Chapter 8
Verification and Validation for CIPRNet

Jeroen Voogd

Abstract In this chapter it is shown that if an appreciable risk is present in the use of Modelling and Simulation (M&S), Verification and Validation (V&V) should be employed to manage and mitigate that risk. The use of M&S in the domain of critical infrastructure (CI) will always be accompanied by such a risk. It is important that a structured approach to V&V is used in order to be more effective and more efficient than just testing without a clear plan. The Generic Methodology for V&V (GM-VV) is a recommended practise in the international M&S community and adopted by large organisations such as NATO. The GM-VV has a number of concepts that allow for a structured approach to V&V. A structured approach to V&V such as the GM-VV leads to a set of handles that allow the best choices for V&V techniques to employ. The choice for a specific technique is dependent on a number of factors such as the needed certainty, the expected residual uncertainty of the proposed technique and its requirements in terms of costs, real-world knowledge, etc. This chapter is divided in 4 parts. The first part has the take away message "You have to do Verification and Validation because there is risk involved", the second "You have to do it in a structured way if you want to do it more effective and more efficient" and the third "You have to choose the appropriate Verification and Validation technique to balance risk, effectiveness and efficiency." In the last part some conclusions are drawn.

1 Do V&V If There Is Risk Involved

In this section we first briefly explain what Modelling and Simulation (M&S) are. It will be made clear that if the M&S results are applied in the real world, M&S Use Risk has to be considered. To manage this risk it is required to have insight into the

J. Voogd (✉)
Modelling, Simulation and Gaming, TNO, Oude Waalsdorperweg 63, PO Box 96864, 2597 AK The Hague, The Netherlands
e-mail: jeroen.voogd@tno.nl

quality and associated risk of the M&S system over its entire life cycle. Verification and Validation are the two processes to obtain this insight. These processes are also briefly explained.

1.1 Modelling and Simulation

Modelling and simulation start—as all system engineering projects do—with a purpose. Then the modelling starts. A possible definition of a model is that it is an abstract representation or specification of a system. A model can represent a system that exists in our material world but can also represent not yet existing systems or combinations thereof. That part of (the imagined) reality that the model is supposed to represent is called the *simuland*. Then further abstractions are applied to the simuland in order to make the model suited for its purpose. Abstraction in this context is a process in which a relative sparse set of relevant (sub)systems, relationships and their inherent qualities are extracted or separated from the more complex (imagined) reality (Fig. 1).

In a simulation the model is used to replicate the simuland behaviour. Thus a simulation is a method, software framework or system to implement and evaluate a model over time i.e., it is a system in which a model is made to execute and is exercised. This model in its executable form is called the M&S system.

The M&S system is provided with input and its output is used within a certain context provided by a frame such as shown in Fig. 2 which is called the Simulation Frame. The model that is executed in the simulation is controlled and observed by means of its ports (ellipses in Fig. 2). Through these ports simulation data, stimuli or settings are entered into the model and simulation output leaving the executed model is observed. During the simulation the model behaves according to a dynamics that represent the state change and behavioural properties of the simuland. The notion of time, behavioural representation and frame are fundamental characteristics of a simulation.

To properly replicate the simuland for the intended use, the model is configured, controlled and stimulated by the Simulation Frame by means of input trajectories,
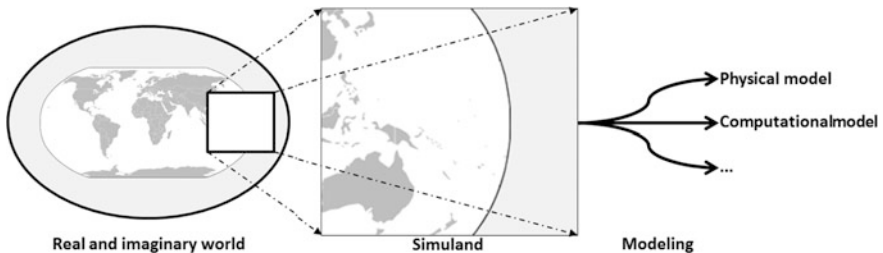

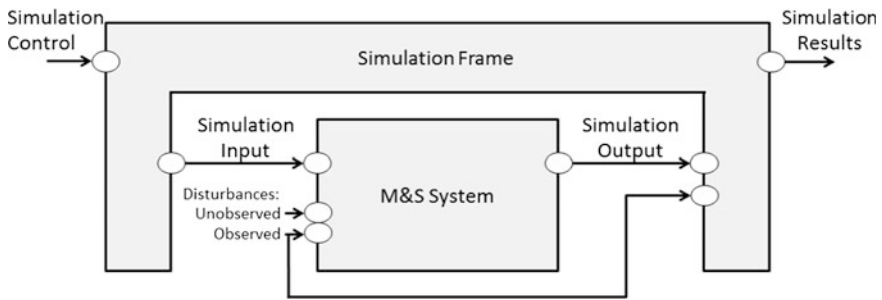**Fig. 1** Modelling is taking abstractions

Fig. 2 Relation between simulation frame and the M&S system

scenario's, parameters, environment variable settings and experimental control settings. Furthermore, environment disturbances coming from connections with live entities may impact the behaviour of the M&S system. During the execution of the model, human input can become part of the displayed behaviour. This can be from trainees, but also from operators such as opponent players to provide stimuli to the trainees or Subject Matter Experts (SMEs) that interfere with the execution of the simulation for some purpose dictated by the Simulation Frame (e.g., keeping the execution within a desired regime).

So, all in all the M&S process consists of cutting away all elements of the real and imaginary world that are not needed for the purpose at hand, then apply various abstraction techniques to make the model suited for use, then the model is executed in order to obtain results (e.g. a trained operator or an optimized CI configuration). These results are then applied in some form or another in the real world.

And that last part is exactly where the risk exists. When the M&S-based solutions to problems are applied in the real world there is a risk that those results are not fully appropriate. There can be a number of causes: the purpose for the M&S endeavour was not what was ultimately needed, maybe the simuland did not contain all needed elements of the real and imaginary world, maybe some abstractions were too large and important details were abstracted away, maybe the implementation and execution of the model or the interpretation of it's results introduced errors.

If the results of M&S are never used in the real world, e.g. if it is used for entertainment purposes or as a hobby, then there is no problem. But for CI this is not the case. The possible sources of errors may for example lead to operators of actual CI taking wrong actions if M&S was used for their training. If it is used for determining the best possible configuration of interconnecting CI, it may result in a system that performs less than desired.

The conclusion is that we need to be sure that the M&S results are fit for purpose before actually applying them to the real world. There are two processes that do exactly that: Verification and Validation. Therefore the take away message of this part is "You have to do Verification and Validation because there is risk involved".

1.2 Verification and Validation

There is no true consensus on the exact definitions of what Verification and Validation (V&V) are. Some definitions are:

Verification. The process of providing evidence justifying the M&S system's correctness [1]. Confirmation, through the provision of objective evidence that specified requirements have been fulfilled [2]. The process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications [3]. The process of determining the degree that a model, simulation, or data accurately represent its conceptual description and its specifications [4].

Correctness. The extent to which an M&S system implementation conforms to its specifications and is free of design and development errors [1].

Validation. The process of providing evidence justifying the M&S system's validity [1]. Confirmation, through the provision of objective evidence that the requirements for a specific intended use or application have been fulfilled [2]. The process of determining the degree to which a model or simulation and its associated data are an accurate representation of the real-world from the perspective of the intended uses of the model [3]. The process of determining the degree to which a model, simulation, or data is an accurate representation of the real world, from the perspective of the intended purpose of the model, simulation or data [4].

Validity. The property of an M&S system's representation of the simuland to correspond sufficiently enough with the referent for the intended use [1]. The property of a model, simulation or federation of models and simulations representations being complete and correct enough for the intended use [5].

A more intuitive explanation can be seen in Fig. 3. There the blue arrows indicate verification: starting from the specification of the M&S system, a simuland is made, which, after modelling, results in an implementation that can be executed to obtain M&S results. At each step one can check if the transformation has been done correctly and the goal is to show that the M&S system adheres to the specification. In literature one often finds that verification assesses if the M&S is built and used right.

