

Chapter 9

Quality Indicators Throughout the Modeling and Simulation Life Cycle

Osman Balci

9.1 Introduction

I met Dr. Tuncer Ören in April 1980 at a meeting in New York City where he presented a paper entitled “Concepts and Criteria to Assess Acceptability of Simulation Studies: A Frame of Reference,” which was later published in the *Communications of the ACM* (Ören 1981). During that time, I was focusing on my Ph.D. dissertation research in the area of simulation model validation. I also presented a paper at that meeting entitled “A Methodology for Cost-Risk Analysis in the Statistical Validation of Simulation Models,” which was later published right after his paper in the same special issue of the *Communications of the ACM* (Balci and Sargent 1981).

Dr. Ören’s seminal paper in the *Communications of the ACM* has expanded my horizon in assessing the acceptability of modeling and simulation applications. I am honored to contribute this chapter presenting quality indicators that can be used for such acceptability assessment. Dr. Ören has published and presented more than 85 articles just on the topic of reliability, quality assurance, and failure avoidance in modeling and simulation since his seminal paper. He has been an internationally recognized leading authority not only in this topic but also in the whole modeling and simulation discipline. Dr. Ören’s linguistic ability is beyond my comprehension. He is the only person I know who can deliver a very technical speech in Turkish without using a single English word! That is unbelievable!

As the saying goes, “Quality is Job 1!” Quality is a critically important issue in almost every discipline. Whether we manufacture a product, employ processes or provide services, quality often becomes a major goal. Achieving that goal is the

O. Balci (✉)

Mobile/Cloud Software Engineering Lab, Department of Computer Science, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, USA
e-mail: balci@vt.edu

challenge. Many associations have been established worldwide for quality, e.g., American Society for Quality (<http://www.asq.org>), Australian Organization for Quality (<http://www.aoq.asn.au>), European Organization for Quality (<http://www.eoq.org>), and Society for Software Quality (<http://www.ssq.org>). Manufacturing companies have quality control departments, business and government organizations have Total Quality Management programs, and software development companies have Software Quality Assurance departments to be able to meet the quality challenge.

Quality can be generically defined for anything X as: Quality of X is the degree to which the X possesses a desired set of characteristics.

The ultimate goal of a modeling and simulation (M&S) project is to develop an M&S application with sufficient quality characteristics. M&S quality assurance (QA) refers to the planned and systematic activities that are established throughout the M&S life cycle to substantiate adequate confidence that an M&S application possesses a set of characteristics required for a set of intended uses.

M&S applications are mostly made up of software or are software based. Software is inherently complex and very difficult to engineer. Under the current state of the art, we continue to face serious technical challenges in developing a reasonably large and complex software product with acceptable accuracy. *Accuracy* refers to the transformational and representational/behavioral correctness and is considered just one of dozens of quality characteristics of an M&S application. M&S accuracy is judged by conducting M&S verification and validation (V&V). As advocated by Balci et al. (2002), we can increase our confidence in the accuracy of large-scale and complex M&S applications by employing a quality-centered evaluation approach.

The purpose of this chapter is to present quality indicators throughout the M&S life cycle that can be used for such a quality-centered evaluation approach. After providing background information and an introduction, we present a life cycle for modeling and simulation, which is applicable for any kind of M&S project. We describe the quality indicators throughout the M&S life cycle based on the experience and knowledge gained by the author in the U.S. Department of Defense large and complex M&S application verification, validation, and accreditation projects. Concluding remarks are stated to conclude this chapter.

9.2 Modeling and Simulation Life Cycle

A life cycle for M&S is presented in Fig. 9.1. This life cycle is a different representation of the same life cycle described by Balci (2012). The author has developed this life cycle based on his many years of experience and knowledge gained in more than a dozen U.S. Department of Defense large and complex M&S application development projects.

