios that explore the ways in which terrorist attacks could evolve."

Furthermore, the review process of the bioterrorism risk assessment (BTRA) model [10] is an example of using subject matter experts to verify and validate a model. Many good recommendations are given to make the model better [41], such as "The Department of Homeland Security should use an explicit risk analysis lexicon for defining each technical term appearing in its reports and presentations," and "To assess the probabilities of terrorist decisions, DHS should use elicitation

techniques and decision-oriented models that explicitly recognize terrorists as intelligent adversaries who observe U.S. defensive preparations and seek to maximize the achievement of their own objectives." [10] concludes that the BTRA is not valid and suggests the DHS not to continue the development of that model.

5 Conclusion

Modeling plays an important role in guiding exploration in scientific research. This chapter has reviewed the concepts of model verification and validation (V&V), illustrated and compared model V&V in the developing process of models, and also discussed techniques for conducting a successful model V&V. Model V&V steps should be integrated with the modeling process and not be separated or treated after the model has been built. In the long run, using a validated model to support decision-making can sometimes improve decisions and make preferred outcomes more likely.

Because of inherent randomness of the systems, uncertainties in the model parameters, and the process of modeling framing, may impact the accuracy of results. Uncertainty quantification (UQ) should be considered in both the modeling process and the experiment process. Having a better understanding of the uncertainties can help the modeler build a more accurate model and thus make the model more effective in practice. This chapter has also illustrated some quantitative model validation techniques in dealing with uncertainty and deciding whether or not to accept the model prediction.

Intelligent adversary model V&V and UQ are different from V&V and UQ for traditional models in literature, where experimental/physical data could be obtained to compare with the model results. Unlike traditional models, the behavior of intelligent adversaries cannot be fully understood, and due to the lack of data and incomplete information about intelligent adversaries' behavior, it is often the case that neither models nor experimental studies of adversarial behavior can be truly validated. This is a new and challenging area in the literature of model V&V. The model V&V techniques that have been discussed in this chapter include basic necessary conditions for intelligent adversary risk analysis, conceptual processes to provide more confidence in intelligent adversary models, risk assessment transparency, comparing models, simulation, experimental data, historical data, proxy models, evaluating effectiveness of real-world deployments, and using subject matter experts. These techniques attempt to take the uncertainties and adaptiveness of the intelligent adversaries into account, which may be helpful in tackling the dilemma of validating intelligent adversary models.

Many intelligent adversary models have appeared in the literature recently, but research on model V&V and UQ with intelligent adversaries is in many ways still in its infancy, with difficult challenges and limited options for overcoming them. More accurate adversarial models and more sophisticated V&V and UQ procedures with respect to adversarial models should be addressed to better understand the risks from intelligent adversaries and to better assist in surveillance and decision-making.

Acknowledgements This research was partially supported by the United States Department of Homeland Security (DHS) through the National Center for Risk and Economic Analysis of Terrorism Events (CREATE) under award number 2010-ST-061-RE0001. This research was also partially supported by the United States National Science Foundation under award numbers 1200899 and 1334930. However, any opinions, findings, and conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the DHS, CREATE, or NSF. The authors assume responsibility for any errors.