Validation, which is the red arrow in Fig. 3, on the other hand, is making sure that the M&S results produced by the M&S system are fit for the customer's needs
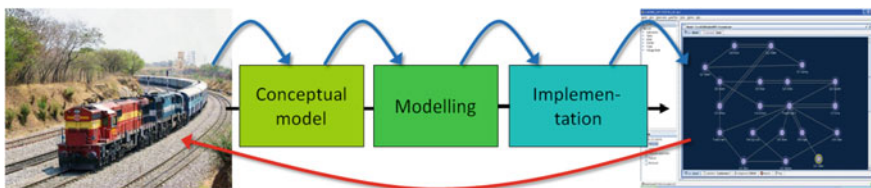

Fig. 3 Verification (*blue arrows*) and Validation (*red arrow*)

in the real world. In literature one often finds that validation assesses if the right M&S is built or procured.

During the execution of V&V it may (and usually will) happen that elements of the M&S System or its use, are found that are not correct or that contribute negatively to the customer's need. Identifying these sources of risk are necessary in order to start managing them. In short: doing V&V provides insight into and advice on the quality of the M&S system over its entire life cycle, and the associated risks.

When studying literature on V&V another term is often found: accreditation. This is, however, a somewhat problematic concept. According to [3] accreditation is "The official certification that a model or simulation and its associated data are acceptable for use for a specific purpose." One problem is that in most countries and for most application domains of M&S there is no official body that can issue such M&S certificates. And besides often the word accreditation is reserved for the official recognition that an organization is allowed to issue certificates. The official certification is called just that: certification (and not accreditation).

In this text the word accreditation or certification is not used. What is assumed is that the result of doing V&V is a body of knowledge on the quality and deficiencies and their associated risks, based on which the customer can decide whether to accept the M&S system or not.

1.3 But How to Do V&V, and How Much?

As described above, for M&S applied to CI it is necessary to identify and manage risk. V&V can be used for that but the question is how should the V&V be approached and how much effort should be spend on it.

The second question is difficult to answer because there is no general answer. Doing V&V can be costly and it should be in balance with the M&S Use Risk involved. Another aspect that has to be considered is the risk tolerance, e.g. in the form of insurance, of the user. What is most important is that the cost spent on the V&V effort should be in balance with the possible costs associated with the risk. The cost of doing V&V should also be significantly less than the possible saving due to risk reduction.

The first question—how to do V&V—is easier to answer. In practice it is often observed that those who develop the simulation also perform the V&V activities. Although they often do a good job, the result does leave something to be desired. After the V&V activities it is not clear anymore which tests were performed and why. The documentation is more often than not a bunch of files on the developer's computer. If after some time things need to be changed in the M&S system and thus some additional V&V activities have to be performed, it is not clear which of the results from the initial V&V activities are still applicable and which tests need to be redone. In short: there is no traceable path from the user's goal to the tests to the results, and no re-usable documentation exists. An unstructured approach to doing V&V may be effective, but often this cannot be shown. It may also be efficient at

first, but again it cannot be shown that the most efficient way of doing V&V tests has been chosen.

So, the question arises if there is a V&V approach that does work well. The first thing to look at is if there are appropriate standards for doing V&V. It turns out that there are a number of V&V approaches for M&S, but these are often domain specific, strongly tied to a specific technology or developer oriented. If that is what is needed, then use them. In general, however, it is not advised to use a developer-oriented approach because the link with the user's goal is not clear. If the V&V effort does not involve a specific domain or technology for which a V&V standard is available, then a more general V&V approach is required.

In order to make sure the V&V effort is effective, the starting point has to be the goal of the user, or to be more precise: the M&S Use Risk associated with the user's goal. From that starting point criteria need to be derived that show what needs to be tested. That derivation and choosing V&V techniques for doing the tests needs to be within the limits of the resources available for the V&V effort, which in practice is always rather limited. The results of the V&V effort need to be documented in such a way that all results can be traced back to the tests and the user's goal, and it should also be such that re-use at a later data is possible. In short: the V&V approach must result in the biggest bang for the buck as well as allow full traceability, otherwise serious questions can be raised about the effectiveness and efficiency of the V&V effort.

The take away message of this section is "You have to do V&V in a structured way if you want to do it more effective and more efficient".

2 Do V&V in a Structured Way to Be More Effective and Efficient

The choice of which V&V method works best in a given situation depends on the individual needs and constraints of an M&S organization, project, application domain or technology. Moreover, V&V usually requires a complex mixture of various activities, methods, tools, techniques and application domain knowledge, which are often tightly coupled with the M&S development process. Therefore, many different approaches to V&V exist that rely on a wide variety of different V&V terms, concepts, products, processes, tools or techniques. In many cases, the resulting proliferation restricts or even works against the transition of V&V results from one M&S organization, project, and technology or application domain to another. Furthermore, history shows that V&V is often more of an afterthought than a built-in part of an M&S development, employment and procurement policy.

The purpose of the Generic Methodology for V&V (GM-VV) is to address these issues by means of providing general applicable guidance for V&V that:

- Facilitates common understanding and communication of V&V within the M&S community.
- Is applicable to any phase of the M&S life-cycle (e.g., development, employment, and reuse).
- Is M&S stakeholders' acceptance decision-making process-oriented.
- Is driven by the M&S stakeholders' needs and M&S use risks tolerances.
- Is scalable to fit any M&S scope, budget, resources and use-risks thresholds.
- Is applicable to a wide variety of M&S technologies and application domains.
- Will result in traceable, reproducible and transparent evidence-based acceptance arguments.
- Can be instantiated on enterprise, project or technical levels alike.
- Facilitates reuse and interoperability of V&V outcomes, tools and techniques.

GM-VV is not aimed to replace the existing V&V approaches, methodologies, standards or policies of M&S organizations, technology and application domains; nor is GM-VV's intent to substitute common enterprise or project management practices prevalent within M&S client or supplier organizations. In addition, GM-VV is not intended to be prescriptive, in that it does not specify a single concrete or unique solution for all V&V applications. Rather, the GM-VV should be tailored to meet the needs of individual V&V applications.

The GM-VV provides a technical framework that focuses on M&S V&V practices. Though interrelated, acceptance decision processes and associated practices such as M&S accreditation and certification are outside the scope of the methodology. GM-VV attains its generic quality from a technical framework that consists of three subparts: the conceptual, implementation and tailoring framework (Fig. 4). This framework is rooted in established international standards and other related practices. The conceptual framework provides the terminology, concepts and principles to facilitate communication and a common understanding and execution of V&V within an M&S context. The implementation framework translates these concepts and principles into a set of generic building blocks to develop consistent V&V solutions for an individual M&S organization, project, and technology or application domain. GM-VV provides a tailoring framework that utilizes these building blocks to develop and cost-efficiently apply such V&V application instances. As such, the GM-VV provides a high-level framework for developing concrete V&V solutions and conducting V&V, into which lower-level practices (e.g., tools, techniques, tasks, acceptability criteria, documentation templates) native to each individual M&S organization, project, technology or application domain can easily be integrated.

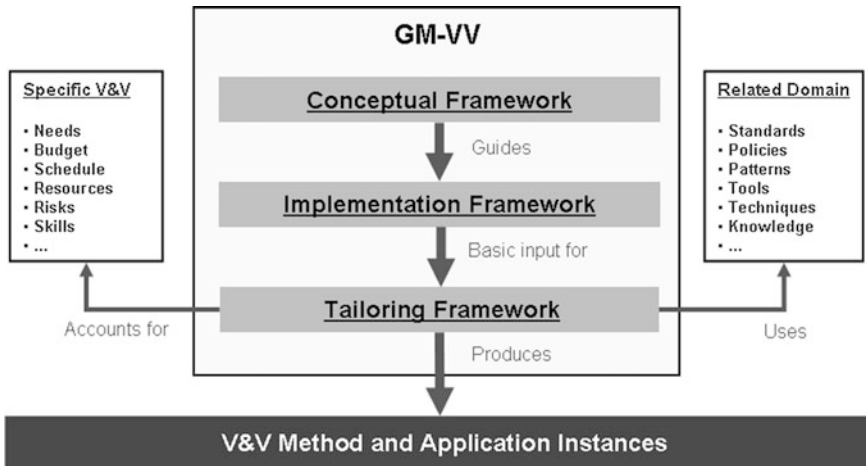Each of the three frameworks will be described in sections below.

Fig. 4 GM-VV technical framework design and operational use concept

2.1 *Conceptual Framework*

This section discusses the GM-VV conceptual framework. This framework provides fundamental and general applicable terminology, semantics, concepts and principles for V&V. The purpose of the framework is to facilitate communication, understanding and implementation of V&V across and between different M&S contexts (e.g., organizations, application domains, standards, technologies). The framework is the foundation upon which the GM-VV implementation framework rests.