An M&S life cycle is a framework for organization of the processes, work products, quality assurance activities, and project management activities required

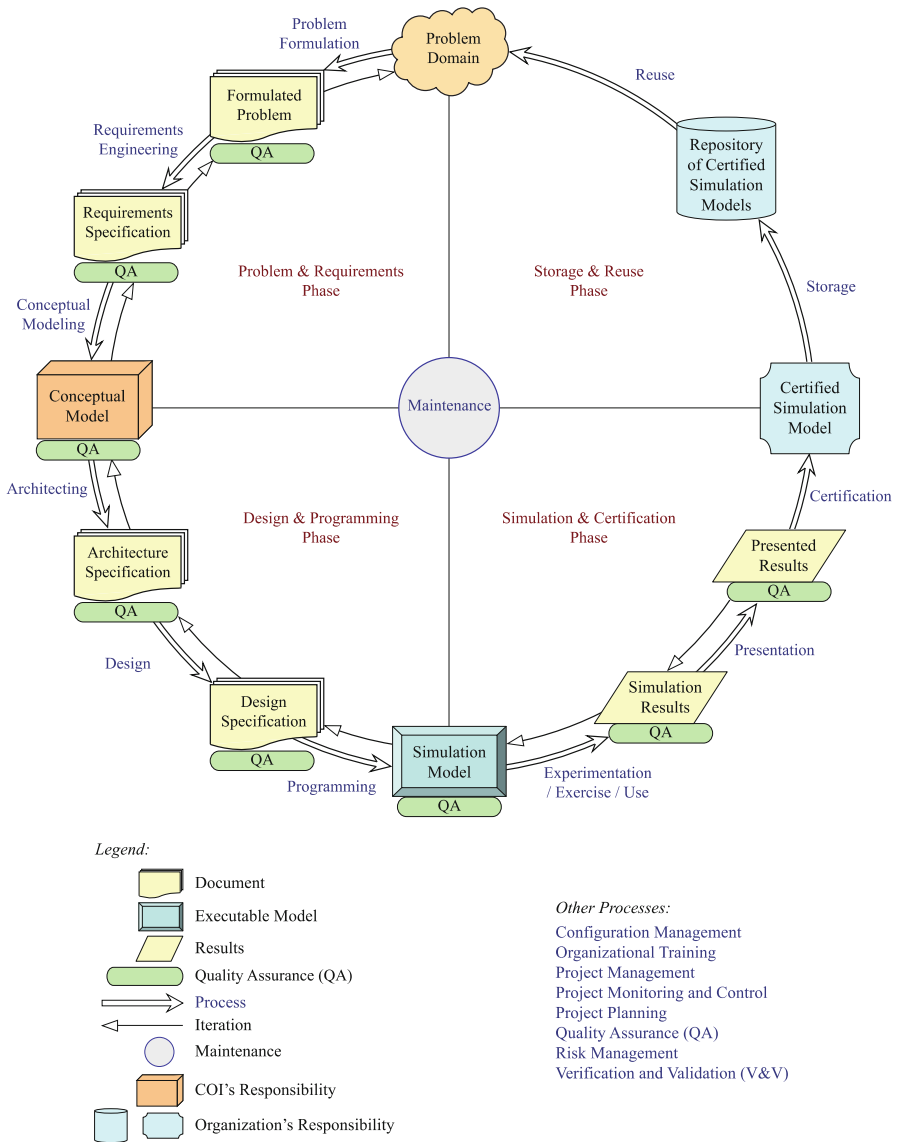


Fig. 9.1 A life cycle for modeling and simulation (Copyright © Osman Balci)

to develop, use, maintain, and reuse an M&S application from birth to retirement. The M&S life cycle is created to modularize and structure an M&S application development and to provide guidance to an M&S developer (engineer), manager, organization, and community of interest (COI).

The M&S life cycle presented in Fig. 9.1 enables to view M&S development from four perspectives (or Ps): Process, Product, People, Project. The M&S life

cycle (a) specifies the *work* products to be created under the designated *processes* together with the integrated verification and validation (V&V) and quality assurance (QA) activities, (b) modularizes and structures M&S development and provides valuable guidance for *project* management, and (c) identifies areas of expertise in which to employ qualified *people*.

The M&S life cycle consists of four phases as depicted in Fig. 9.1: problem and requirements phase, design and programming phase, simulation and certification phase, and storage and reuse phase. It consists of eleven major processes organized in a logical order, as depicted in Fig. 9.1, starting with Problem Formulation and culminating with Reuse. A process, represented by a double-line arrow, is executed to create a work product. For example, we execute the process of Requirements Engineering to create a Requirements Specification document or the process of Design to create a Design Specification document. A work product is created in different forms, i.e., document, model, executable, results, or repository, as shown with different symbology in Fig. 9.1.

The M&S life cycle should not be interpreted as strictly sequential or linear. The sequential representation of the double-line arrows is intended to show the direction of workflow throughout the life cycle. The life cycle is iterative in nature and reverse transitions are expected. For example, an error identified during V&V of the executable model may require changes in the requirements specification and redoing the earlier work. We typically bounce back and forth between the processes until we achieve sufficient confidence in the quality of the work products.