References

1. AIAA: AIAA guide for the verification and validation of computational fluid dynamics simulation. AIAA-G-077-1998, Reston (1998)
2. Balci, O., Sargent, R.G.: A Methodology for cost-risk analysis in the statistical validation of simulation models. Commun. ACM. **24**(4), 190–197 (1981)
3. Banks, D.: Adversarial Risk Analysis: Principles and Practice. Presentation on First Conference on Validating Models of Adversary Behaviors, Buffalo (2013)
4. Banks, J., Carson II J.S., Nelson, B.L.: Discrete-Event System Simulation, 2nd edn. Prentice Hall International, London, UK (1996)
5. Borgonovo, E.: Measuring uncertainty importance: investigation and comparison of alternative approaches. Risk Anal. **20**(5), 1349–1361 (2006)
6. Coleman, H.W., Steele, W.G.: Experimentation, Validation, and Uncertainty Analysis for Engineers. Wiley, Hoboken (2009)
7. DoD: DoD directive No 5000.59: Modeling and Simulation (M&S) Management. Defense Modeling and Simulation Office, Office of the Director of Defense Research and Engineering (1994)
8. DoD: Verification, Validation, and Accreditation (VV&A) Recommended Practices Guide. Defense Modeling and Simulation Office, Office of the Director of Defense Research and Engineering (1996)
9. DoD: Special Topic on "Subject Matter Experts and Validation, Verification and Accreditation", DoD Recommended Practices Guide (RPG) for Modeling and Simulation VV&A, Millennium Edition (2000)
10. DHS: Department of Homeland Security Bioterrorism Risk Assessment: A Call for Change. Available at http://www.nap.edu/catalog/12206.html (2006). Accessed in Nov 2015
11. Elovici, Y., Kandel, A., Last, M., Shapira, B. Zaafrany, O.: Using Data mining Techniques for Detecting Terror-Related Activities on the Web. Available at http://www.ise.bgu.ac.il/faculty/mlast/papers/JIW_Paper.pdf. Accessed in Nov 2015
12. Ezell, B.C., Bennett, S.P., Winterfeldt, D., Sokolowski, J, Collins, A.J.: Probabilistic risk analysis and terrorism risk. Risk Anal. **30**(4), 575–589 (2010)
13. Ferson, S., Oberkampf, W.: Validation of imprecise probability models. Int. J. Reliab. Saf. **3**(1), 3–22 (2009)
14. Garrick, B.J., Hall, J.E., Kilger, M., McDonald, J.C., O'Toole, T., Probst, P.S., Parker, E.R., Rosenthal, R., Trivelpiece, A.W., Arsdale, L.V., Zebroski, E.L.: Confronting the risks of terrorism: making the right decisions. Reliab. Eng. Syst. Saf. **86**(2), 129–176 (2004)
15. Gass, S.I.: Decision-aiding models: validation, assessment, and related issues for policy analysis. Oper. Res. **31**(4), 603–631(1983)
16. Gruhl, J., Gruhl, H.: Methods and Examples of Model Validation-an Annotated Bibliography. MIT Energy Laboratory Working Paper MIT-EL 78-022WP (1978)
17. Guikema, S.: Modeling intelligent adversaries for terrorism risk assessment: some necessary conditions for adversary models. Risk Anal. **32**(7), 1117–1121 (2012)
18. Guikema, S., Reilly, A.: Perspectives on Validation of Terrorism Risk Analysis Models. Presentation on First Conference on Validating Models of Adversary Behaviors, Buffalo (2013)

19. Guo, L., Huang, S., Zhuang, J.: Modeling parking behavior under uncertainty: a static game theoretic versus a sequential neo-additive capacity modeling approach. Netw. Spat. Econ. **13**(3), 327–350(2013)
20. Hausken, K., Zhuang, J.: The impact of disaster on the interaction between company and government. Eur. J. Oper. Res. **225**(2), 363–376(2013)
21. Hemez, F.M., Doebling, S.W.: Model validation and uncertainty quantification. For publication in the proceeding of IMAC-XIX, the 19th International Model Analysis Conference, Kissimmee, 5–8 Feb 2001
22. Hills, R.G., Leslie, I.H.: Statistical validation of engineering and scientific models: validation experiments to application. Sandia Technical Report (SAND2003-0706) (2003)
23. Holt, C.A., Kydd, A., Razzolini, L., Sheremeta, R.: The Paradox of Misaligned Profiling: Theory and Experimental Evidence. Available at http://www.people.vcu.edu/~lrazzolini/Profiling.pdf (2014). Accessed in Nov 2015
24. IEEE: IEEE Standard Glossary of Software Engineering Terminology. IEEE Std 610.12-1990, New York (1991)
25. International Terrorism: Attributes of Terrorist Events (ITERATE). Available at http://library.duke.edu/data/collections/iterate. Accessed in Nov 2015
26. Jiang, X., Mahadevan, S.: Bayesian risk-based decision method for model validation under uncertainty. Reliab. Eng. Syst. Saf. **92**(6), 707–718 (2007)
27. John, R., Rosoff, H.: Validation of Proxy Random Utility Models for Adaptive Adversaries. Available at http://psam12.org/proceedings/paper/paper_437_1.pdf (2014). Accessed in November, 2015
28. Jose, V.R.R., Zhuang, J.: Technology Adoption, Accumulation, and Competition in Multiperiod Attacker-Defender Games. Mil. Oper. Res. **18**(2), 33–47 (2013)
29. LaFree, G., Dugan, L.L.: Introducing the global terrorism database. Terror. Political Violence **19**(2), 181–204 (2007)
30. Landry, M., Malouin, J.L., Oral. M.: Model validation in operations research. Eur. J. Oper. Res. **14**(3), 207–220 (1983)
31. Law, A.M., Kelton, W.D.: Simulation Modeling and Analysis, 2nd edn. McGraw-Hill, New York (1991)
32. Ling, Y., Mahadevan, S.: Quantitative model validation techniques: new insights. Reliab. Eng. Syst. Saf. **111**, 217–231 (2013)
33. Liu, Y., Chen, W., Arendt, P., Huang, H.: Toward a better understanding of model validation metrics. J. Mech. Des. **133**(7), 1–13(2011)
34. Louisa, N., Johnson, C.W.: Validation of Counter-terrorism Simulation Models. Available at http://www.dcs.gla.ac.uk/~louisa/Publications_files/ISSC09_Paper_2.pdf (2009). Accessed in Nov 2015
35. Mason, R., McInnis, B., Dalal, S.: Machine Learning for the Automatic Identification of Terrorist Incidents in Worldwide News Media. In: 2012 IEEE International Conference on Intelligence and Security Informatics (ISI), Washington, DC, pp. 84–89 (2012)
36. McCarl, B.A.: Model validation: an overview with some emphasis on risk models. Rev. Market. Agric. Econ. **52**(3), 153–173 (1984)
37. McCarl, B.A., Spreen, T.H.: Validation of Programming Models. Available at http://agecon2.tamu.edu/people/faculty/mccarl-bruce/mccspr/new18.pdf (1997). Accessed in Nov 2015
38. Merrick, J., Parnell, G.S.: A comparative analysis of PRA and intelligent adversary methods for counterterrorism risk management. Risk Anal. **31**(9), 1488–1510 (2011)
39. Morral, A.R., Price, C.C., Ortiz, D.S., Wilson, B., LaTourrette, T., Mobley, B.W., McKay, S., Willis, H.H.: Modeling Terrorism Risk to the Air Transportation System: An Independent Assessment of TSA's Risk Management Analysis Tool and Associated Methods. RAND report. Available at http://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND_MG1241.pdf (2012). Accessed in Nov 2015
40. NRC: Review of the Department of Homeland Security's Approach to Risk Analysis. Available at https://www.fema.gov/pdf/government/grant/2011/fy11_hsgp_risk.pdf (2010). Accessed in Nov 2015