2.1.1 Links to Systems Engineering

Within the GM-VV, M&S systems are considered to be systems of systems that have a lifecycle and are subject to system engineering practices. Moreover, models and simulations are considered to be part of a larger system in which they are used. From this perspective, M&S is a systems engineering specialization. V&V is an intrinsic part of the systems engineering process [6–9]. Therefore, the GM-VV considers the V&V of M&S as a specialization of systems engineering V&V. Hence, the GM-VV can be integrated with, complement or extend the V&V processes within such existing systems engineering methodologies or standards.

2.1.2 M&S-Based Problem Solving Approach

The basic premise of the GM-VV is that models and simulations are always developed and employed to fulfil the specific needs of their end users (e.g., trainers, analysts, decision makers). Modelling and simulation is thus considered to be a problem solving process that transforms a simple statement of an end user's need into an M&S-based solution for the problem implied in the need. The GM-VV assumes that V&V always takes place within such larger context. This context is abstracted by means of defining four interrelated worlds (Fig. 5). Together, these four worlds define a generic lifecycle and process view of M&S-based problem solving. A view that serves as a common basis, in which V&V for M&S (e.g., concepts, principles, processes, products, techniques) can be understood, developed or applied.

These four worlds can be described as follows:

- **Real World**: The Real World is, as the name suggests, the actual real-life world of which we are part of. It is where the need for some solution arises and where the solution is applied to obtain the desired outcomes. It is also where the real operational and other business risks exist in case the M&S based problem solution is not fit for purpose. Stakeholders from this world may for example be CI facility owners that need well trained operators as well as the general public that wishes to use these facilities and desire a stable service.
- **Problem World**: In the Problem World the needs of the Real World are further examined and solved. For some needs the problem may be training, in which case the Problem World is actually the "Training World", or if the need involves analysis it is the "Analysis World". Here the generic "Problem World" is used. The problem solution may consist of different parts, for example a training program may consist of class room training, simulator based training and live training; an analysis may consist of a literature study, simulation based analysis and expert interviews. In the Problem World the complete problem is solved. Thus the simulation based component (i.e., M&S results) may only be part of the solution.
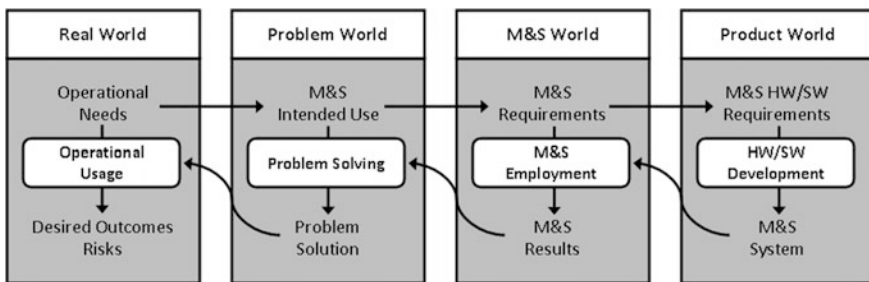

Fig. 5 Four worlds view of M&S based problem solving

Stakeholders from within the Problem World are those parties involved in the complete solution (e.g., organizations) such as education centres and their trainers in case of training, analysts in case of an analysis problem. Stakeholders from the Real World or their experts are typically also involved in the Problem World.

- **M&S World**: In the M&S World the focus is on the M&S based components of the Problem Solution. Here M&S (sub)systems are defined and used. It starts with the specified M&S intended use from the Problem World from which the requirements are derived such as the M&S System components that are needed, which scenarios are to be used and which personnel (trainers, scenario builders, etc.) are needed. After the M&S System becomes available from the "Product World" the execution takes place and the M&S Results are constructed. Stakeholders from within the M&S World are trainers, trainees, analysts or other controllers that control the simulation.

- **Product World**: The Product World takes the M&S requirements from the M&S World and determines the M&S hardware and software requirements. The M&S System is constructed and delivered to the M&S World. Stakeholders within the Product World are those organizations that build and deliver the M&S System such as programmers, model developers, system or software architects and managers of repositories with reusable models.

When the M&S problem solving process described by the four-worlds view is properly executed, the resulting solution should satisfy the originally identified needs with a minimal level of (use) risk in the Real World.

The M&S system, M&S requirements, M&S results and other development artefacts (e.g., conceptual model, software design, code) are thus always directed toward contributing to and satisfying the Real World operational needs. The degree of success of such M&S in satisfying these needs depends on how well they are specified, designed, developed, integrated, tested, used, and supported. These M&S activities require the contribution of individuals or organizations that have a vested interest in the success of the M&S asset, either directly or indirectly. An individual or organization with such interest is referred to in GM-VV as a stakeholder. Stakeholders can play one or more roles in each of the four worlds such as M&S user/sponsor, supplier, project manager, software developer, operator, customer, or subject matter expert (SME). Depending upon their role, stakeholders may hold different responsibilities in the M&S life-cycle processes, activities or tasks.

2.1.3   V&V Problem Solving Approach

Within the four-world context, stakeholders exist who are responsible for making acceptance decisions on the use of M&S. Within the GM-VV, these stakeholders are referred as V&V User/Sponsor. In this context the V&V User/Sponsor could be an M&S User/Sponsor, Accreditation Authority or any other domain specific role

that uses the outcomes of the V&V. V&V Users/Sponsors face the problem of having to make a judgment on the development and suitability of the M&S system or results for an intended use. The key issue here is that it is not possible to demonstrate with absolute certainty that the M&S system or results will meet the Real World needs prior to its actual use. Consequently, there is always a probability that the M&S-based solution is not successful when used (i.e., fails). Such a failure would result in an undesirable impact (i.e., a risk) on the operational environment. Therefore, an M&S system or result is only acceptable to the V&V User/Sponsor if he or she has sufficient confidence that the use of an M&S system or result satisfies the Real World needs without posing unacceptable risks (e.g., costs, liabilities). This M&S acceptability is something relative to different V&V Users/Sponsors: what is acceptable to one V&V User/Sponsor may not be acceptable for another. The V&V User/Sponsor's decision-making process therefore requires appropriate evidence-based arguments to justify his or her acceptance decision.

The basic premise of GM-VV is that V&V are performed to collect, generate, maintain and reason with a body of evidence in support of the V&V Users/Sponsors acceptance decision-making process. Here, validation is referred to as the process that establishes the V&V User/Sponsor's confidence as to whether or not they have built or procured the right M&S system or result for the intended use (i.e., M&S validity). In other words "Did we build the right M&S system?". To ensure that the M&S system or results at delivery can be demonstrated to be valid, it is necessary to ensure that the M&S system is built and employed in the right manner. Here verification is referred to as the process of establishing V&V User/Sponsors confidence in whether the evolving M&S system or result is built right (i.e., M&S correctness). In other words "Did we build the M&S system right?". The GM-VV considers V&V as a specific problem domain of M&S with its own needs, objectives and issues. This domain is referred to as the V&V World (Fig. 6).

The V&V world groups the products, processes and organizational aspects that are needed to develop an acceptance recommendation that can be used by the V&V User/Sponsor in his or her acceptance decision procedure(s). This recommendation included in a V&V report is the key deliverable of a V&V effort and contains evidence-based arguments regarding the acceptability of an M&S system or results. Here the GM-VV premise is that the acceptance decision itself is always the responsibility of the V&V User/Sponsor and decision procedure(s) may involve trade-off aspects beyond the V&V effort scope.