9.3 Quality Indicators

In this section, we present quality indicators throughout the M&S life cycle that can be employed under a quality-centered assessment approach for large-scale and complex M&S projects.

9.3.1 Formulated Problem Quality Indicators

Formulated problem quality can be assessed by employing the following indicators (Balci 2012):

1. What are the chances that the *real* problem is not completely identified due to the possibility that
 - 1.1. People might have personalized problems?
 - 1.2. Information showing that a problem exists might not have been revealed?
 - 1.3. The problem context is too complex for the analyst to comprehend?
 - 1.4. Root problems might have arisen in contexts with which people have had no experience?

- 1.5. Cause and effect may not be closely related within the problem context?
 - 1.6. The analyst might have been unable to distinguish between facts and opinions?
 - 1.7. The analyst might have been misguided deliberately or accidentally?
 - 1.8. The level of extraction of problem context was insufficiently detailed?
 - 1.9. The problem boundary was insufficient to include the entire real problem?
 - 1.10. Inadequate standards or definition of desired conditions exist?
 - 1.11. The root causes might be time dependent?
 - 1.12. A root cause might have been masked by the emphasis on another?
 - 1.13. Invalid information might have been used?
 - 1.14. Invalid data might have been used?
 - 1.15. Assumptions might have concealed root causes?
 - 1.16. Resistance might have occurred from people suspicious of change?
 - 1.17. The problem was formulated under the influence of a solution technique?
 - 1.18. The real objectives might have been hidden accidentally, unconsciously, or deliberately?
 - 1.19. Root causes might be present in other unidentified systems, frameworks, or structures?
 - 1.20. The formulated problem may be out of date?
2. Stakeholders and Decision Makers
 - 2.1. Do you know or can you think of any stakeholders and decision makers, other than the ones identified by the analyst, who might be aided by the solution of the problem?
 - 2.2. Are all active stakeholders (e.g., users of the solution system, administrators of the solution system, trainers of the solution system users) identified? (An *active stakeholder* is the one who will actively interact with the solution system once it is operational and in use.)
 - 2.3. Are all passive stakeholders (e.g., developers, decision makers about the use of the solution system, logistics personnel, manufacturer, owners/sponsors if they do not use/operate the solution system) identified? (A *passive stakeholder* is the one who will not actively interact with the solution system once it is operational and in use.)
3. Constraints
 - 3.1. Do you know or can you think of any other constraints, which should have been identified by the analyst?
 - 3.2. Are there any incorrect or irrelevant constraints?
 - 3.3. Are there any constraints that make the formulated problem infeasible to solve?
4. Objectives
 - 4.1. How well are the objectives stated?
 - 4.2. Do you believe any objectives to be inconsistent, ambiguous, or conflicting in any way?

- 4.3. How realistic are the objectives?
 - 4.4. Are there any priorities specified for the case where only some of the objectives are achievable?
 - 4.5. Do you know or can you think of any relevant stakeholders and decision makers whose objectives are conflicting with any of those specified?
 - 4.6. In case of multiple objectives, do you agree with the way the objectives are weighted?
 - 4.7. Do you agree that the stated objectives are the real objectives of the stakeholders and decision makers involved?
 - 4.8. Do you know or can you think of any associated objective, which is disguised or hidden accidentally, unconsciously, or deliberately?
 - 4.9. How often could the stated objectives change?
5. Data and Information
- 5.1. Are there any sources of data and information used by the analyst that you believe to be unreliable?
 - 5.2. Are there any data and information used by the analyst that you believe to be out of date or need to be updated?
 - 5.3. Are there any data and information, which you believe to be not sufficiently accurate?
6. Assumptions
- 6.1. How well are the assumptions stated?
 - 6.2. Are there any invalid assumptions based on which the problem is formulated?
 - 6.3. Are there any invalid inferences or conclusions drawn by the analyst?