41. NRC: Bioterrorism Risk Assessment. Biological Threat Characterization Center of the National Biodefense Analysis and Countermeasures Center. Fort Detrick, MD (2008)
42. Oberkampf, W., Barone, M.: Measures of agreement between computation and experiment: validation metrics. J. Comput. Phys. **217**(1), 5–36 (2006)
43. Oberkampf, W., Trucano, T.: Verification and validation in computational fluid dynamics. Progr. Aerosp. Sci. **38**(3), 209–272 (2002)
44. Oberkampf, W.L.: Bibliography for Verification and Validation in Computational Simulation. Sandia Report (1998)
45. Oberkampf, W.L., Trucano, T.G., Hirsch, C.: Verification, validation, and predictive capability in computational engineering and physics. Appl. Mech. Rev. **57**(5), 345–384 (2004)
46. Oden, J.T.: A Brief View of Verification, Validation, and Uncertainty Quantification. Available at http://users.ices.utexas.edu/~serge/WebMMM/Talks/Oden-VVUQ-032610.pdf (2009). Accessed in Nov 2015
47. O'Keefe, R.M., O'Leary, D.E.: Expert system verification and validation: a survey and tutorial. Artif. Intell. Rev. **7**, 3–42 (1993)
48. Oliver, R.W.: What Is Transparency? McGraw-Hill, New York (2004)
49. Pace, D.K.: Modeling and simulation verification and validation challenges. Johns Hopkins APL Technical Digest. **25**(2), 163–172 (2004)
50. Rakesh, K., Sarin, L. Keller, R.: From the editors: probability approximations, anti-terrorism strategy, and bull's-eye display for performance feedback. Decis. Anal. **10**(1), 1–5(2013)
51. Rebba, R., Mahadevan, S.: Validation of models with multivariate output. Reliab. Eng. Syst. Saf. **91**(8), 861–871 (2006)
52. Roach, P.J.: Verification and Validation in Computational Science and Engineering. Hermosa Publishers, Albuquerque (1998)
53. Salari, K., Knupp, P.: Code Verification by the Method of Manufactured Solutions. Sandia National Laboratories, SAND2000-1444 (2000)
54. Saltelli, A., Tarantola, S.: On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal. J. Am. Stat. Assoc. **97**(459), 702–709 (2002)
55. Sankararaman, S., Mahadevan, S.: Model validation under epistemic uncertainty. Reliab. Eng. Syst. Saf. **96**(9), 1232–1241(2011)
56. Sargent, R.G.: An assessment procedure and a set of criteria for use in the evaluation of computerized models and computer-based modeling tools. Final technical report RADC-TR-80-409, U.S. Air Force (1981)
57. Sargent, R.G.: Some subjective validation methods using graphical displays of data. In: Proceedings of the 1996 Winter Simulation Conference, Coronado, California (1996)
58. Sargent, R.G.: Verification and validation of simulation models. In: Proceedings of the 2009 Winter Simulation Conference, Austin, Texas, pp. 162–176 (2009)
59. Schlesinger, S., Crosbie, R.E., Innis, G.S., Lalwani, C.S., Loch, J., Sylvester, R.J., Wright, R.D., Kheir, N., Bartos, D.: Terminology for model credibility. Simulation **32**(3), 103–104 (1979)
60. Shan, X., Zhuang, J.: Cost of equity in homeland security resource allocation in the face of a strategic attacker. Risk Anal. **33**(6), 1083–1099 (2013)
61. Shan, X., Zhuang, J.: Hybrid defensive resource allocations in the face of partially strategic attackers in a sequential defender-attacker game. Eur. J. Oper. Res. **228**(1), 262–272 (2013)
62. Shieh, E., An, B., Yang, R., Tambe, M., Baldwin, C., DiRenzo, J., Maule, B. , Meyer, G.: PROTECT: a deployed game theoretic system to protect the ports of the United States. In: AAMAS, Valencia, Spain (2012)
63. Sornette, D., Davis, A.B., Vixie, K.R.,Pisarenko, V., Kamm, J.R.: Algorithm for model validation: theory and applications. Proc. Natl. Acad. Sci. U. S. A. **104**(16), 6562–6567 (2007)
64. START: Global Terrorism Database[data file]. Available at http://www.start.umd.edu/gtd. Accessed in Nov 2015
65. Streetman, S.: The Art of the Possible in Validating Models of Adversary Behavior for Extreme Terrorist Acts. Presentation on First Conference on Validating Models of Adversary Behaviors, Buffalo (2013)