The development of an acceptance recommendation in the V&V world is driven by the V&V needs that are traceable to the V&V User/Sponsor's acceptance decision or procedure(s) needs (e.g., budget, responsibilities, risks, liabilities). Therefore, the extent, rigor and timeframe of a V&V effort depend on these needs. Depending on these needs, the V&V effort could span the whole or specific M&S lifecycle phase of the four worlds; could focus on one specific or multiple (intermediate) M&S products; and should match the development paradigm that was used (e.g., waterfall, spiral). Each case may require a separate acceptance recommendation with its own scope and development timeline. Moreover, the way the V&V effort interacts with the four M&S-based problem worlds also varies from
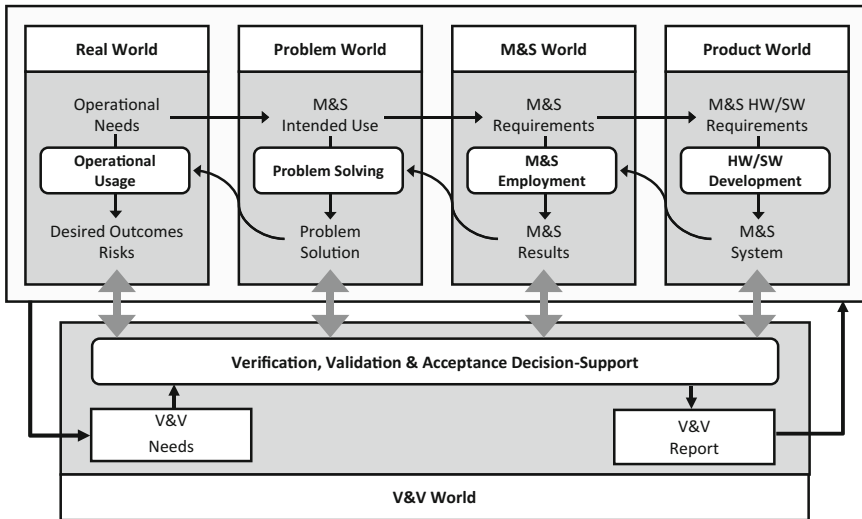
Fig. 6 V&V world and four-world interfacing

case to case. These mutual dependencies are depicted in Fig. 6 with bidirectional arrows that interface the V&V world with each of the four M&S-based problem solving worlds. Two classical types of V&V that can be identified based on the time frame of their execution are [6, 10–12]:

- **Post hoc V&V**: V&V conducted in retrospect on an M&S system after development or on M&S results after M&S system employment.
- **Concurrent V&V**: V&V conducted in prospective throughout the whole M&S life cycle to manage and improve the quality of newly developed M&S systems or results.

The GM-VV supports both V&V time frames but is not limited to these distinct types. A V&V effort can be post hoc, concurrent, iterative, recursive or even be a recurrent effort in the case where legacy M&S products are updated or reused for a different intended-use.

2.1.4 Acceptance Recommendation, Acceptability Criteria and Evidential Quality

The objective of a V&V effort is to develop evidence upon which an acceptance recommendation is based. This V&V objective is articulated as an acceptance goal. This high-level goal should be translated into a set of concrete and assessable acceptability criteria for the M&S system or result(s). Relevant and convincing
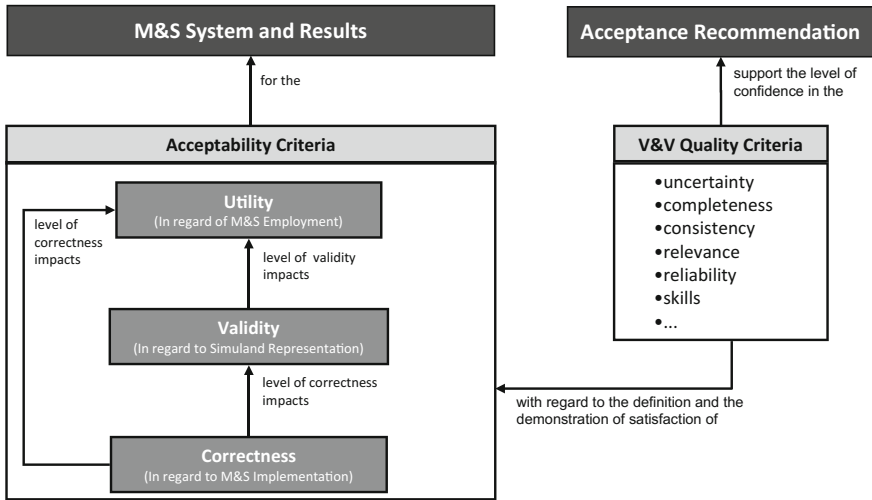
Fig. 7 Utility, validity, correctness and V&V quality criteria relationships

evidence should then be collected or generated to assess the satisfaction of these criteria. When it is convincingly demonstrated to what extent the M&S system or result(s) does or does not satisfy all these acceptability criteria, a claim can be made on whether or not the M&S system or result(s) is acceptable for its intended use (i.e., acceptance claim).

The GM-VV identifies three types of M&S properties for which acceptability criteria could be set (Fig. 7):

- **Utility**: this property refers to the extent to which the M&S system or result(s) is useful in solving the M&S user/sponsor's needs. Utility properties could comprise sub-types such as M&S value (e.g., measures of effectiveness, measures of performance), cost (e.g., money, time) and use risks (e.g., impact, ramifications).
- **Validity**: this property refers to the extent to which the M&S system's representation corresponds to the simulated simuland (i.e., system of interest) from the perspective of the intended use. The level of validity impacts the utility.
- **Correctness**: this property refers to the extent to which the M&S system implementation conforms to its specifications (e.g., conceptual model, design specification); and is free of design and development defects (e.g., semantic errors, syntactic errors, numerical errors, user errors). The level of correctness impacts both validity and utility.

These three types of M&S properties include but not limited to capability, accuracy, usability and fidelity [13, 14]. To make an acceptance decision, the V&V User/Sponsor needs to know whether the M&S system or results are (un)acceptable, as well as the evidential value of this acceptance claim (i.e., strength). The required evidential strength to establish sufficient trust in such a claim depends on the use

risks and the V&V User/Sponsor responsibilities (i.e., liability). The convincing force that can be placed on such a claim depends on the quality of the whole V&V effort. For this purpose, the GM-VV identifies quality properties that can be associated with identifying and defining the acceptability criteria; and developing convincing evidence for demonstrating their satisfaction (Fig. 7).

- **V&V Quality**: this property refers to how well the V&V effort is performed (e.g., rigor) with regard to developing the acceptability criteria, collecting evidence, and assessing to what extent the M&S satisfy the acceptability criteria (e.g., evidential value, strength).

Typical examples of V&V quality properties are the completeness, correctness, consistency, unambiguous and relevance of the acceptability criteria or their supporting items of evidence. In the process of collecting or generating evidence, quality properties could comprise independence of applied V&V techniques or persons, knowledge gaps and uncertainties of referent data for the simuland [15], skill level of V&V personnel, and reliability and repeatability of V&V techniques. Relevance and warrants for any assumption made in a V&V effort could also be addressed in the form of quality properties.

The defined acceptability criteria, the collected evidence and assessment of the satisfaction of these criteria are the basis for developing the arguments underlying the acceptance claim. This acceptance claim provides the V&V User/Sponsor with a recommendation regarding the acceptability of the M&S system or result for the intended use. In practice, an acceptance recommendation is not necessarily just a yes or no claim, in the sense that an M&S system or results can be accepted only if it meets all of the acceptability criteria and cannot be accepted if it does not. Meeting all the acceptability criteria means the claim can be made that the M&S system or result should be accepted to support the intended use without limitations. In case not all acceptability criteria are met, alternative weaker acceptance claims with underlying arguments can be constructed. Such alternative acceptance claims could, for example, provide recommendations regarding conditions or restrictions under which the M&S system or result can still be used (i.e., limit the domain of use); or on modifications that, when implemented, will lead to an unconditionally acceptable M&S system or results for the intended use. Another rationale for alternative acceptance claims is when convincing or sufficient evidence is lacking (e.g., access to data prohibited, or referent system unavailable for testing). In any case, an acceptance recommendation always requires well-structured supporting arguments and evidence for the V&V User/Sponsor to make the right acceptance decision. Depending on the identified M&S use risk, the V&V User/Sponsor can also decide not to take any actions when not all acceptability criteria are met by the M&S system. In that case, the V&V User/Sponsor simply accepts the risks associated with the M&S system use.