9.3.2 *Requirements Quality Indicators*

M&S requirements quality can be assessed by employing the following indicators:

1. *M&S Requirements Accuracy* is the degree to which the requirements possess sufficient transformational (verity) and representational (validity) correctness.
 - 1.1. *M&S Requirements Verity* is assessed by conducting M&S requirements verification. M&S requirements verification is substantiating that the M&S requirements are transformed from higher levels of abstraction into their current form with sufficient accuracy judged with respect to the M&S intended uses. M&S requirements verification addresses the question of “Are we creating the M&S requirements *right*?”
 - 1.2. *M&S Requirements Validity* is assessed by conducting M&S requirements validation. M&S requirements validation is substantiating that the M&S requirements represent the *real* needs of the application sponsor with

sufficient accuracy. M&S requirements validation addresses the question of “Are we creating the *right* M&S requirements?”

2. *M&S Requirements Clarity* is the degree to which the M&S requirements are unambiguous and understandable.
 - 2.1. *M&S Requirements Unambiguity* is the degree to which each statement of the requirements can only be interpreted one way.
 - 2.2. *M&S Requirements Understandability* is the degree to which the meaning of each statement of the requirements is easily comprehended by all of its readers.
3. *M&S Requirements Completeness* is the degree to which all parts of a requirement are specified with no missing information, i.e., each requirement is self-contained. For example, “radar search pulse rate must be 10” is an incomplete requirement because it is missing the “per second” part. The requirement “missile kill assessment delay must follow the Uniform probability distribution” is incomplete because it is missing the range parameter values. Also use of the placeholder “TBD” (to be determined or to be defined), “TBR” (to be resolved), “TBP” (to be provided), and use of the phrases such as “as a minimum,” “as a maximum,” and “not limited to” are indications of incomplete requirements specification.
4. *M&S Requirements Consistency* is the degree to which (a) the requirements are specified using uniform notation, terminology, and symbology, and (b) any one requirement does not conflict with any other.
5. *M&S Requirements Feasibility* is the degree of difficulty of (a) implementing a single requirement, and (b) simultaneously meeting competing requirements. Sometimes it may be possible to achieve a requirement by itself, but it may not be possible to achieve a number of them simultaneously.
6. *M&S Requirements Modifiability* is the degree to which the requirements can easily be changed.
7. *M&S Requirements Stability* is (a) the degree to which the requirements are changing while the M&S application is under development, and (b) the possible effects of the changing requirements on the project schedule, cost, risk, quality, functionality, design, integration, and testing of the M&S application.
8. *M&S Requirements Testability* is the degree to which the requirements can easily be tested. A testable requirement is the one that is specified in such a way that pass/fail or assessment criteria can be derived from its specification. For example, the following requirement specification is not testable: “The probability of kill should be estimated based on the simulation output data.” The following requirement specification is testable: “The probability of kill should be estimated by using a 95 % confidence interval based on the simulation output data.”
9. *M&S Requirements Traceability* is the degree to which the requirements related to a particular requirement can easily be found. Requirements should be specified in such a way that related requirements are cross-referenced. When it is

necessary to change a requirement, those requirements affected by the changed requirement should be easily identified by using the cross-references.

9.3.3 *Conceptual Model Quality Indicators*

The quality of a conceptual model created for a particular problem domain can be assessed by employing the following indicators (Balci et al. 2011):

1. How well does the conceptual model assist in designing not just one simulation model but also many in a particular problem domain?
2. How well does the conceptual model assist in designing any type of simulation model?
3. How well does the conceptual model assist in achieving *reusability* in simulation model design? (Balci et al. 2011)
4. How well does the conceptual model assist in achieving *composability* in simulation model design? (Balci et al. 2011)
5. How well does the conceptual model enable effective communication among the people involved in a large-scale M&S project such as stakeholders, potential users, managers, analysts, and M&S developers?
6. How well does the conceptual model assist in overcoming the complexity of designing large-scale complex simulation models in a particular problem domain?
7. How well does the conceptual model provide a multimedia knowledge base covering the areas of expertise needed for designing large-scale complex simulation models in a particular problem domain?
8. How well does the conceptual model help a subject matter expert (SME) involved in an M&S project to understand another SME's work?
9. How well does the conceptual model facilitate the collaboration among the SMEs for designing a large-scale complex simulation model in a particular problem domain?
10. How well does the conceptual model assist in verification, validation, and certification (VV&C) of an M&S application in a particular problem domain?
11. How well does the conceptual model support effective and efficient VV&C of an M&S application in a particular problem domain?
12. How well does the conceptual model assist in the specifications of test designs, test cases, and test procedures for an M&S application in a particular problem domain?
13. How well does the conceptual model assist in proper formulation of *intended uses* (objectives) for an M&S application in a particular problem domain?
14. How well does the conceptual model assist in the generation of new M&S requirements?
15. How well does the conceptual model provide significant economic benefits through its repeated use?