66. Tambe, M.: Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned. Cambridge University Press, New York (2011)
67. Tambe, M., Shieh, E.: Stackelberg Games in Security Domains: Evaluating Effectiveness of Real-World Deployments. Presentation on First Conference on Validating Models of Adversary Behaviors, Buffalo (2013)
68. Tetlock, P.E., Gardner, D.: Superforecasting: The Art and Science of Prediction. Crown, New York (2015)
69. Terrorism in Western Europe: Events Data (TWEED). Available at http://folk.uib.no/sspje/tweed.htm. Accessed in Nov 2015
70. Thacker, B.H., Riha, D.S., Millwater, H.R., Enright, M.P.: Errors and uncertainties in probabilistic engineering analysis. In: Proceedings of the 42nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference and Exhibit, Seattle, Washington (2001)
71. Thacker, B.H., Doebling, S.W., Hemez, f. M., Anderson, M.C., Pepin, J.E., Rodriguez, E.A.: Concepts of Model Verification and validation. Available at http://www.ltas-vis.ulg.ac.be/cmsms/uploads/File/LosAlamos_VerificationValidation.pdf (2004). Accessed in Nov 2015
72. The Federal Emergency Management Agency (FEMA). Available at http://www.fema.gov/. Accessed in Nov 2015
73. The National Research Council (NRC). Available at http://www.nationalacademies.org/nrc/. Accessed in Nov 2015
74. Toubaline, S., Borrion, H., Sage, L.T.: Dynamic generation of event trees for risk modeling of terrorist attacks. In: 2012 IEEE Conference on Technologies for Homeland Security (HST), Waltham, MA, pp. 111–116 (2012)
75. Urschel, J., J. Zhuang.: Are NFL coaches risk and loss averse? Evidence from their use of kickoff strategies. J. Quant. Anal. Sports **7**(3), Article 14(2011)
76. U.S. GAO: Guidelines for Model Evaluation. PAD-79-17, Washington, DC (1979)
77. U.S. Government Accountability Office (U.S. GAO). Available at http://www.gao.gov/. Accessed in Nov 2015
78. Zhang, J., Zhuang, J.: Modeling a Multi-period, Multi-target Attacker-defender Game with Multiple attack types. Working paper (2015)
79. Zhang, J., Zhuang, J.: Defending Remote Border Security with Sensors and UAVs based on Network Interdiction Methods. Working paper (2015)
80. Zhuang, J., Bier, V.: Balancing terrorism and natural disasters-defensive strategy with endogenous attacker effort. Oper. Res. **55**(5), 976–991(2007)
81. Zhuang, J., Saxton, G., Wu, H.: Publicity vs. impact in nonprofit disclosures and donor preferences: a sequential game with one nonprofit organization and N donors. Ann. Oper. Res. **221**(1), 469–491(2014)