2.1.5   V&V Argumentation Approach: Structured Reasoning with Arguments

Developing an acceptance recommendation that meets the V&V User/Sponsor needs usually involves the identification and definition of many interdependent acceptability criteria, particularly for large-scale and complex M&S systems or for M&S-based solutions used in safety–critical, real-world environments. Demonstrating the satisfaction of acceptability criteria requires evidence. Collecting the appropriate evidence is not always simple and straightforward, or even not always possible due to various practical constraints (e.g., safety, security, costs, schedule). In many cases, the collected evidence comprises a large set of individual items or pieces of evidence that may be provided in different forms or formats, and may originate from various sources (e.g., historical, experimental data, SME opinion). Moreover, the strength of each item of evidence may vary and the total set of collected evidence may even contain contradicting items of evidence (i.e., counter evidence). The quality of this effort determines the value of an acceptance recommendation for the V&V User/Sponsor. Therefore, the arguments underlying an acceptance recommendation should be developed in a structured manner using a format where the reasoning is traceable, reproducible and explicit. Alternative approaches to implement such reasoning exist and may be incorporated within the GM-VV technical framework to tailor it the specific needs of an M&S organization or domain. An example of such an approach is the V&V goal-claim network approach (Fig. 8). A V&V goal-claim network is an information and argumentation structure rooted in both goal-oriented requirements engineering and claim-argument-evidence safety engineering principles [16–19].
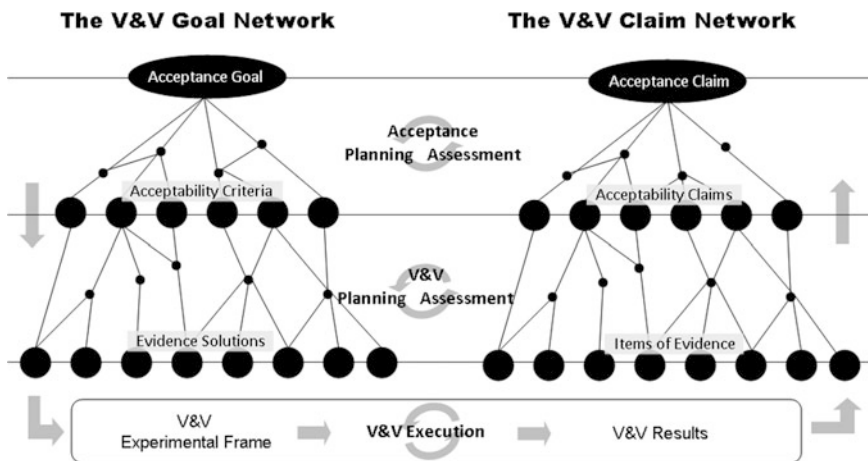

Fig. 8 V&V goal—claim network

Figure 8 provides an abstract illustration of a V&V goal-claim network. The left part of the goal-claim network is used to derive the acceptability criteria from the acceptance goal; and deriving solutions for collecting evidence to demonstrate that the M&S asset satisfies these criteria as indicated by the top-down arrows. The acceptance goal reflects the V&V needs and scope (e.g., simuland, intended use). Evidence solutions include the specification of tests/experiments, referent for the simuland (e.g., expected results, observed real data), methods for comparing and evaluating the test/experimental results against the referent. Collectively, they specify the design of the V&V experimental frame used to assess the M&S system and its results. When implemented, the experimental frame produces the actual V&V results. After a quality assessment (e.g., for errors, reliability, strength), these results can be used as the items of evidence in the right part of the goal-claim network. These items of evidence support the arguments that underpin the acceptability claims. An acceptability claim states whether a related acceptability criterion has been met or not. Acceptability claims provide the arguments for assessing whether or to what extent the M&S system and its results are acceptable for the intended use. This assessment, as indicated by the bottom-up arrows in Fig. 8, results in an acceptance claim inside the V&V goal-claim network. As such a V&V goal-claim network encapsulates, structures and consolidates all underlying evidence and argumentation necessary for developing an appropriate and defensible acceptance recommendation. The circular arrows in Fig. 8 represent the iterative nature of developing a V&V goal-claim network during planning, execution and assessment phases of a V&V effort.

2.1.6   V&V Organizational and Management Approach

In order to facilitate efficient and high quality V&V, the V&V effort inside the V&V world should be executed in a controlled and organized way. The basic premise of the GM-VV is that the acceptance recommendation for an M&S asset is developed and delivered by means of a managed project. Moreover, GM-VV assumes that V&V is conducted by a person, a team of people or a dedicated organization with assigned responsibilities, obligations and functions. Therefore, GM-VV identifies three organizational levels at which V&V efforts can be considered. In order of the lowest to the highest organizational level these levels are:

- Technical Level: concerns the engineering aspects of a V&V effort that are necessary to develop and deliver an acceptance recommendation,
- Project Level: concerns the managerial aspects related to the proper execution of the technical actions of a V&V effort,
- Enterprise Level: concerns the strategic and enabling aspects to establish, direct and support the execution or business environment for V&V efforts.

The core GM-VV concept on the V&V project level is the concept of a managed project. A V&V project can be viewed as a unique process comprised of coordinated and controlled activities that address: V&V effort planning in terms like cost, timescales and milestones; measuring and checking progress against this planning; and selecting and taking corrective actions when needed. A V&V project could be a separate project alongside the M&S project of which the M&S asset is part, or be an integral part of this M&S project itself (e.g., subproject, work package). A separate V&V project is particularly relevant in the case when a level of independence must be established between the M&S development and V&V team/organization. On the V&V project level, GM-VV also provides derived concepts such as a V&V plan and report to manage the technical V&V work.

For CIPRNet all three levels are important. For CI it is important to have a good set of tools and techniques to do the technical V&V activities. Since with the application of M&S systems for serious CI application there is always M&S Use Risk involved, for each project run by the to be established EISAC (European Infrastructures Simulation and Analysis Centre), V&V activities should be executed. A project approach is suited for that. Doing V&V from within EISAC means that EISAC should have support for the V&V activities at the highest level: the enterprise level.

The core GM-VV concept on the V&V enterprise level is the concept of an enterprise entity. A V&V enterprise entity can be viewed as an organization that: establishes the processes and lifecycle models to be used by V&V projects; initiates or defers V&V projects; provides resources required (e.g., financial, human, equipment); retains reusable knowledge and information from current V&V projects; and leverages such knowledge and information from previous V&V projects. The V&V enterprise provides the environment in which V&V projects are conducted. GM-VV defines two types of enterprise entities:

- **V&V Client**: the person or organization that acquires V&V products or services,
- **V&V Supplier**: the person or organization that develops and delivers V&V products or services.

A V&V agreement is arranged between a V&V client and V&V supplier to provide products and/or services that meet the V&V client's needs. Both these V&V entities could be organizations (e.g., companies) separate from the organization that develops or acquires M&S or it could be different units (e.g., department, division, group) within a single M&S supplier or client organization. Typically, a separate V&V supplier is an organization that has the provision of independent V&V products and services to external V&V clients as its core business. Though depending on their business model, an M&S supplier or client organization could have their own V&V supplier entity that may provide V&V services and products to internal and external V&V clients alike.

2.1.7 V&V Levels of Independence: Acceptance, Certification and Accreditation

An independent V&V (IV&V) authority is often described as an organization or a person that is employed to conduct V&V, independent of the developer's team or organization [6, 10, 12]. The need for IV&V is mostly driven by:

- risks and liabilities taken by the V&V User/Sponsor's acceptance decision,
- level of trust the V&V User/Sponsor has in the M&S developer,
- authoritative policies and regulations that may demand independent V&V for the M&S intended use,
- lack of specialist skills, tools and techniques by user, sponsor or developer to perform V&V.

In practice however, it is highly incumbent upon the V&V User/Sponsor acceptance decision needs and complexity of the M&S system as to which parts and to what extent V&V should be conducted in an independent manner. Therefore, the GM-VV adopts a sliding scale of independence for V&V [15], which can be selected accordingly to match the V&V needs. The justification and selection of a proper level of independence is supported within GM-VV through the use of the V&V argumentation network. Within this sliding scale for independent V&V, certification and accreditation can be located in the right part of the scale (Fig. 9).

2.1.8 V&V Information and Knowledge Management

V&V of M&S is an information and knowledge intensive effort. In particular, during the V&V of large scale, distributed or complex M&S applications, care must be taken to preserve or reuse information and knowledge. Therefore, GM-VV
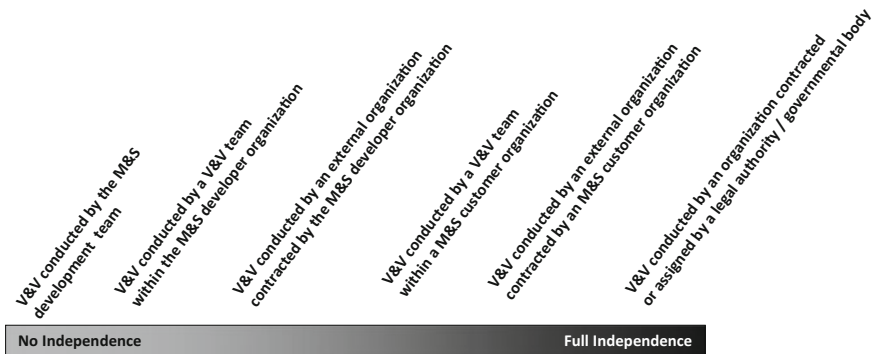

Fig. 9 Levels of independent V&V

applies the memory concept on both the V&V project and enterprise levels. A memory is viewed as a combination of an information and knowledge repository and a community of practice [20]. The repository is a physical place where information, knowledge objects, and artefacts are stored. The community of practice is composed of the people who interact with those objects to learn, understand context and make decisions.