9.3.4 *Architecture Quality Indicators*

The following indicators can be employed for assessing how well a specified architecture such as High Level Architecture (HLA) (IEEE 2000) enables an M&S application to possess a desired set of quality characteristics under a set of indented uses (Balci and Ormsby 2008):

1. *Adaptability* is the degree to which the architecture enables the M&S application to be easily modified to satisfy changing requirements.
2. *Compliance with standards* is the degree to which the architecture enables the M&S application to comply with required standards.
3. *Dependability* is the degree to which the architecture enables the M&S application to (a) deliver services when requested, (b) deliver services as specified, (c) operate without catastrophic failure, and (d) protect itself against accidental or deliberate intrusion.
 - 3.1. *Availability* is the degree to which the architecture enables the M&S application to function according to its requirements at a given point in time. Availability refers to the ability of the M&S application to deliver services when requested.
 - 3.2. *Reliability* is the degree to which the architecture enables the M&S application to perform its required functions without failure under prescribed conditions in a specified period of time for a specific purpose. Reliability refers to the ability of the M&S application to deliver services as specified.
 - 3.3. *Safety* is the degree to which the architecture enables the M&S application to operate, normally or abnormally, without threatening people or the environment. Safety refers to the ability of the M&S application to operate without catastrophic failure.
 - 3.4. *Security* is the degree to which the architecture enables the M&S application to provide protection and authentication of information in transit or stationary, as well as the confidentiality of sensitive information. Security refers to the ability of the M&S application to protect itself against accidental or deliberate intrusion.
4. *Deployability* is the degree to which the architecture enables the M&S application to be easily transformed to run on more than one hardware, software, or network environment.
5. *Extensibility* is the degree to which the architecture enables the M&S application to (a) be capable of growing by including more and a greater diversity of subsystems, and (b) facilitate the extension of its capabilities by modifying current features or adding new features.
6. *Interoperability* is the degree to which the architecture enables the M&S application to exchange data with other systems or subsystems and be able to use the data that has been exchanged.

7. *Maintainability* is the degree to which the architecture enables the M&S application to facilitate changes for: (a) adaptations required as the system's external environment evolves (*adaptive maintenance*), (b) fixing bugs and making corrections (*corrective maintenance*), (c) enhancements brought about by changing customer requirements (*perfective maintenance*), and (d) preventing potential problems or for reengineering (*preventive maintenance*).
8. *Modifiability* is the degree to which the architecture enables the M&S application to be easily changed.
9. *Openness* is the degree to which the architecture enables the M&S application to possess interface specifications of its components or subsystems that are fully defined, publicly available, nonproprietary, and maintained by recognized standards bodies.
10. *Performance* is the degree to which the architecture enables the M&S application to execute its work in a speedy, efficient, and productive manner.
11. *Scalability* is the degree to which the architecture enables the M&S application to continue to function correctly as its workload (e.g., number of users, size of network, and amount of processing) is increased within anticipated limits.
12. *Survivability* is the degree to which the architecture enables the M&S application to satisfy and continue to satisfy specified critical requirements (e.g., security, reliability, real-time responsiveness, and accuracy) under adverse conditions.
13. *Testability* is the degree to which the architecture enables the M&S application to facilitate the creation of test criteria and conduct tests to determine whether those criteria have been met.
14. *Usability* is the degree to which the architecture enables the M&S application to be easily employed for its intended uses.

9.3.5 *Simulation Model Design Quality Indicators*

The following indicators can be employed for assessing the quality of a simulation model design:

1. *Adaptability* is the degree to which the simulation model design can accommodate changing requirements.
2. *Complexity* is the degree to which the simulation model design can be understood without difficulty and can easily be communicated to others.
3. *Composability* is the degree to which the simulation model design is capable of being constituted by combining modules, parts, or elements.
4. *Correctness* is the degree to which the simulation model design possesses sufficient transformational, representational, and behavioral accuracy.
5. *Detailedness* is the degree to which the simulation model design possesses sufficient level of detail to enable its programming into an executable model.