The V&V project memory provides the means to manage information and knowledge produced and used during the lifetime of an individual V&V project. V&V is often an iterative and recurrent process linked to an M&S system's life-cycle, hence V&V products for an M&S system may have different configurations. Therefore, a V&V project memory may also retain records on possible different V&V product configurations. The V&V enterprise memory retains the total body of information and knowledge from past and current V&V projects to sustain and support the cost-effective execution of future V&V projects. Such reusable information could be, for example, M&S technology or domain specific recommended practices, acceptability criteria, V&V goal-claim network design patterns, V&V tools and techniques, or policies and standards. On a more strategic level, a V&V enterprise memory could retain information and knowledge on V&V project costs and maturity as well.

2.2 Implementation Framework

The GM-VV implementation framework translates the GM-VV basic concepts into a set of generic V&V building blocks (i.e., components). These may be used to develop a tailored V&V solution that fits the V&V needs of any particular M&S organization, project, application, and technology or problem domain. The implementation framework has three interrelated dimensions: product, process and organization (Fig. 10). The underlying principle of this framework is that the V&V needs of the V&V User/Sponsor in the M&S four-world view are addressed by one or more V&V products, those being the V&V report and possibly other custom V&V products the V&V User/Sponsor may need. These V&V products in general require intermediate products (i.e., information artefacts) and associated processes to produce them. The V&V processes are executed by a corresponding V&V organization that is responsible for the development and delivery of the V&V products. In general the V&V effort should result in a V&V report to be delivered to the customer containing one or more of the information artefacts. Individual needs will drive which V&V products are required.

Fig. 10 GM-VV implementation framework dimensions
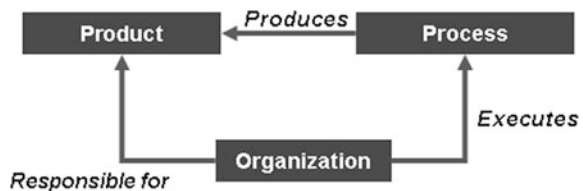
As indicated in Fig. 10, the GM-VV implementation framework consists of three key dimensions:

- **Products**: the information artefacts that may be delivered, developed or used throughout a V&V effort. These artefacts can have multiple instances, representational and documentation formats.
- **Processes**: the set of activities and tasks that comprise V&V execution as well as those management tasks that increase the efficiency and effectiveness of the V&V effort. These activities and tasks are inspired by the IEEE standard system life-cycle processes model [2] and can be carried out recursively, concurrently, and iteratively.
- **Organization**: the roles played either by people or by organizations in the V&V effort. The roles are defined in terms of responsibilities and obligations. Depending on the M&S organization, project and application domain needs; several roles could be played by separate organizations, separate people in one organization or by a single person.

The V&V effort culminates in a V&V report that is comprised of the information generated throughout the execution of the V&V and acceptance decision-support process (Fig. 6). The following sub-sections provide an overview of the information artefacts, activities and roles that are implemented or produced during this execution. They are ordered according to the GM-VV technical, project and enterprise levels.

It is important to re-emphasize the tailorable nature of the methodology. GM-VV provides all the elementary information artefacts, activities, tasks and roles to address the most common technical, project and enterprise level aspects of a V&V effort. Depending on the M&S project and organizational needs one could choose not to implement all GM-VV components or one could choose to adjust them accordingly. This is particularly relevant for M&S organizations that already have some project and enterprise level components in place, and only require technical level V&V (intermediate) products, processes and roles to conduct their V&V effort. The overall tailoring and application concepts of the GM-VV implementation framework are provided in the next section.

2.3 Tailoring Framework

GM-VV recognizes that a particular M&S organization, project, application, technology or problem domain may not need all these components or use them directly as-is. Therefore, the GM-VV components are intended to be selected, combined and modified accordingly, to obtain an effective and efficient V&V effort of sufficient rigor. This is particularly relevant for M&S projects and organizations that already have some project and enterprise level components in place, and only require technical level V&V (intermediate) products, processes and roles to conduct their V&V effort.

The basic premise of the GM-VV tailoring concept is that the GM-VV should first be cast into a concrete V&V method fit for an organization or application domain, and secondly this instance should be optimized for a V&V project. This tailoring concept is implemented by means of a framework that refers to all three levels of the GM-VV implementation framework. The objective of this GM-VV tailoring framework is to adapt each GM-VV (intermediate) product, process and role to satisfy the specific requirements and constraints of:

- An organization that is employing the GM-VV (e.g., company policies, standards)
- A domain in which the GM-VV is employed (e.g., standards, regulations, technologies)
- A V&V supplier entity delivering V&V products or services (e.g., standards, processes)
- A V&V project (e.g., time, budget, scale, complexity, risk, resources).

As described above tailoring is accomplished in two phases. In the first phase of the GM-VV tailoring framework, the implementation framework components are utilized to establish concrete V&V solution instances on one or more of the three organizational levels (i.e. a permanent V&V organization, V&V project or technical V&V approach). In here, the GM-VV recognizes that a particular M&S organization, project, technology or problem domain may not need all three organizational levels or all components on a single organizational level nor even use them directly as-is. Therefore, the GM-VV implementation framework organizational levels and components are selected, combined and modified accordingly, to obtain a concrete tailored V&V solution. For instance an M&S organization may already have an M&S project and enterprise level in place, and only require technical level V&V (intermediate) products, processes and roles to conduct their technical V&V work. Successful application of the first phase of the tailoring framework results in a modified or new V&V solution instance conforming to the GM-VV architectural templates (i.e. in a structure and organizational manner). Four tailoring approaches can be used for this: extension, omission, specialization and balancing, which are discussed below.

In the second phase these same tailoring approaches are applied throughout the operational lifetime (i.e. permanent organization or project) or execution (i.e. technical approach) of each V&V solution instance. This type of tailoring comprises run-time optimization of the instantiated V&V processes at all three organizational levels. At a technical level this could imply the application of a risk-based V&V approach, such as the MURM [21], to prioritize the acceptability criteria, allocate and specific V&V techniques and tools based on V&V User/Sponsor risk tolerance levels. On the project level this could be the alignment of technical V&V activities with the progress of the M&S system's life-cycle phases, balance and allocate the available V&V resources to each phase M&S life-cycle or (work) products. On the enterprise level this could mean balancing the

cost-risk of new investments in training of personnel or V&V tool infrastructure development against a future V&V project order intake volume.

The GM-VV tailoring framework applies four basic tailoring approaches:

- **Tailoring by Extension**: adaptation of the implementation framework by adding custom V&V products, processes, activities, tasks and roles. For example, a V&V Client organization or application domain may require additional custom artefacts not foreseen by the GM-VV.
- **Tailoring by Reduction**: adaptation of the implementation framework by deleting products, processes, activities, tasks and roles due to constraints such as inaccessibility of data and information protected by intellectual property rights, security or technical restrictions.
- **Tailoring by Specialization**: adaptation of the implementation framework by adding or using domain specific V&V methods, techniques and data that are unique for a V&V project, organization or application.
- **Tailoring by balancing**: adaptation of the implementation framework by fitting a suitable cost-benefit-ratio towards an acceptance recommendation. The level of acceptable M&S use risk should drive the rigor and resources employed for V&V. Therefore, in this approach one tries to balance aspects such as:

  - M&S use-risk tolerances and thresholds
  - criticality and scope of the acceptance decision
  - scale and complexity of the M&S system
  - information security, with

  V&V project resource variables such as

  - time schedule
  - budget
  - V&V personnel skills
  - infrastructure.
  - Hence, balancing establishes the suitable and feasible level of rigor for the V&V effort.

Tailoring by these four approaches should be performed in accordance with the three dimension design principle of the GM-VV implementation framework (Fig. 10), to obtain a consistent and coherent V&V method and project. For example, each new or specialized product needs a corresponding process (activities, tasks) and role (responsibilities, obligations).

Successful application of the tailoring framework results in a modified or new V&V method instance conforming to the GM-VV. This consists of concrete V&V organization, products and processes, which should achieve the V&V objectives of an M&S organization, project, technology, or application domain.

The first three types of tailoring are mainly of importance at the start of a V&V effort. The tailoring by balancing is important during the V&V effort.

### 2.3.1 Risk Decomposition and Tailoring by Balancing

As described above it is advised to use a decomposition of the top goal into smaller and smaller goals up to the point that a test can be devised that is within resources and is likely to deliver suitable evidence. During the balancing tailoring during the execution of the V&V work priorities need to be determined. These priorities together with the resources available are used to decide which goals will be further expanded and which will be left undeveloped. The basis for that decision and thus for the prioritization is risk. What is needed is to determine the contribution of a goal to the overall M&S use-risk. If a goal has a high contribution of risk it must be taken into consideration in the V&V work. If it has a very low contribution it can. In that case it should be explicitly be recorded that that goal is not used in the rest of the V&V work such that at the end a feeling for how complete the V&V work is can be obtained.

An evidence solution for a goal with a (relatively) high contribution to the overall risk should likely result in a high confidence in the evidence. For a goal with a low contribution to the M&S use-risk risk it may be sufficient to have evidence that contains some uncertainty, i.e. if the evidence is just an indication that the goal is met it may already suffice.

To find the contribution to the overall M&S use risk for a node it is necessary to make a risk decomposition in the same way as the decomposition of the Acceptance Goal, see Fig. 8. In practice it is difficult to make an exact risk decomposition, therefore it is advised to use a somewhat simpler approach as indicated in Fig. 11. The red stands for high contribution to the overall M&S use risk, orange for medium contribution and green for low contribution. During the decomposition nodes with a low contribution to the overall use risk may be left undeveloped. At the bottom of Fig. 11 the contribution to the risk is an indication of how convincing the evidence should be which is important for specifying which type of tests are required.

If after evidence collection it turns out not all goals are met, the contribution to the overall risk may be used during the acceptance decision to decide what to do. If it concerns a node with low contribution to the overall M&S use risk, it may be decided to leave things as is and accept the small risk. If it is goal with a medium or

Fig. 11 Risk decomposition
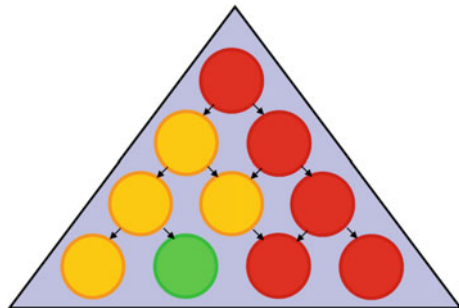
high contribution to risk it can be decided to either change the M&S system such that the identified problems are corrected, or the purpose for which the user intends to use the M&S System should be made smaller such that the current state of the M&S System will be fit for purpose.

2.4 Why Is This Structured Approach so Much More Effective and Efficient

The above-described structured approach to doing V&V has a number of advantages that make it more effective and more efficient than doing V&V in a less structured way. Below some of the key advantages are discussed.

The right starting point for the V&V effort leads to more effective results
The V&V effort should start from the perspective of risk. Who runs the real risk in an M&S endeavour? It is not the modeller, not the implementer (maybe there is a risk of repetitive strain injury) and not the person who executes the simulation (maybe if there is a moving base simulator). In general the real M&S use risk is found when the M&S based results are applied in the real word. Therefore V&V processes that are developer oriented might miss the real risk. Also, when studying the 4-world view in Fig. 5 it may become clear that possibly many more aspects may need to be considered than just the domain knowledge as coded in a simulation. Thus organizational aspects that may make or break the use of simulation, the level of proficiency of all people involved, the processes used to derive the products such as the Operational Needs, etc. may all play a significant role and may need to be included in determining the overall utility and thus in the V&V approach. If such a very broad scope is used it becomes clear that a domain oriented V&V process may also miss some aspects. Therefore a general methodology that starts at the true M&S use risk and that can incorporate domain specific elements as well as other aspects will result in a more effective V&V result because the right starting point can be chosen and all relevant aspects included.

Balancing resources with needs leads to efficiency and effectiveness
A structured decomposition of the Acceptance Goal into all aspects that are relevant and on top of that a decomposition of the contribution of the M&S use risk attached to the Acceptance Goal leads to the possibility to spend the available resources for the V&V effort wisely. Based on priorities related to the contribution to the overall M&S use risk it can be decided which parts of the decomposition requires more or less effort. When available resources do not allow testing all aspects to their maximum, i.e. in all practical situations, it can be decided to let the goals with low contribution to risk remain undeveloped. In that case it should be explicitly recorded that that goal is not used in the rest of the V&V work, see "Knowledge of the completeness of the V&V effort leads to effectiveness" below. If nodes are developed to the point where tests can be defined, the contribution to the M&S use risk can be used to make choices for tests. Low contribution to the risk allow for cheaper tests

that may not give a high convincing force. A high contribution to the risk means that sufficient convincing force must be required of the evidence, possibly meaning more expensive tests need to be performed.

The structured approach to V&V makes it possible to balance the resources during the construction of the goal network and the evidence solutions. This means that the V&V effort uses the available resources in an efficient way, allowing for the best possible answer for the given resources, which means the highest possible effectiveness.

Re-usable domain knowledge leads to more efficient and effective results
The top part of the decomposition of the Acceptance Goal, see Fig. 8, contains domain knowledge because it is the user's perspective that is encoded and the role of the M&S system in that domain. From an V&V enterprise point of view, see Sect. 2.1.6, this domain knowledge can be re-used if other V&V projects are executed on (almost) the same domain or for (almost) the same purpose. In that case the domain knowledge can be re-used and even extended to be more complete. Of course, for each new project in which existing domain knowledge is re-used it must be made sure that no irrelevant aspects are taken into account. Over the course of several projects the domain knowledge becomes more and more complete, which helps in not forgetting possibly important aspects. The re-use of domain knowledge thus leads to more to a more effective and more efficient V&V effort. It is, however, needed that a good discipline in documenting the V&V effort is used.

Distribute the V&V work among experts leads to efficiency
In the lower part of the goal-network many different aspects covering many different disciplines can be found. The expansion—if needed—of these goals and the execution of associated tests likely requires different experts and facilities. Using the natural break up of a structured approach to V&V, e.g. the tree structure in Fig. 8, it becomes easier to assign experts to different groups of goals and tests. For CI simulation it may be that organizations do not wish to have other experts test their simulation assets, in that case each partner can be assigned a set of goals for which they need to provide evidence. It would be better, however, to have a certain level of independence (see Sect. 2.1.7). The structured approach leads to more efficient execution of the V&V effort by clearly indicating which expertise should be handled by which expertise.

Complete one branch while waiting for others to complete leads to efficiency
In the structured approach as presented above, it becomes clear that if one branch of the tree structure is fully developed and ready for execution of the tests, there may be no need to wait for other parts to also become fully developed. The parts that are ready to go to the test phase can start independently of the rest. This may even lead to the discovery of problems with the M&S System that can already be corrected before tests of other branches are executed. This leads to a more overall efficient V&V effort.

Knowledge of the completeness of the V&V effort leads to effectiveness
During the balancing of the resources in building the goal network and the specification of the evidence solutions the important decisions on when goals with a low

contribution to the M&S use risk are left undeveloped and which tests are chosen in the specification of the evidence solutions should be unambiguously be recorded. That makes it possible to get a feeling for how complete the V&V effort as a whole is. This completeness should be translated into an uncertainty in the Acceptance Recommendation to the customer. Thus if insufficient resources were allocated to the V&V work, the conclusion might state that the available evidence indicates that the M&S system is fit for purpose, but that the V&V effort as a whole has left too many aspects out of consideration and that thus a high level of uncertainty is present in that statement.

The statement on completeness of the V&V effort will allow the decision maker to make a much better decision, which leads to better effectiveness of the use of the V&V results.

Standardized documentation leads to efficiency
An often observed problem with unstructured V&V efforts is that it results in either very little documentation or it results in a lot of documents that are unorganized and scattered over different places, usually in the form of computer files that are difficult to find and for which it is hard to recall what its content means and in what piece of evidence it was used.

A structured approach should adopt some standard approach to documentation. This documentation should be such that the Acceptance Recommendation should be completely traceable through the claim network, via the evidence collection, through the goal network back to the Acceptance Goal. Also all decisions due to tailoring should be well documented and immediately clear where they influence the Acceptance Recommendation.

A standardized approach to documentation is also important on the V&V enterprise level where it can be expected that re use of previous V&V projects will lead to efficiency.