6. *Efficiency* is the degree to which the simulation model design enables the M&S application to fulfill its purpose without waste of resources.
7. *Flexibility* is the degree to which the simulation model design accommodates modifications.
8. *Integrity* is the degree to which the simulation model design enables the M&S application to be capable of controlling access to sensitive information by unauthorized persons or other applications.
9. *Interoperability* is the degree to which the simulation model design enables the M&S application in a distributed environment to exchange data with other applications and to be able to use the data that has been exchanged.
10. *Maintainability* is the degree to which the simulation model design facilitates changes for: (a) adaptations required as the model's external environment evolves (*adaptive maintenance*), (b) fixing bugs and making corrections (*corrective maintenance*), (c) enhancements brought about by changing customer requirements (*perfective maintenance*), and (d) preventing potential problems or for reengineering (*preventive maintenance*).
11. *Modularity* is the degree to which the simulation model design has the highest level of cohesion and the lowest level of coupling.
12. *Cohesion* is the degree to which the elements included within a simulation model design component are highly related to each other.
13. *Coupling* is the degree to which the simulation model design components depend on each other in terms of their internal logic.
14. *Portability* is the degree to which the simulation model design can easily be transformed to enable the M&S application to run on more than one hardware or software platform.
15. *Reusability* is the degree to which the simulation model design facilitates the reuse of its components in the development of other simulation model designs.
16. *Testability* is the degree to which the simulation model design facilitates the creation of test criteria and conducting tests to determine whether those criteria have been met.

9.3.6 M&S Application Quality Indicators

M&S application quality can be assessed by employing the following indicators (Balci 2004; Pressman 2010; Sommerville 2011):

1. *M&S Application Dependability* is the degree to which the M&S application (a) delivers services when requested, (b) delivers services as specified, (c) operates without catastrophic failure, and (d) protects itself against accidental or deliberate intrusion.
 - 1.1. *M&S Application Availability* is the probability that the M&S application functions according to its requirements at a given point in time. Availability

refers to the ability of the M&S application to deliver services when requested.

- 1.2. *M&S Application Reliability* is the degree to which the M&S application performs its required functions without failure under prescribed conditions in a specified period of time for a specific purpose. M&S application reliability refers to the ability of the M&S application to deliver services as specified.
 - 1.2.1. *M&S Application Accuracy* is the degree to which the M&S application possesses sufficient transformational and representational/behavioral accuracy.
 - 1.2.1.1. *M&S Application Verity* is assessed by conducting M&S application verification, which is substantiating that the M&S application is transformed from one form into another with sufficient accuracy. M&S application verification addresses the question of “Are we building the M&S application *right*?”
 - 1.2.1.2. *M&S Application Validity* is assessed by conducting M&S application validation, which is substantiating that the M&S application possesses sufficient representational and behavioral accuracy. M&S application validation addresses the question of “Are we building the *right* M&S application?”
 - 1.2.2. *M&S Application Mean Time to Failure* (MTTF) is the average time between observed M&S application failures. MTTF = 300 h means that, on the average, one failure can be expected to occur every 300 h.
 - 1.2.3. *M&S Application Mean Time to Restore* (MTTR) is the average time it takes to restore the M&S application after failure.
 - 1.2.4. *M&S Application Recoverability* is the degree to which the M&S application provides mechanisms to enable users to recover from errors.
 - 1.3. *M&S Application Safety* is the ability of the M&S application to operate, normally or abnormally, without threatening people or the environment. M&S safety refers to the ability of the M&S application to operate without catastrophic failure. The safety may be an issue particularly for training simulations.
 - 1.4. *M&S Application Security* is the degree to which the M&S application provides protection and authentication of information in transit or stationary, as well as the confidentiality of sensitive information. M&S security refers to the ability of the M&S application to protect itself against accidental or deliberate intrusion.
2. *M&S Application Functionality* is the degree to which the M&S application completely captures all of the desired functional modules that need to be present.