Efficiency for recurrent testing
In practice it may occur that a M&S system had been used for some time and that subsystems are being replaced or upgraded. In that case the structured approach described above makes it immediately clear which parts of the goal network are affected by the change and which tests should be re-done for the new M&S system. This leads to a very efficient way of doing recurrent testing.

3 Choose the Appropriate Verification and Validation Technique

There are many different V&V techniques, see e.g. [22–25]. The V&V techniques in those references are categorized in Table 1.

Table 1 Examples of V&V techniques

Informal	Formal	Static	Dynamic
· Audit · Desk checking · Documentation checking · Face validation · Inspections · Reviews · Turing test · Walkthroughs	· Induction · Inductive assertions · Inference · Logical deduction · Lambda calculus · Predicate calculus · Predicate transformation · Proof of correctness	· Cause-effect graphing · Control analysis · Calling structure analysis · Concurrent process analysis · Control flow analysis · State transition analysis · Data analysis · Data dependency analysis · Data flow analysis · Fault/failure analysis · Interface analysis · Model interface analysis · User interface analysis · Semantic analysis · Structural analysis · Symbolic evaluation · Syntax analysis · Traceability assessment	· Acceptance testing · Alpha testing · Assertion checking · Beta testing · Bottom-Up testing · Comparison testing · Compliance testing · Authorization testing · Performance testing · Securitytesting · Standards testing · Debugging · Execution testing · Execution monitoring · Execution profiling · Execution tracing · Fault/failure insertion testing · Field testing · Functional (Black-Box) testing · Graphical comparisons · Interface testing · Data interface testing · Model interface testing · User interface testing · Object-flow testing · Partition testing · Predictive validation · Product testing · Regression testing · Sensitivity analysis · Special input testing · Boundary value testing · Equivalence partitioning testing · Extreme input testing · Invalid input testing · Real-time input testing · Self-driven input testing · Stress testing · Trace-driven input testing · Statistical techniques · Structural (White-Box) testing

Table 1 (continued)

Informal	Formal	Static	Dynamic
			• Branch testing • Condition testing • Data flow testing • Loop testing • Path testing • Statement testing • Submodel/module testing • Symbolic debugging

The four broad categories of V&V techniques can be described as:

- **Informal V&V techniques** are usually executed and interpreted by humans. Typically these require few resources and can be executed in a short time. The convincing force, however, depends on the trust in the humans doing the work and the process they use.
- **Formal V&V techniques** are based on mathematical proofs of correctness. The application of formal methods, however, is often limited due to large resource costs even for relatively small M&S systems and their use. If applicable, the convincing forces of the V&V results are very strong.
- **Static V&V techniques** can be applied early in the development process because no execution is required. It is typically applied in the concept phase and parts of the development phase. Typically specialized tools are used to do automated checks. The required resources are normally limited. It is required to have access to documentation and half-products. The strength of the convincing force is dependent on the rigor of the tests.
- **Dynamic V&V techniques** require execution of the M&S System in part or as a whole. The dynamic properties of the M&S System are studied and checked. Typically specialized tools are used to do automated measurements and checks. The required resources are normally limited. Dynamic V&V techniques may require access to parts of the M&S System that are usually not available. The strength of the convincing force is dependent on the rigor of the M&S System check.

It is difficult to state in general which V&V techniques (i.e. what type of tests) should be used. So in this text we provide a basis to choose the right V&V technique. There are a number of important aspects that determine which V&V techniques are appropriate for a given situation:

- Contribution to the M&S Use Risk

    - It is clear that a relatively high contribution to the M&S Use Risk requires evidence that can be trusted. This requires a rigorous V&V technique, i.e. one for which the expected residual uncertainty is low. When possible formal techniques should be used. In practice however, this is often prohibitively

expensive and (combinations of) techniques have to be used that are with the available means but still deliver sufficiently trustworthy evidence.

- Available means

  - The available means are a set of limiting factors such as budget, time, expert knowledge, access to testing facilities, etc. The whole V&V effort has to be run within these limits. That means that during the construction of the goal network only those criteria can be considered that contribute highly to the over M&S Use Risk and collectively are likely to remain within the available means. The collection of evidence solutions has to be chosen such that the expected results of executing the tests delivers the lowest overall residual uncertainty.

- Referent data

  - The Referent data is the knowledge of the real world. It is needed during the tests to compare the simulation results with. If no or little referent data is available only tests that do not (heavily) depend on referent data can be chosen, e.g. expert opinion or examination of the conceptual model.

- M&S system availability

  - For dynamic testing it is evident that (parts of) the M&S system itself has to be available. Some types of tests require access to M&S system internals in order to make "measurements" that are not visible to the end user. For other tests it is necessary to have access to development documents such as the conceptual model.

Summarizing: the tests all have different costs and different expected residual uncertainty. The contribution to the M&S User Risk should be the basis for choosing the best V&V techniques. A set of evidence solutions need to be chosen such that collectively the best possible result for the given available resources is obtained.

Take away message: You have to choose the appropriate Verification and Validation techniques to balance risk, effectiveness and efficiency.

4   Conclusion

As a very brief summary of the text above it can be stated that:

- You have to do Verification and Validation because there is risk involved,
- You have to do it in a structured way if you want to do it more effective and more efficient,
- You have to choose the appropriate Verification and Validation technique to balance risk, effectiveness and efficiency.

References

1. SISO (2012) GM-VV Vol. 1: introduction & overview, SISO-GUIDE-001.1-2012
2. IEEE, Systems and software engineering—system life-cycle processes, IEEE Std 15288-2008, Jan 2008
3. DoDI 5000.61, DoD modeling and simulation (M&S) verification, validation, and accreditation (VV&A), 9 Dec 2009
4. Roza M, Jacquart J, Giannoulis C (2009) Common validation, verification and accreditation framework for simulation, REVVA-2 Reference Manual. Report: Europa 111–104. Mar 2009
5. IEEE, IEEE recommended practices for verification, validation and accreditation of a federation—an overlay to the high level architecture (HLA) FEDEP, IEEE Std 1516.4-2007, Dec 2007
6. US DoD, Safety Management College, Systems engineering fundamentals, SEF-Guide 01-01
7. Wasson CS (2006) System analysis, design and development: concepts, principles and practices. Wiley, Hoboken
8. INCOSE (2002) Systems engineering handbook, a "How To" guide for all engineers. Version 2.0. INCOSE
9. Grady JO (1998) System verification and validation. CRC Press LLC, Boca Raton
10. US Department of Defense Verification, Validation, and Accreditation Recommended Practices Guide, RPG Build 4.0, Nov 2011. http://vva.msco.mil/
11. UK MoD, A generic process for the verification & validation of modeling and simulation & synthetic environments systems, DEF STAN 03-44 Issue 2, 31 Mar 2008
12. Australian DoD (2005) DSO, Simulation verification, validation and accreditation guide
13. Gross DC et al (1999) Report from the fidelity implementation study group. 1999 Spring SIW Proceeding, 99S-SIW-167
14. Roza ZC (2004) Simulation fidelity theory and practice: a unified approach to defining, specifying and measuring realism of simulations. Delft University Press Science, Delft
15. Oberkampf WL, Roy CJ (2010) Verification and validation in scientific computing. Cambridge University Press, Cambridge
16. Lamsweerde van A (2001) Goal oriented requirements engineering: a guided tour. In: Fifth IEEE international symposium on requirements engineering
17. Anwer S, Ikram N (2006) Goal oriented requirement engineering: a critical study of techniques. In: XIII Asia Pacific software engineering conference
18. Kelly TP (1998) Arguing safety—a systematic approach to managing safety case. Master's thesis, University of York, Sept 1998
19. Mayo P (2002) Structured safety case evaluation: a systematic approach to safety case review. Master's Thesis, University of York
20. Wikipedia, Information management, information lifecycle management and knowledge management
21. Risk-based methodology for verification, validation and accreditation (VV&A): the M&S use risk methodology (MURM). Johns Hopkins University Applied Physics Laboratory Report, NSAD-R-2011-011, Apr 2011

22. Balci O (1997) Verification, validation, and accreditation of simulation models. In: Andradhttir S, Healy KJ, Withers DH, Nelson BL (eds) Proceedings of the 1997 winter simulation conference, pp 135–141
23. Sargent (2010) Verification and validation of simulation models. In: Johansson B, Jain S, Montoya-Torres J, Hugan J, Yücesan E (eds) Proceedings of the 2010 winter simulation conference, pp 166–183
24. DoD (2001) V&V techniques, RPG Reference Document, vva.msco.mil
25. Petty D (2011) Model verification and validation methods, I/ITSEC tutorial 1101