- 2.1. *M&S Application Capabilities* is the degree to which the M&S application is capable of performing its feature set, e.g., capability of simulating a particular combat at the soldier level of granularity.
- 2.2. *M&S Application Detailedness* is the degree to which the M&S application is characterized by abundant use of detail or thoroughness of treatment.
- 2.3. *M&S Application Feature Set* is the degree to which the M&S application provides the set of features that need to be present, e.g., simulating a particular combat.
- 2.4. *M&S Application Generality* is the degree to which the M&S application can be used for a wide range of intended uses.
3. *M&S Application Performance* is the degree to which the M&S application executes its work in a speedy, efficient, and productive manner.
 - 3.1. *M&S Application Algorithmic Efficiency* is the degree to which the algorithms used in the M&S application provide the optimal execution time.
 - 3.2. *M&S Application Architectural Efficiency* is the degree to which the M&S application architecture enables the optimal execution time.
 - 3.3. *M&S Application Communication Efficiency* is the degree to which the M&S application fulfills its purpose of communicating with its user over a network without waste of resources. Communication efficiency is influenced by the communication protocol (e.g., HTTP or RMI) used by the M&S application, encryption/decryption of the communication, or the existence of a firewall.
 - 3.4. *M&S Application Resource Use Efficiency* is the degree to which the M&S application fulfills its purpose without waste of resources such as CPU, main memory, and hard disk space.
4. *M&S Application Supportability* is the degree to which the M&S application can be supported.
 - 4.1. *M&S Application Compatibility* is the degree to which the M&S application can be integrated into or used with other M&S applications, products, or systems.
 - 4.2. *M&S Application Configurability* is the degree to which the M&S application can easily be set up or configured for a particular application or intended use.
 - 4.3. *M&S Application Conformity* is the degree to which the M&S application adheres to standards and conventions.
 - 4.4. *M&S Application Installability* is the degree to which the M&S application can easily be prepared for use.
 - 4.5. *M&S Application Interoperability* is the degree to which the M&S application in a distributed environment (e.g., federation of models) can exchange data with one or more other M&S applications and be able to use the data that has been exchanged.
 - 4.6. *M&S Application Localizability* is the degree to which the M&S application can easily be adopted, preferably via preferences or options, (a) to satisfy

- the needs of languages other than English, and (b) to local standards such as decimal separator, currency symbol, time zone, calendar, etc.
- 4.7. *M&S Application Maintainability* is the degree to which the M&S application facilitates changes for: (a) adaptations required as the M&S application's external environment evolves (*adaptive maintenance*), (b) fixing bugs and making corrections (*corrective maintenance*), (c) enhancements brought about by changing customer requirements (*perfective maintenance*), and (d) preventing potential problems or for reengineering (preventive maintenance or software reengineering).
 - 4.7.1. *M&S Application Adaptability* is the degree to which the M&S application can accommodate changes to its external environment.
 - 4.7.2. *M&S Application Correctability* is the degree to which the M&S application facilitates changes for fixing bugs and making corrections.
 - 4.7.3. *M&S Application Extensibility* is the degree to which the M&S application capabilities can be extended by modifying current features or adding new features.
 - 4.7.4. *M&S Application Preventability* is the degree to which the M&S application facilitates changes for preventing potential problems or for reengineering.
 - 4.8. *M&S Application Portability* is the degree to which the M&S application can easily be transformed to run on more than one hardware or software platform.
 - 4.9. *M&S Application Testability* is the degree to which the M&S application facilitates the creation of test criteria and conducting tests to determine whether those criteria have been met.
 5. *M&S Application Usability* is the degree to which the M&S application can easily be employed for its intended use.
 - 5.1. *M&S Application Documentation Quality* is the degree to which the M&S application external documentation (e.g., user manuals, reference guides, online help) possesses a desired set of characteristics.
 - 5.2. *Ease of Experimentation or Exercise Specification* is the degree to which a simulation experiment (for analysis) or a simulation exercise (for training) can easily be specified.
 - 5.3. *Ease of M&S Application Input Specification* is the degree to which the input conditions and data of the M&S application are easily specified under a set of prescribed intended uses.
 - 5.4. *M&S Application Ease of Learning* is the ease with which the M&S application can be learned.
 - 5.5. *M&S Application Output Understandability* is the degree to which the meaning of the M&S application output is easily comprehended by its users under a set of prescribed intended uses.

9.3.7 *Simulation Results Quality Indicators*

The simulation results make up the solution to the problem (for problem solving), show effectiveness of simulation-based training (for training purposes), or indicate some benefit in using the simulation model (e.g., for research). The quality of the results obtained from a simulation model by way of experimentation (for problem solving), exercise (for training purposes) or otherwise use can be assessed by employing the following indicators:

1. How reliable is the random number generator used as judged by the community?
2. How theoretically accurate are the algorithms used for random variate generation?
3. How accurately are the random variate generation algorithms translated into executable code?
4. How well are the simulation experiments designed to gather the desired information at minimal cost and to enable the analyst to draw valid inferences?
5. How accurately are the designs of experiments translated into executable code?
6. How appropriate are the statistical techniques used for the analysis of simulation output data?
7. How well are the assumptions underlying the statistical techniques used satisfied?
8. How appropriately is the problem of the initial transient (or the start-up problem) addressed?
9. How correctly are identical experimental conditions created for each of the alternative operating policies compared?

9.3.8 *Presented Results Quality Indicators*

The life cycle process of *presentation* consists of (a) interpretation of the simulation results, (b) documentation of the simulation results, and (c) communication of the simulation results to the decision makers. Simulation results must be interpreted because all simulation models are descriptive in nature. A *descriptive model* is a model that describes the behavior of a system without any value judgment on the “goodness” or “badness” of such behavior. For example, a simulation result can be “average waiting time is 5 minutes” without indicating how good or bad the value 5 is. That value must be interpreted and judged before presenting it to the decision makers. Due to the complexity of some simulation results, failing to properly interpret, document, and communicate the simulation results may lead to wrong decisions in spite of the fact that the simulation results are sufficiently credible.

The quality of the presented results can be assessed by employing the following indicators:

1. How accurately are the simulation results interpreted?
2. How properly are the simulation results documented?

3. How correctly are the simulation results communicated to the decision makers?
4. How accurately are the simulation results converted from the technical jargon into a language the decision makers can understand?
5. How accurately are the simulation output data transformed into visualizations, spreadsheets, tabulations, and/or graphical representations for effective presentation?

9.4 Concluding Remarks

Accuracy undoubtedly stands out to be the most important quality indicator of an M&S application. It is assessed by conducting verification, validation, and testing (VV&T) (Balci 2003). Tremendous amount of literature exists on VV&T. More than 75 VV&T techniques have been described in the published literature (Balci 1998).

Assessment of accuracy alone, however, is not sufficient for judging the acceptability of a large-scale and complex M&S application. An M&S application may be sufficiently accurate, but it may not satisfy other quality indicators such as the ones described in this chapter. Gaining an acceptable level of confidence in the accuracy of a large-scale and complex M&S application may not be feasible due to the complexity. Therefore, a total quality-centered assessment approach should be used to gain sufficient level of confidence in certifying the acceptability of a large-scale and complex M&S application.

Quality assessment activities must be tied to a well-structured M&S life cycle (Balci 2012). Quality assessment is not a stage but a continuous activity carried out hand in hand throughout the entire M&S life cycle. The use of an effective M&S life cycle is critically important for success in gaining sufficient confidence in M&S application acceptability.

References

- Balci O (1998) Chapter 10. Verification, validation, and testing. In: Banks J (ed) *The handbook of simulation*. Wiley, New York, pp 335–393
- Balci O (2003) Verification, validation, and certification of modeling and simulation applications. In: *Proceedings of the 2003 Winter Simulation Conference* (New Orleans, LA, Dec. 7–10). IEEE, Piscataway, pp 150–158
- Balci O (2004) Quality assessment, verification, and validation of modeling and simulation applications. In: *Proceedings of the 2004 Winter Simulation Conference* (Washington, DC, Dec. 5–8). IEEE, Piscataway, pp 122–129
- Balci O (2012) A life cycle for modeling and simulation. *Simulation Trans Soc Model Simul Int* 88 (7):870–883
- Balci O, Ormsby WF (2008) Network-centric military system architecture assessment methodology. *Int J Syst Syst Eng* 1(1–2):271–292

- Balci O, Sargent RG (1981) A methodology for cost-risk analysis in the statistical validation of simulation models. *Commun ACM* 24(4):190–197
- Balci O, Nance RE, Arthur JD, Ormsby WF (2002) Expanding our horizons in verification, validation, and accreditation research and practice. In: *Proceedings of the 2002 Winter Simulation Conference* (San Diego, CA, Dec. 8–11). IEEE, Piscataway, pp 653–663
- Balci O, Arthur JD, Ormsby WF (2011) Achieving reusability and composability with a simulation conceptual model. *J Simul* 5(3):157–165
- IEEE (2000) IEEE standard for modeling and simulation (M&S) high level architecture (HLA) – framework and rules, IEEE Standard No. 1516–2000. IEEE, Piscataway
- Ören TI (1981) Concepts and criteria to assess acceptability of simulation studies: a frame of reference. *Commun ACM* 24(4):180–189
- Pressman RS (2010) *Software engineering: a practitioner’s approach*, 7th edn. McGraw-Hill, New York
- Sommerville I (2011) *Software engineering*, 9th edn. Addison-Wesley, Reading